

**Selectie bij overgangen in het
onderwijs**

Een beknopte literatuurstudie

Cor Sluiter

Inhoudsopgave

Voorwoord	7
1 Inleiding	9
2 Een methodologische verkenning van overgangsbepalingen in het onderwijs.....	9
2.1 Het vinden van een goede balans tussen het geven en ontzeggen van kansen	9
2.2 Voorspellen bij overgangsbepalingen.....	13
2.3 Kwaliteitseisen voor bij overgangsbepalingen in te zetten procedures	14
2.4 Het subjectieve karakter van overgangsbepalingen	18
2.5 Doorlopende leerlijnen	19
2.6 Tot slot.....	20
3 De overgang naar het vo	21
3.1 Stand van zaken	21
3.2 Recente ontwikkelingen	24
3.3 Evaluatie	26
4 De overgang naar het mbo	29
4.1 Stand van zaken	29
4.2 Recente ontwikkelingen	31
4.3 Evaluatie	31
5 De overgang naar hbo en wo	33
5.1 Stand van zaken	33
5.2 Recente ontwikkelingen	36
5.3 Evaluatie	38
6 De overgang van bachelor naar masterstudie	41
7 Slotbeschouwing	43
Literatuur	47

'Tests derive their worth from their measurement and decision-making qualities. These qualities cannot be assumed by fiat, but have to be based on theoretical and empirical research. Analogously to the health science concept of evidence-based treatment, it is proposed that evidence-based testing be used in the behavioral sciences. Evidence-based testing is testing that is justified by sound research.'

(Mellenbergh, 2011, 455-456)

Voorwoord

In het Werkprogramma voor 2013 van de Onderwijsraad is het adviesonderwerp 'Overgangen in het onderwijs' opgenomen. De raad heeft in dit kader bijzondere aandacht voor vormen van selectie die worden toegepast bij de overgangen in het stelsel. Hij heeft namelijk aanwijzingen dat er in de afgelopen periode een toename is geweest van de inzet van allerlei vormen van selectie bij overgangen in het onderwijs. Instellingen investeren steeds meer in procedures die er in hun ogen toe kunnen bijdragen dat leerlingen, deelnemers of studenten zoveel mogelijk in een opleiding terecht komen die in de praktijk ook geschikt blijkt te zijn. En blijkbaar zien zij selectie als een middel om de kans op uitval te verminderen. De vraag die zich hierbij voordoet is of het middel niet erger is dan de kwaal. Selectiemechanismen kunnen er immers ook toe leiden dat de toegankelijkheid van het stelsel in gevaar komt, omdat bepaalde groepen leerlingen door steeds strengere en vroegere selectie minder onderwijskansen krijgen. Om gefundeerde uitspraken te kunnen doen over deze ontwikkelingen wilde de Onderwijsraad graag in kaart gebracht zien wat de mogelijkheden en beperkingen zijn van selectiemechanismen bij overgangen in het onderwijs.

Daarom is aan Cito het verzoek gedaan een beknopte literatuurstudie uit te voeren waarin de empirische stand van zaken op dit gebied wordt weergegeven. Deze literatuurstudie zou een beeld moeten geven van de voorspellende waarde van verschillende selectiemethoden die kunnen worden gehanteerd. De raad is geïnteresseerd in de betrouwbaarheid en de bruikbaarheid van instrumenten zoals bestaande examens en toetsen, specifieke toelatingstests en motivatiegesprekken. Daarbij gaat het in principe om selectiemechanismen bij alle overgangen in het onderwijsstelsel, dat wil zeggen de overgangen tussen primair en voortgezet onderwijs; de overgang van voortgezet onderwijs naar middelbaar beroepsonderwijs; de overgang van voortgezet onderwijs en middelbaar beroepsonderwijs naar het hoger beroepsonderwijs; de overgang van voortgezet onderwijs en hoger beroepsonderwijs naar het wetenschappelijk onderwijs en de overgang van de bachelor- naar de masterstudie in het hoger onderwijs.

Deze notitie bevat de resultaten van de uitgevoerde literatuurstudie rond het voornoemde thema. Daarbij is naast de empirische stand van zaken ook de nodige aandacht besteed aan het maken van enkele kanttekeningen die betrekking hebben op verschillende voetangels en klemmen waar men mee geconfronteerd wordt bij het doen van onderzoek naar de bruikbaarheid van selectiemechanismen en het ontwikkelen van selectie-instrumenten. Dit laatste om de Onderwijsraad een bruikbaar referentiekader te geven voor het interpreteren van de stand van zaken. Cito hoopt dat deze studie de Onderwijsraad van dienst kan zijn bij het verder vormgeven van de visie op het thema selectie.

Bij het uitvoeren van deze studie ben ik ondersteund door het Kenniscentrum van Cito. Mijn dank daarvoor gaat uit naar mijn collega's Anneke van Roosmalen en Astrid Segers.

Arnhem, januari 2013
Cor Sluijter

1 Inleiding

De rendementsvraag in het onderwijs staat weer volop in de belangstelling. De feiten lijken voor zich te spreken: een fors percentage van de leerlingen in het voortgezet onderwijs, de deelnemers in het mbo en de studenten in hoger onderwijs behaalt het diploma bij de oorspronkelijk beoogde opleiding niet. Bovendien verloopt de weg naar het einddiploma lang niet voor iedereen via de kortste route en/of binnen de normale opleidingstermijn. De nodige leerlingen, deelnemers en studenten lopen tijdens hun schoolloopbaan vertraging op, of stromen op of af naar een andere opleiding. De meest recente cijfers van het Centraal Bureau voor de Statistiek (CBS; 2012) spreken voor zich.

In het schooljaar 2010/2011 verliet volgens het CBS 3,2 procent van de leerlingen onder de drieëntwintig jaar het vo of mbo zonder startkwalificatie: een diploma op minimaal havo-, vwo- of mbo-2-niveau. Binnen het voortgezet onderwijs loopt het percentage vroegtijdig schoolverlaters uiteen van 0,7 procent in het derde en vierde leerjaar van de havo en het derde, vierde en vijfde leerjaar van het vwo tot 9,8 procent in de eindexamenklas van het vmbo. Binnen het mbo ligt de uitval zonder diploma op ongeveer achttien procent. Bijna drie op de tien vroegtijdig schoolverlaters keren binnen zeven jaar terug in het onderwijs. Ruim een kwart van de vroegtijdig schoolverlaters uit het vo en één op de vijf uit het mbo haalt uiteindelijk alsnog een startkwalificatie.

Verder meldt het CBS dat ongeveer één op de zes leerlingen in vmbo-gt, havo en vwo tijdens de opleiding afstroomt naar een lager niveau dan het aanvangsniveau. Een iets groter deel stroomt daarentegen vanuit de verschillende niveaus in het vmbo en vanuit de havo op naar een hoger niveau. Van de vmbo'ers die niet van onderwijssoort veranderen loopt ongeveer een kwart vertraging op door zittenblijven. Van de havisten is dat zelfs bijna veertig procent. En bij de vwo'ers bedraagt dit percentage eveneens ongeveer vijfentwintig procent.

De analyses van het CBS maken daarnaast duidelijk dat vmbo'ers na het behalen van het diploma meestal doorstromen naar het mbo, havisten naar het hbo en vwo'ers naar het wo. Van de geslaagden voor een mbo-4-opleiding stroomt ongeveer twee vijfde door naar het hbo. Van de eerstejaars in het mbo is na vijf jaar driekwart geslaagd. Voor de meeste opleidingen in het hoger beroepsonderwijs staat een studieduur van vier jaar. In de praktijk rondt ongeveer vijfenvijftig procent van de studenten een voltijdstudie binnen vijf jaar af. Eén op de zeven studenten haakt al in het eerste jaar af. Aan het begin van het zesde jaar staat nog ruim een kwart ingeschreven en heeft ongeveer één op de vijf studenten het hoger onderwijs zonder diploma verlaten. In het wo heeft na zeven jaar iets meer dan de helft van de voltijdstudenten met vwo als vooropleiding een masterdiploma behaald. Ongeveer tweevijfde is dan nog bezig met zijn studie en zeven procent is gestopt met de opleiding.

Dit alles lijkt te pleiten voor het inzetten van middelen voor het verhogen van het rendement. Selectie lijkt één van die middelen te zijn. Verschillende politici, opleidingsmanagers, onderwijskundigen en andere stakeholders hebben zich daarom met enthousiasme op dit thema gestort. Illustratief in dit kader is een opmerking van het CBS (2012, p. 86): "Zowel bij havo als vwo

lagen de slagingspercentages vijf jaar eerder nog 4 procentpunten hoger dan in 2010/'11. Wellicht was de keuze voor havo of vwo de afgelopen paar jaar voor enkelen te hoog gegrepen en had een aantal vwo'ers meer succes op de havo gehad en was een aantal havisten beter af geweest op het vmbo".

Met deze laatste opmerking roept het CBS impliciet op tot grotere strengheid bij het nemen van overgangsbeslissingen voor en in het voortgezet onderwijs. In hoeverre het gehoor geven aan een dergelijke oproep voor verschillende onderwijsvormen verstandig is en welke mogelijkheden er zijn voor het verbeteren van overgangsbeslissingen is het onderwerp van deze notitie.

Een deel van de leerlingen, deelnemers en studenten in het laatste leerjaren van het vo, mbo of ho is daar, zoals boven aangegeven, niet zonder vertraging en omwegen beland. Naast tussentijdse op-, af- en doorstroom is ook sprake van het 'stapelen' van diploma's. In de praktijk blijkt het behaald hebben van een vmbo-diploma tegenwoordig niet meer te leiden tot directe toelating tot 4-havo. Ook het behaald hebben van het havo-diploma leidt niet meer direct tot toelating tot 5-vwo. Doorstromende leerlingen blijken namelijk in de praktijk minder succesvol te zijn dan leerlingen die de betreffende vorm van onderwijs al volgden. Zie voor de overgang van vmbo naar havo bijvoorbeeld Neuvel, van Esch en Westerhuis (2011) Om uitval zoveel mogelijk te voorkomen is er voor ontvangende scholen beleidsruimte om aanvullende eisen te stellen. Formeel is er dus sprake van selectie. Om willekeur bij deze selectie te voorkomen en transparantie te bevorderen heeft de VO-raad voor de overgang van vmbo naar havo een toelatingscode ontwikkeld die duidelijkheid geeft over de aanvullende eisen die scholen stellen en aanzet tot verbetering van de aansluiting en de samenwerking tussen vmbo. Voor de concrete inhoud hiervan zij verwezen naar het Ministerie van OCW (2011). Voor de overgang van havo naar vwo zijn (nog?) geen concrete afspraken gemaakt. Dit document laat om redenen van beknoptheid en overzichtelijkheid overgangsbeslissingen binnen het vo verder buiten beschouwing. Ditzelfde geldt voor overgangsbeslissingen binnen de overige onderwijssectoren met uitzondering van de overgang van de bachelor- naar de masterstudie.

2

Een methodologische verkenning van overgangsbepalingen in het onderwijs

Selectie is van alle tijden. Zo bestond in China tijdens de Sui-dynastie, rond het jaar 600, al een examensysteem voor het selecteren van ambtenaren. (Têng, 1943). Omdat er concreet bewijs is voor de validiteit van de gebruikte toelatingsexamens (Suen & Wu, 2006) en er van iedere honderd kandidaten maar vijf werden aangenomen (Miyazaki, 1981), mogen we aannemen dat de kandidaten die het examen behaalden vervolgens ook hun functie naar tevredenheid zullen hebben uitgevoerd. Waarschijnlijk zal slechts bij hoge uitzondering een ambtenaar niet goed gefunctioneerd hebben bij het uitoefenen van zijn ambt. Dit historische voorbeeld lijkt te pleiten voor het inzetten van selectiemechanismen in het onderwijs. Het lijkt eenvoudig: selecteer strenger en het rendement gaat omhoog. Dergelijke berichten zijn ook in de media terug te vinden. Zie bijvoorbeeld een artikel in NRC (22, februari 2011) dat ingaat op het feit dat bij 'tien opleidingen die op dit moment scholieren mogen keuren alvorens ze toe te laten' de rendementscijfers aanzienlijk hoger zijn dan het landelijk gemiddelde. Het is echter gevaarlijk om uitsluitend op basis van in isolatie waargenomen effecten te besluiten dat selectie hét middel bij uitstek is om het rendement van een geheel onderwijsstelsel te verhogen. De destijds door de Chinese overheid gehanteerde selectieratio zal er toe geleid hebben dat een hoog percentage personen ten onrechte werd afgewezen. Dat was destijds geen probleem, omdat het vooral ging om het goed laten functioneren van de Chinese overheid. Het probleem waar de Nederlandse overheid mee geconfronteerd wordt, is van een andere aard dan in het oude China: het gaat erom het aanwezige potentieel binnen de Nederlandse bevolking optimaal te benutten. Dat betekent dat er bij alle overgangsbepalingen een goede balans gevonden moet worden tussen het geven en ontzeggen van kansen aan leerlingen, deelnemers en studenten.

En het vinden van een goede balans tussen het geven en ontzeggen van kansen is niet de enige uitdaging. Dat is namelijk pas mogelijk als er op het moment van de overgangsbepaling informatie ter beschikking staat die voldoende sterk samenhangt met toekomstige prestaties in de onderwijsvormen waartussen de keuze is. Verder kan er nog een reeks andere generieke kanttekeningen gemaakt worden bij het fenomeen overgangsbepaling en bij de instrumenten die daarbij in te zetten zijn. Deze drie zaken komen in dit hoofdstuk aan de orde. In de daaropvolgende hoofdstukken wordt vervolgens meer specifieke aandacht gegeven aan de stand van zaken bij de overgangsbepalingen die het Nederlands onderwijsstelsel kent. Het betreft respectievelijk de overgang naar het vo; de overgang naar het mbo; de overgang naar hbo en wo en de overgang van bachelor- naar masteropleiding. Hoofdstuk 7, ten slotte, is gereserveerd voor een eindbeschouwing.

2.1. Het vinden van een goede balans tussen het geven en ontzeggen van kansen

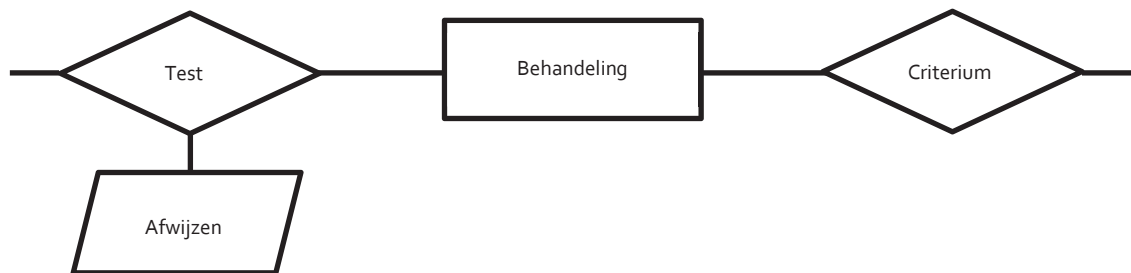
Omdat het voorspellen van de toekomstige leerprestaties van leerlingen slechts tot op zekere hoogte mogelijk is, zullen zich altijd fouten voordoen bij het nemen van overgangsbepalingen. De

informatie van het CBS is dus zeker niet verrassend; uitval, vertraging, op- en afstroom zullen zich altijd voor blijven doen. De vraag is alleen of en hoe het mogelijk is verbetering aan te brengen.

Voor een conceptuele beschrijving van overgangsbepalingen in het onderwijs is het paradigma van de psychometrische besliskunde (Cronbach en Gleser, 1965) geschikt. Het algemene uitgangspunt is dat de beslisser of de beslissende instantie¹ de keuze heeft uit minstens twee acties. Iedere actie heeft een aantal uitkomsten. Over de uitkomst van de acties waaruit de beslisser kan kiezen, bestaat onzekerheid, omdat de informatie waarover men beschikt op het moment van beslissen niet meer dan een feilbare indicatie geeft van het toekomstig gedrag van personen dat relevant is voor de bepaling. In het bijbehorende begrippenkader was oorspronkelijk sprake van drie verschillende basis typen bepalingen; namelijk selectie, plaatsing en classificatie. Later is de beheersingsbepaling daar door Van der Linden (1985) aan toegevoegd. In dit artikel en een ander (Van der Linden, 1990) worden de verschillende basistypen bepalingen beschreven op grond van volgende drie elementen:

- de 'test' op basis waarvan de bepaling genomen wordt;
- de 'behandeling(en)' waaruit gekozen moet worden;
- het criterium met behulp waarvan het welslagen van de behandeling(en) te meten is.

Van der Linden (1990) merkt bij deze elementen op dat ze generiek van karakter zijn. De term 'test' kan verwijzen naar het resultaat op een enkele test of toets, maar ook op een test- of toetsbatterij, een reeks achtergrondgegevens, of welke informatie dan ook. Zo lang er maar sprake is van een verband met het criterium. Ditzelfde geldt voor de term 'criterium'; deze kan betrekking hebben op (een) uiteenlopende (set van) criteria, zoals resultaten op een examen, prestaties na het verlaten van het onderwijs, het gemiddelde cijfer in het eerste jaar van een onderwijsvorm, het al dan niet overgaan naar het volgende leer- of studiejaar, het al dan niet behalen van een diploma, het al dan niet krijgen van een negatief bindend studieadvies, etcetera.

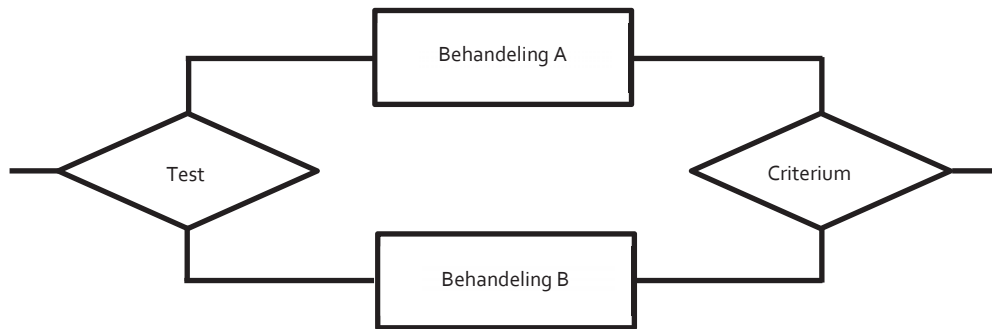


Figuur 1; Stroomdiagram van een selectiebepaling (van der Linden, 1990)

Bij een selectiebepaling gaat het in de psychometrische besliskunde om het selecteren van personen voor een bepaalde behandeling op grond van hun testresultaten. Een voorbeeld van een selectiebepaling in het onderwijs is het al dan niet aannemen van personen voor een bepaalde opleiding op grond van hun resultaten op een toelatingsexamen. Een selectiebepaling is met behulp van een stroomdiagram als hierboven (Figuur 1) weer te geven.

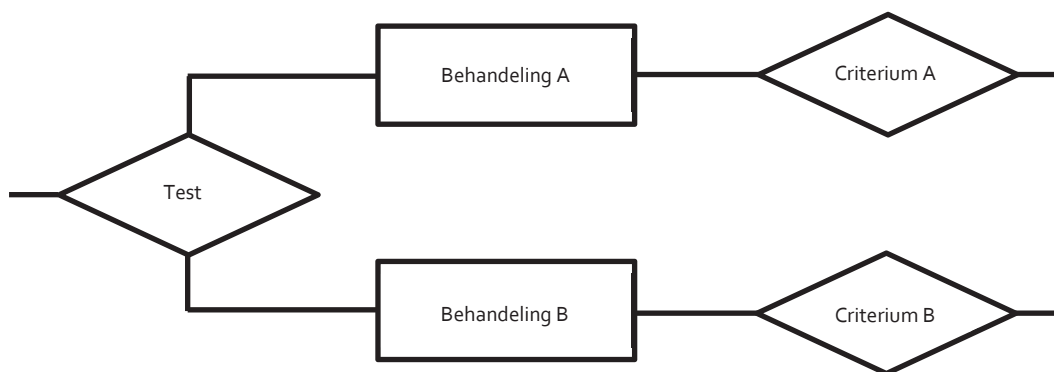
¹ In het vervolg van deze tekst zal uitsluitend nog de term 'beslisser' gehanteerd worden. Bedoeld wordt echter altijd de persoon of groep personen die de bepaling formeel neemt.

Bij een plaatsingsbeslissing worden personen aan kwalitatief verschillende behandelingen toegewezen op grond van hun testresultaten en stelt men het welslagen van iedere behandeling vast met behulp van één en hetzelfde criterium. Van een plaatsingsbeslissing in het onderwijs is bijvoorbeeld sprake, wanneer men besluit leerlingen op grond van hun prestaties op een toets individueel of klassikaal onderwijs te laten volgen en na afronding aan alle leerlingen hetzelfde afsluitende examen geeft. Een dergelijk type beslissing is in een stroomdiagram als volgt (Figuur 2) af te beelden.



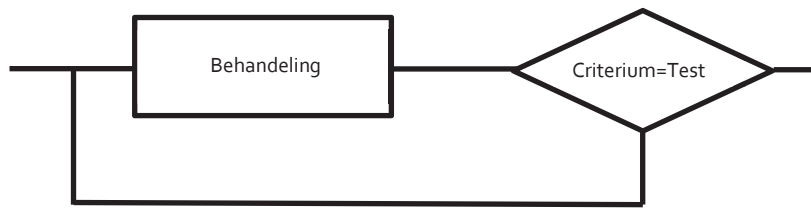
Figuur 2; Stroomdiagram van een plaatsingsbeslissing (van der Linden, 1990)

Bij een classificatiebeslissing wijst men eveneens personen toe aan kwalitatief verschillende behandelingen op grond van hun testresultaten. Alleen wordt bij een classificatiebeslissing het succes van de onderscheiden behandelingen vastgesteld met behulp van verschillende criteria. Het meest bekende voorbeeld van een classificatiebeslissing in het onderwijs treffen we aan bij de overgang van het primair naar het voortgezet onderwijs. Leerlingen worden dan mede op basis van de resultaten op een toets of test (veelal de Eindtoets Basisonderwijs van Cito) ondergebracht in verschillende vormen van voortgezet onderwijs die verschillen van niveau en met verschillende examens worden afgerond. De onderstaande figuur bevat een weergave van een classificatiebeslissing met behulp van een stroomdiagram.



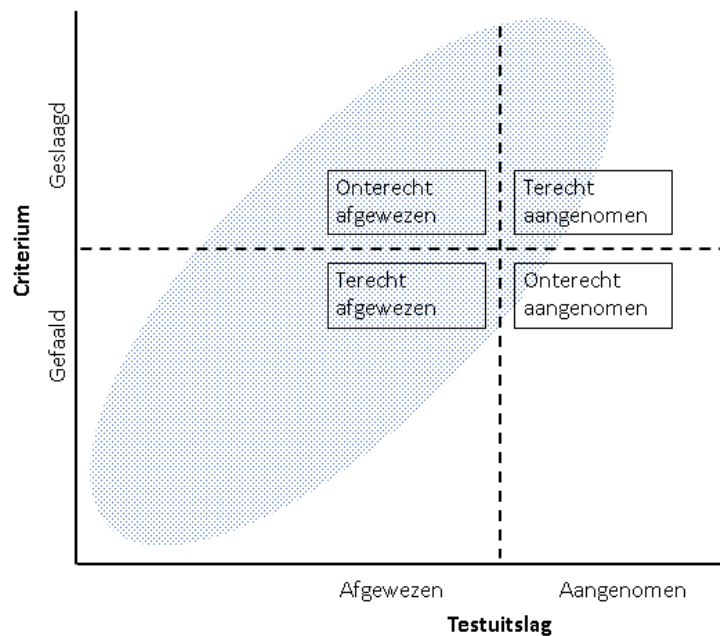
Figuur 3; Stroomdiagram van een classificatiebeslissing (van der Linden, 1990)

Bij de beheersingsbeslissing, tot slot, wordt met behulp van een test vastgesteld of een behandeling voldoende effect heeft gesorteerd. Test en criterium vallen samen, in die zin dat de testscore een met meetfouten behepte weergave van de criteriumscore is. Het criterium bij de beheersingsbeslissing is vanuit het standpunt van de klassieke testtheorie gezien de ware score op de test en vanuit het perspectief van de itemresponstheorie de latente vaardigheid. Feitelijk is in alle situaties in het onderwijs waar sprake is van een opleiding met een afsluitend examen sprake van beheersingsbeslissingen. De onderstaande figuur bevat een stroomdiagram dat laat zien welke vorm een beheersingsbeslissing heeft.



Figuur 4; Stroomdiagram van een classificatiebeslissing (van der Linden, 1990)

Zoals eerder al aangegeven, heeft de beslisser de keuze uit minstens twee acties. In het geval van de selectiebeslissing is dat aannemen of afwijzen. Over de uitkomst van de acties waaruit de beslisser kan kiezen, bestaat onzekerheid, omdat de informatie waarover men beschikt op het moment van beslissen niet meer dan een feilbare indicatie geeft van de prestaties van personen op het criterium. Bij de selectiebeslissing bijvoorbeeld, kan een kandidaat terecht afgewezen of aangenomen worden, of onterecht afgewezen of aangenomen. Figuur 5 geeft de uitkomstenruimte bij een selectiebeslissing grafisch weer. De vorm van de afgebeelde puntenwolk laat zien dat er in dit fictieve geval een relatief sterke positieve samenhang bestaat tussen de informatie die bij selectie gebruikt wordt en het uiteindelijke studiesucces. De verticale stippellijn maakt duidelijk dat in deze situatie een relatief groot deel van de kandidaten wordt afgewezen. De horizontale stippellijn geeft aan dat van de aangenomen kandidaten verreweg het grootste deel de studie met succes afrondt. Maar zelfs in deze gunstige fictieve situatie is er naast een kleine groep onterecht aangenomen kandidaten een veel minder kleine groep onterecht afgewezen kandidaten.



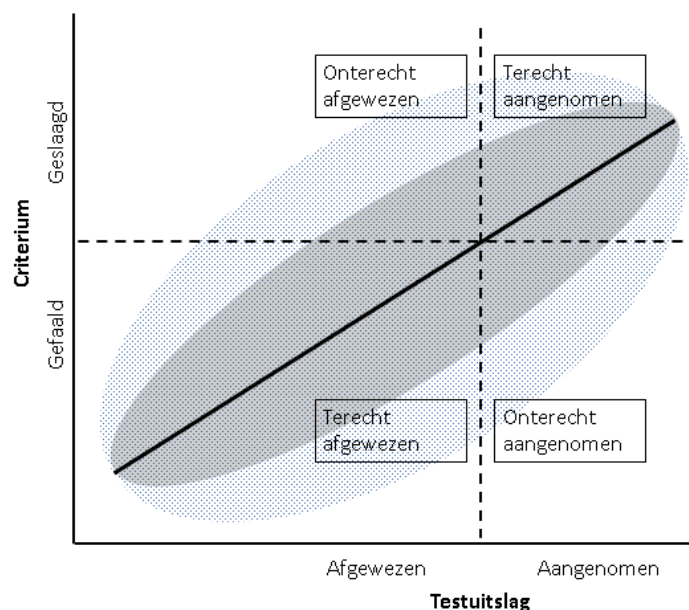
Figuur 5; Beslissinguitkomsten bij selectie

De figuur maakt duidelijk dat er in de praktijk altijd sprake zal zijn van een percentage leerlingen, deelnemers of studenten over welke een verkeerde beslissing genomen wordt. Hoe groot dat percentage is, hangt af van de sterkte van de samenhang tussen test en criterium. Het is afhankelijk van de strengheid of mildheid van de beslisser of er eerder sprake zal zijn van onterecht aannemen

of onterecht afwijzen. Om beslissingen te kunnen optimaliseren, is het daarom nodig de voorkeuren van de beslisser voor de mogelijke beslissingsuitkomsten te kwantificeren. Om dat te kunnen doen, maakt de psychometrische besliskunde gebruik van begrip 'utiliteit' (Von Neuman & Morgenstern, 1953). De subjectieve waarden die de beslisser, gegeven een bepaalde actie, hecht aan de scores op de nameting worden vastgelegd in een zogeheten utiliteitsfunctie. Er worden in de literatuur verschillende mathematische functies gehanteerd om utiliteitsfuncties weer te geven (zie bijvoorbeeld Hambleton & Novick, 1973; Van der Linden & Mellenbergh, 1977 en Novick & Lindley, 1978). Ook is het mogelijk om utiliteitsfuncties empirisch te bepalen (Van der Gaag, 1990). Wanneer de utiliteitsfuncties voor de mogelijke acties en de simultane kansverdeling van scores op voor- en nameting bekend zijn, kan voor iedere mogelijke grensscore² op de test de verwachte utiliteit berekend worden. Een rationele beslisser dient dan te kiezen voor de grensscore die de hoogste verwachte utiliteit oplevert.

2.2. Voorspellen bij overgangsbepalingen

Hierboven is aangegeven dat het er bij overgangsbepalingen in het onderwijs om gaat tussen het geven en ontzeggen van kansen een balans te vinden die de waarden weerspiegelt die de beslisser aan de verschillende beslissingsuitkomsten hecht. Voorwaarde is dan natuurlijk wel dat er op het moment van beslissen informatie ter beschikking staat die sterk samenhangt met toekomstige prestaties van leerlingen, deelnemers of studenten. Ieder instrument, of het nu een test of toets is, of welke procedure dan ook, moet dergelijke informatie leveren.



Figuur 6; Samenhang tussen test en criterium

Figuur 6 maakt duidelijk dat overgangsbepalingen in het onderwijs een lager percentage onterechte beslissingen zullen kennen, naarmate er een sterkere samenhang is tussen de informatie

² Ook deze term is generiek bedoeld; grensscore kan betrekking hebben op een enkele score op een toets of test, maar ook op een reeks scores op een toets- of testbatterij of een andere vorm van kwantitatieve informatie. Kortom iedere minimale combinatie van gegevens die leidt tot het aannemen van een leerling, deelnemer of student.

op basis waarvan de beslissing genomen wordt en het criterium dat de mate van welslagen van de opleiding weerspiegelt. De grootste puntenwolk representeert – net als in figuur 5 - een relatief sterke samenhang tussen test en criterium. Hoe hoger de samenhang tussen test en criterium, des te scherper zal de ellipsvorm van de puntenwolk worden (in het geval van lineaire samenhang) en hoe lager het percentage onterechte beslissingen zal kunnen zijn. In het theoretische geval dat er een perfecte samenhang is tussen test en criterium, is het aantal onjuiste beslissingen gelijk aan nul. Deze situatie wordt gerepresenteerd door de rechte lijn door het midden van de twee afgebeelde puntenwolken. Hoewel voor individuele personen uiteraard niet valt te bepalen of sprake is van een misclassificatie, is het wel mogelijk op basis van de betrouwbaarheid van een toets of een reeks toetsen het percentage misclassificaties te schatten (Sanders, 2011). Voor de centrale examens havo en vwo is dat bijvoorbeeld recentelijk concreet uitgerekend (Van Rijn, 2009; Van Rijn, Beguin & Verstralen, 2009).

Mellenbergh (2011, pp 373-420) staat uitgebreid stil bij het voorspellen van toekomstig gedrag en selectie van kandidaten. Hij beschrijft onder meer de verschillende categorieën predictoren en criteria die gebruikt kunnen worden. Bij predictoren kan sprake zijn van geobserveerde variabelen met verschillende categorieën: nominale (dichotoom, bijvoorbeeld sexe, en polytoom; bijvoorbeeld politieke voorkeur); ordinale, bijvoorbeeld hoogst behaalde diploma en continue, bijvoorbeeld leeftijd. Predictoren kunnen ook latent zijn. Dat houdt in dat ze niet te observeren zijn, maar dat er wél een schatting van hun waarde te maken is op basis van geobserveerde variabelen. De geschatte rekentaaligheid van een leerling op basis van een rekentoets is bijvoorbeeld een latente variabele. De geschatte mate van introversie van een persoon op basis van een persoonlijkheidstest is een ander voorbeeld van een latente variabele. Criteria zijn meestal geobserveerde variabelen en vaak dichotoom of ordinaal polytoom of continue van aard. Afhankelijk van de aard van predictor(en) en criterium moeten verschillende predictiemodellen ingezet worden. Een eerste belangrijke opmerking ten aanzien van bij een voorspelling in te zetten predictoren is dat deze ethisch en maatschappelijk gezien aanvaardbaar moeten zijn. Zo geeft het ethisch en juridisch gezien bijvoorbeeld geen pas om toekomstige prestaties in een baan te voorspellen op basis van etnische afkomst. Een tweede belangrijke opmerking is dat collineariteit ongewenst is: predictoren mogen niet sterk correleren. De voorspellende waarde van een reeks predictoren is namelijk hoger, naarmate ze sterker met het criterium in verband staan en conditioneel op deze samenhang onderling minder sterk met elkaar samenhangen.

Om de geldigheid van een predictiemodel vast te kunnen stellen moet kruisvalidering plaatsvinden. Bij kruisvalidering wordt een onderzoeksgroep random in tweeën gesplitst. In de ene helft wordt de samenhang tussen predictor of een set predictoren en een criterium vastgesteld op basis van verzamelde gegevens. In de andere helft wordt het geschatte predictiemodel gebruikt om de scores te voorspellen van de overige groepsleden op basis van hun waargenomen scores op de predictoren. Dat biedt de gelegenheid om de voorspelde scores op het criterium in deze groep te vergelijken met hun werkelijke scores en zo de kwaliteit van het predictiemodel te bepalen.

2.3. Kwaliteitseisen voor bij overgangsbepalingen in te zetten procedures

Tests en toetsen kunnen informatie leveren die samenhangt met toekomstige prestaties van leerlingen, deelnemers en studenten. Een test is een gestandaardiseerde en systematische

meetprocedure die het mogelijk maakt kwantitatieve uitspraken te doen over eigenschappen van personen met het doel om betere beslissingen ten aanzien van die personen te nemen (Sluijter, 1998). Merk op dat dit een brede definitie is, waar bijvoorbeeld ook het beoordelen via intakegesprekken of motivatiebrieven, het afnemen van specifieke toelatingsexamens, etcetera, valt. Ook het begrip toets heeft een plaats binnen deze definitie. Dat is namelijk een verbijzondering van het begrip test. De term toets is gereserveerd voor instrumenten die door onderwijs, zelfstudie of praktische ervaring verworven kennis, inzicht, vaardigheden of competenties meten. Tests kunnen bij juist gebruik een belangrijke bijdrage leveren aan de kwaliteit van beslissingen over personen: *'... the proper use of well-constructed and validated tests provides a better basis for making some important decisions about individuals ... than would otherwise be available.'* (American Psychological Association, American Educational Research Association & National Council on Measurement in Education, 1985). De gedachte achter het gebruik van tests is dat het nemen van beslissingen over personen met de grootst mogelijke zorgvuldigheid dient te gebeuren, zowel in het belang van de persoon of instantie die de beslissing neemt, als in het belang van de persoon of personen op wie de beslissing betrekking heeft.

Volgens Sluijter (1998) kunnen tests de kwaliteit van beslissingen over personen in verschillende opzichten verbeteren. In de eerste plaats kan het gebruik van tests de kans op een correcte beslissing verhogen. Een tweede voordeel van het gebruik van tests is dat de objectiviteit van het beslisproces erdoor toeneemt. Willekeur bij het nemen van beslissingen kan voorkomen worden. Een derde voordeel is dat de transparantie van het beslisproces voor alle betrokkenen groter is. Aan alle stakeholders valt duidelijk te maken volgens welke procedure een beslissing tot stand gekomen is of tot stand zal komen. Een vierde voordeel is dat tests de gelegenheid bieden om op snelle en relatief eenvoudige wijze beslissingen te nemen over grote groepen personen. Een vijfde voordeel is dat het gebruik van tests kwantitatieve gegevens oplevert. Dat maakt het mogelijk wetenschappelijk onderzoek uit te voeren om de kwaliteit en bruikbaarheid aan te tonen van tests en om de beslisprocessen waar tests een rol bij spelen verder te verbeteren.

Om een uitspraak te kunnen doen over de kwaliteit van een test moet allereerst onderzocht worden in hoeverre de test voldoet aan twee fundamentele eisen. De eerste eis is die van meetnauwkeurigheid, of betrouwbaarheid (Haertel, 2006). Een test voldoet beter aan deze eis, naarmate de scores die personen erop behalen consistentere, en beter reproduceerbaar zijn, kortom vrijer zijn van meetfouten. Betrouwbaarheid is een noodzakelijke, maar onvoldoende voorwaarde voor de kwaliteit van een test.

De tweede eis is die van validiteit. Het concept validiteit kent een lange ontwikkelingsgeschiedenis; zie daarvoor Kane (2006). Oorspronkelijk werd validiteit gedefinieerd als de mate waarin een test meet wat men ermee beoogt te meten: *'The validity of a test is the extent to which... it measures what it purports to measure.'* (Garret, 1937, p.324, geciteerd in Angoff, 1988). Tot omstreeks 1950 was het gebruikelijk de correlatie tussen de score op een test en een criteriumvariabele te rapporteren als indicatie voor de validiteit van deze test (Angoff, 1988; Shepard, 1993). Sinds 1954 zijn er verschillende publicaties verschenen die de opvattingen over validiteit in een bepaalde periode codificeren. Het betreft hier opeenvolgende versies van de 'Standards for Educational and Psychological Testing', met als meest recente de versie van 1999 (American Psychological Association, American Educational Research Association & National Council on Measurement in Education, 1999) als ook verschillende hoofdstukken in de opeenvolgende edities van 'Educational Measurement' met als meest recente versie het hoofdstuk van Kane (2006). Deze publicaties geven

een duidelijk beeld van de veranderingen die zich de laatste decennia hebben voorgedaan in de opvattingen omtrent dit begrip.

De hierboven genoemde strikt operationele benadering is vanaf 1950 langzaam maar zeker verdrongen door een meer theoretisch georiënteerde zienswijze, waarin verschillende soorten validiteit een plaats innamen. Hoewel zich vanaf de jaren vijftig een ware wildgroei van begrippen heeft voltrokken waarin de term validiteit voorkwam (zie bijvoorbeeld van Berkel, 1984), was tot halverwege de jaren tachtig de als canon geaccepteerde indeling die naar inhouds-, construct- en criterium gerelateerde validiteit. De inhoudsvaliditeit van een test had betrekking op de mate waarin de items in de test een welomschreven en afgebakend universum representeren van mogelijk in de test op te nemen items. De constructvaliditeit van een test had betrekking op de mate waarin testcores zijn toe te schrijven aan verklarende concepten en constructen die deel uitmaken van het theoretisch kader dat aan de ontwikkeling van de test ten grondslag ligt. De criteriumgerelateerde validiteit van een test ten slotte, was de correlatie tussen de scores op de test en de scores op een variabele die een operationalisatie is van een actueel of toekomstig gegeven waar de test betrekking op moet hebben.

De klassieke driedeling is tegenwoordig achterhaald. Veel theoretici, onder aanvoering van Messick (1989) kwamen tot de conclusie dat de drie genoemde soorten validiteit verschillende aspecten zijn van een en hetzelfde overkoepelende begrip. En dat valideren een continu proces zou moeten zijn waarin men op gezette tijden op grond van empirische gegevens en theoretische overwegingen bepaalt in hoeverre het geoorloofd is om de scores op een test op een bepaalde manier te interpreteren en te gebruiken. : ... validity is a unified though faceted concept and ... validation is scientific inquiry.' (Messick, 1989, p. 14). In een valideringsonderzoek moet volgens Messick (1990) niet alleen aandacht zijn voor de betekenis van testcores, maar ook voor de consequenties van het gebruik van een test: '*... test validation is empirical evaluation of the meaning and consequences of measurement.*'. Cronbach (1988), Shepard (1993, 1997) en Linn (1997) stellen zich eveneens op dit standpunt. Het is in hun ogen niet voldoende om uitsluitend te onderzoeken in hoeverre het gewettigd is tot bepaalde uitspraken over eigenschappen van personen te komen. Ook de beslissingen die uit deze uitspraken voortvloeien, dienen onderwerp van onderzoek te zijn. Een valideringsonderzoek zal dan ook aandacht moeten besteden aan de gevolgen van het gebruik van een test: ... since validity depends on the uses to which results are put, it needs to include an evaluation of the consequences of those uses.' (Linn, 1997, p. 14).

Met deze laatste stelling zijn Popham (1997) en Mehrens (1997) en ook Borsboom, Mellenbergh en van Heerden (2004) het niet eens. Zij vinden dat valideringsonderzoek zich volledig zou moeten concentreren op de interpretatie van testcores. Hoewel het in kaart brengen van de gevolgen van het gebruik van een test belangrijk is, zou dit geen deel mogen uitmaken van het valideringsonderzoek.

Het is een academische vraag of een test valide genoemd mag worden, indien deze onbedoelde negatieve consequenties heeft. Het staat natuurlijk buiten kijf dat men bij het inzetten van een test ook rekening moet houden met dergelijke negatieve consequenties. Kane heeft in dit kader een interessante observatie (2006, p 54-55): '*Test-based accountability programs ... blur the distinction between intended consequences and social consequences by adopting as their major purpose the improvement of educational outcomes for all students ... These testing programs have moved ... to the use of testing as the engine of reform and accountability in education. Since these testing programs are*

intended to improve (or 'reform') educational institutions, it seems reasonable to evaluate them as educational programs. Program evaluations include the evaluation of intended and unintended outcomes of the program being evaluated.'

Een fenomeen waar veel naar verwezen wordt in de onderwijsliteratuur is het 'backwash' of 'washback'- effect. Het is een bekend gegeven dat de inhoud van toetsen die van belang zijn voor kandidaten hun leergedrag en motivatie sterk stuurt. Recent heeft Joosten-ten Brinke (2011) hier bijvoorbeeld nog op gewezen rond het thema toetsing in beroepsgerichte opleidingen. Het backwash-effect bij examens houdt bijvoorbeeld in dat studenten leren alleen maar aandacht besteden aan die onderdelen van het curriculum waarvan duidelijk is dat ze ook werkelijk geëxamineerd worden. Voor toetsen die een rol spelen bij overgangsbepalingen betekent het dat leerlingen aan het beheersen van de inhoud ervan relatief meer aandacht zullen besteden dan aan kennis en vaardigheden die niet aan bod komen. Omdat dergelijke toetsen zich uit het oogpunt van efficiëntie en het vermijden van collineariteit per definitie beperken tot een kleine set vaardigheden, heeft dit negatieve gevolgen voor het beheersen van die andere –toch niet onbelangrijke– vaardigheden. In het vervolg van deze notitie zal nog op dit punt worden teruggekomen. Het backwash-effect is een voorbeeld van wat in deze notitie in navolging van Frederiksen en Collins (1989) een systeemeffect genoemd zal worden.

Het nemen van overgangsbepalingen in het onderwijs dient met de grootst mogelijke zorgvuldigheid te gebeuren. Overgangsbepalingen in het onderwijs zijn 'high-stakes'. Er is sprake van zware consequenties voor belanghebbenden. Voor leerlingen, deelnemers en studenten geldt dat dergelijke bepalingen hun schoolcarrière en vervolgens hun loopbaan en maatschappelijke positie sterk in positieve of negatieve zin kunnen beïnvloeden. Voor de overheid geldt dat hoe optimaler schoolcarrières verlopen, hoe lager de kosten van het onderwijsstelsel zullen zijn en hoe beter de wereldwijde concurrentiepositie van Nederland. Voor instellingen geldt dat hun rendement toeneemt naarmate ze beter presteren op dit vlak en dat hun aantrekkelijkheid hierdoor navenant zal stijgen. Daarom is het van belang dat elk instrument dat bij het nemen van overgangsbepalingen ingezet wordt, voldoet aan de eisen van betrouwbaarheid en validiteit. Ontwikkelaars hebben de plicht om concreet aan te tonen dat hun procedures zinvol in te zetten zijn bij overgangsbepalingen. Daartoe zou gebruik gemaakt kunnen worden van het beoordelingssysteem voor de kwaliteit van tests van de Commissie Testaangelegenheden Nederland van het Nederlands Instituut van Psychologen (COTAN, 2010) of het beoordelingssysteem voor de kwaliteit van toetsen en examens van het Research Center voor Examinering en Certificering (RCEC, 2011).

De criteria die de COTAN hanteert zijn:

- Uitgangspunten bij de testconstructie;
- Kwaliteit van het testmateriaal;
- Kwaliteit van de handleiding;
- Normen;
- Betrouwbaarheid;
- Begripsvaliditeit;
- Criteriumvaliditeit.

Instrumenten kunnen op deze beoordelingscriteria op basis van de antwoorden op een reeks bij de criterium gestelde vragen op ieder criterium als goed, voldoende of onvoldoende beoordeeld

worden. Het beoordelingssysteem van het RCEC is een adaptatie van het systeem van de COTAN. Het systeem van de COTAN is oorspronkelijk ontwikkeld voor het beoordelen van tests. Het systeem van het RCEC is een afgeleide hiervan speciaal ontwikkeld voor het beoordelen van toetsen en examens.

Deze twee beoordelingssystemen leggen echter weinig tot geen nadruk op het in kaart brengen van de onbedoelde consequenties van het inzetten van tests, zoals boven beschreven. Eerder in deze notitie is echter al duidelijk gemaakt dat het inzetten van op zichzelf valide (in de opvatting van bijvoorbeeld Borsboom et al.) instrumenten bij overgangsbepalingen niet volstaat. Beleidsmakers zouden zich moeten realiseren dat het inzetten van dergelijke instrumenten effecten met zich mee kan brengen die op stelselniveau ongewenst zijn.

2.4. Het subjectieve karakter van overgangsbepalingen

Psychologisch onderzoek van beslisprocessen op verschillend terreinen, waaronder het onderwijs (Cooksey, 1988), heeft duidelijk gemaakt dat de mens in de regel geen goede voorspeller en beslisser is (Meehl, 1954; Kahneman, Slovic & Tversky, 1982; Grove, Zald, Lebow, Snitz, & Nelson, 2000). Dit laatst genoemde onderzoek betreft een meta-analyse van maar liefst 136 studies uit onder meer de klinische psychologie, de arbeids- en organisatiepsychologie en de onderwijspsychologie die de oorspronkelijke conclusies van Meehl ondersteunt. Beoordelaars zijn niet goed in het integreren van relevante informatie over kandidaten tot een voorspelling over toekomstige prestaties. Daarom constateert Born (2007) in het kader van personeelsselectie bijvoorbeeld dat bepalingen het best gebaseerd kunnen zijn op een formele, rekenkundige formule die voortkomt uit eerder empirisch onderzoek met een gevonden verband tussen de informatie over kandidaten en hun latere werkgedrag. Drenth en Sijtsma (2006, p.416) stellen bijvoorbeeld in dit kader: *'Dit betekent dat ... voor de bepaling van iemands kans op succes een statistische combinatie van testgegevens nauwkeuriger is dan een intuïtief weegproces. De actuarische verwerking leidt tot betere resultaten, dus tot minder foute bepalingen'*. Ook de psychometrische besliswetenschap gaat ervan uit dat bepalingen het best op een actuarische wijze tot stand kunnen komen. Mellenbergh (2011, p. 418), bijvoorbeeld, stelt dat: *'Theoretically the decision-theoretic approach has to be preferred above other approaches because it yields a coherent framework to choose optimal selection ratios. Moreover, decision theory requires decision makers to make their value judgments explicit.'* Daar kan echter tegen in worden gebracht dat veel beslissers niet bekend zijn met het specificeren van utiliteiten én dat er vaak sprake is van een beslissende instantie die uit verschillende personen bestaat met uiteenlopende opvattingen. Dat maakt toepassing in de praktijk lastig.

Desondanks treffen we bij verschillende overgangsbepalingen in het onderwijs regelingen aan die veel kenmerken hebben van de actuarische benadering. Bijvoorbeeld bij de eindexamens in het voortgezet onderwijs en bij de entreetoetsen taal en rekenen voor de pabo's. In andere gevallen worden tests gebruikt in combinatie met andere bronnen van informatie gebruikt en is niet duidelijk hoe de aanwezige informatie tot een eindoordeel geïntegreerd wordt.

2.5. Doorlopende leerlijnen

Het Nederlands onderwijsstelsel is geen statisch maar een dynamisch systeem. De eerder afgebeelde stroomdiagrammen van beslissingen zijn een simplificatie van de werkelijkheid. Feitelijk is er in de verschillende onderwijssectoren geen sprake van een en dezelfde 'behandeling', maar is sprake van –beperkte- kwaliteitsverschillen die hun invloed hebben op het studiesucces van leerlingen, deelnemers en studenten. De gehanteerde procedures bij overgangsbepalingen kunnen subtiel of principieel verschillen, afhankelijk van de vraag of een instelling kansen wil bieden, of meer de nadruk legt op een hoog rendement. Dit kan ook gelden voor –tussentijdse- criteria die ingezet worden om studiesucces te bepalen. De examens in het vo, mbo, hbo en wo kunnen beschouwd worden als 'eindcriteria' die er voor zorgen dat het niveau van de verschillende opleidingen geborgd wordt. Terzelfdertijd zijn die examens ook te beschouwen als tests op basis waarvan selectie voor een verdere schoolcarrière of loopbaan plaatsvindt.

Uitvoering van de plannen die gemaakt zijn in het kader van doorlopende leerlijnen naar aanleiding van het advies van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2007) leiden tot veranderingen in het beheersingsniveau van de taal- en rekenvaardigheid in het hele onderwijsstelsel. Het rapport van de Expertgroep Doorlopende Leerlijnen geeft een systematische beschrijving van de kennisdomeinen taal en rekenen. Hierbij wordt aangegeven wat leerlingen op belangrijke momenten in hun schoolloopbaan moeten kennen en kunnen. Deze belangrijke momenten noemt de Expertgroep 'drempels'. Er zijn vier drempels vastgesteld in een schoolloopbaan van primair onderwijs tot uitstroom naar de arbeidsmarkt of instroom in het hoger onderwijs. Het niveau dat leerlingen moeten behalen om een drempel goed te kunnen nemen, wordt het fundamentele niveau (F) genoemd. Er zijn ook streefniveaus (S) geformuleerd voor leerlingen die meer aankunnen (S).

Op 1 augustus 2010 is de Wet Referentieniveaus Nederlandse taal en rekenen (Rijksoverheid, 2010) van kracht geworden. Deze wet geldt voor alle onderwijssectoren en draagt daarmee zorg voor inhoudelijke samenhang tussen deze sectoren. In de wet is vastgesteld dat de referentieniveaus verplicht zijn van primair onderwijs tot uitstroom naar de arbeidsmarkt en instroom naar het hbo. Het is de bedoeling dat in het vmbo en mbo niveau 1, 2 en 3 referentieniveau 2F beheerst zal gaan worden en in havo, vwo en mbo niveau 4 het referentieniveau 3F.

Om te bepalen of leerlingen en studenten het gevraagde referentieniveau beheersen, is er naast een doorlopende leerlijn ook een doorlopende toetslijn opgesteld met verplichte centrale toetsen en examens. In het voortgezet onderwijs zullen de referentieniveaus in het eindexamen Nederlands getoetst worden en wordt een centrale rekentoets ingevoerd. In de mbo niveaus 2, 3 en 4 worden centraal ontwikkelde examens voor de Nederlandse taal en voor rekenen ingevoerd om te kunnen bepalen of taal en rekenen voldoende beheerst worden.

De implementatie van de plannen zou moeten leiden tot een verhoging van de eisen in alle onderwijssectoren voor wat betreft beheersing van de Nederlandse taal en de rekenvaardigheid. Dit zal leiden tot veranderingen in de opleidingen en mogelijk ook leiden tot een toename in de uitval van leerlingen, deelnemers en studenten, c.q. het rendement van opleidingen. Dit zal ook tot aanpassingen, bijvoorbeeld herijking, moeten leiden van de instrumenten die ingezet worden voor het ondersteunen van overgangsbepalingen.

2.6. Tot slot

In dit hoofdstuk is aandacht besteed aan verschillende thema's. In paragraaf 2.1 is duidelijk gemaakt dat het paradigma van de psychometrische besliskunde mogelijkheden biedt om het geven en ontzeggen van kansen aan leerlingen, deelnemers en studenten te optimaliseren. In paragraaf 2.2. is aandacht besteed aan de tweede uitdaging bij het verbeteren van overgangsbepalingen in het onderwijs: het zo goed mogelijk voorspellen van toekomstig studiegedrag. In paragraaf 2.3 is ingegaan op de kwaliteitseisen waar procedures die ingezet worden bij overgangsbepalingen aan moeten voldoen. Daarbij is in het bijzonder aandacht gegeven aan het feit dat er zich ook bij het inzetten van procedures van goede kwaliteit voor het onderwijsstelsel onbedoelde negatieve consequenties kunnen zijn. In paragraaf 2.4, is aangegeven dat mensen geen goede voorspellers en beslissers zijn en dat daarom actuarische regels de voorkeur verdienen bij besluitvorming in het kader van overgangsbepalingen. In paragraaf 2,5 tot slot is gewezen op mogelijke veranderingen die implementatie van de referentieniveaus in het onderwijs met zich mee zal brengen voor overgangsbepalingen. De informatie in dit hoofdstuk is bedoeld om een referentiekader te creëren waarmee de inhoud van de volgende hoofdstukken beter te duiden is.

3

De overgang naar het vo

3.1. Stand van zaken

Bij de overgang van po naar vo is sprake van een classificatiebeslissing. Hiervoor zetten veel poscholen toetsen in. De wettelijke grondslag hiervoor ligt in het Inrichtingsbesluit bij de Wet op het Voortgezet Onderwijs (Rijksoverheid, 1993). In deze regeling staat dat de toelating tot het eerste leerjaar vwo, havo en mavo mede gebaseerd moet zijn op *'een onderzoek naar de geschiktheid voor het volgen van het onderwijs aan de school waarvoor de toelating wordt gevraagd'*. Een *'onderzoek naar de kennis en het inzicht'* van de leerlingen in het laatste leerjaar van het primair onderwijs is één van de genoemde onderzoeksvormen. De andere drie zijn het volgen van een *'proefklasse'*, een *'toelatingsexamen, door de leraren van de school afgenomen, dat zich ten minste uitstrekt over de vakken Nederlandse taal en rekenen'* of een *'psychologisch onderzoek'*.

Er lijkt dus een grote keuzevrijheid te zijn voor scholen wat betreft het in te zetten instrumentarium. In de praktijk is er echter weinig variatie. Dat heeft onder meer te maken met het feit dat de Inspectie van het Onderwijs bij het beoordelen van de eindresultaten van het onderwijs gebruik maakt van instrumenten die ook in te zetten zijn bij het bepalen van de geschiktheid van leerlingen. Het betreft hier het Drempelonderzoek 678 van 678 Onderwijs Advisering, de Eindtoets Basisonderwijs van Cito of het Schooleindonderzoek van de vier Landelijke Protestantse Christelijke Schoolbegeleidings-diensten (Staatscourant, 9 mei 2012). De praktijk laat dan ook zien dat maar weinig scholen voor andere instrumenten kiezen dan de hiervoor genoemde; waarschijnlijk vanwege de extra inspanningen en kosten die dat met zich meebrengt.

Het Drempelonderzoek (678 Onderwijs Advisering) bestaat uit vijf toetsen op het gebied van technisch lezen, spelling, woordenschat, begrijpend lezen en rekenen. Met behulp van een zogenoemd plaatsbepalingsformulier worden de individuele scores van leerlingen vergeleken met scores van havo/vwo-leerlingen, praktijkschoolleerlingen en leerlingen in de vmbo-leerwegen; gemengd-theoretisch, basis- en kaderberoepsgericht, wel of niet met leerwegondersteuning. Het didactisch profiel van de leerling wordt vertaald in een plaatsingsadvies op basis van een "veilige" plek voor de leerling. Een leerling is veilig als ongeveer een kwart van de leerlingen in zijn omgeving minder weten/kunnen. Het Drempelonderzoek wordt binnen het basisonderwijs in enkele regio's al in de maanden september en oktober door leerlingen in groep acht gemaakt. Het grootste deel van de afnames vindt plaats tussen de herfstvakantie en de kerst. In het voortgezet onderwijs wordt het Drempelonderzoek gebruikt in de maanden voor de zomervakantie om aangemelde leerlingen vanuit het basisonderwijs zorgvuldiger te kunnen classificeren. Het instrumentarium is door de COTAN op de eerder genoemde kwaliteitscriteria beoordeeld als 'goed' of 'voldoende'.

De Eindtoets Basisonderwijs van Cito meet wat een leerling in vergelijking met andere leerlingen in acht jaar basisonderwijs geleerd heeft op het gebied van taal, rekenen-wiskunde, studievoordigheden en -facultatief- wereldoriëntatie. De toets verstrekt scholen informatie over de kans op succes in de verschillende brugklassen/onderwijstypen. Van de Eindtoets Basisonderwijs verschijnt elk jaar een nieuwe versie. Op basis van het totaal aantal goed op de onderdelen taal, rekenen-

wiskunde en studievoordigheden wordt de zogenaamde standaardscore berekend. De standaardcores liggen op een schaal waarnaar de scores op de toetsen in de verschillende jaargangen van de Eindtoets te vertalen zijn. Door de standaardcoreschaal is Cito in staat de scores van jaar tot jaar te vergelijken ondanks verschillen in moeilijkheid tussen diverse jaargangen. Verder vindt elk jaar toelatings- en doorstromingsonderzoek plaats om de relatie tussen de standaardcore en brugklatype te onderzoeken. De Eindtoets basisonderwijs is door de COTAN op alle eerder genoemde kwaliteitscriteria beoordeeld als 'goed', met uitzondering van het kwaliteitscriterium begripsvaliditeit; daar is het oordeel 'voldoende'.

Het Schooleindonderzoek bestaat uit een combinatie van een intelligentietest en een toets, facultatief aan te vullen met een persoonlijkheidsvragenlijst. De betreffende test is de Nederlandse Intelligentietest voor Onderwijsniveau (de NIO), uitgegeven door Boom testuitgevers. De NIO bestaat uit zes subtests op het gebied van verbaal inzicht (Synoniemen, Analogieën en Categorieën) en rekenkundig-ruimtelijk inzicht (Getallen, Rekenen en Uitslagen). Voor de verbale schaal, de rekenkundig-ruimtelijke schaal en het totaal wordt informatie verstrekt over hoe goed de leerling presteert in vergelijking met leerlingen in de derde klas van verschillende typen voortgezet onderwijs. De NIO is door de COTAN op alle eerder genoemde kwaliteitscriteria beoordeeld als 'goed'. De leervorderingen worden bij het Schooleindonderzoek gemeten met behulp van toetsmateriaal dat ontwikkeld is door het Cito, vergelijkbaar met de Cito-eindtoets. Gedragskenmerken worden in beeld gebracht met behulp van de Nederlandse Persoonlijkheidsvragenlijst – junior (NPV-J), uitgegeven door Pearson Assessment B.V. De NPV-J kent vijf onderdelen: inadequatie, volharding, sociale Inadequatie, recalcitrantie en dominantie. Ook deze test heeft een positieve COTAN-beoordeling.

Een ander instrument dat regelmatig ingezet wordt, is de Drempeltest van Boom testuitgevers. De Drempeltest meet de aanleg van de leerlingen op acht gebieden: verbaal-theoretisch (VT): woordenschat, zinsbouw, logisch redeneren met woorden, soortbegrip u wiskundig-praktisch (WP): ruimtelijk twee- en driedimensionaal, rekenkundig redeneervermogen, logisch redeneren met figuren. De drempeltest is door de COTAN op alle kwaliteitscriteria beoordeeld als 'goed', of 'voldoende'.

Verder is sprake van verschillende andere regionale procedures die kenmerken vertonen van het al genoemde schooleindonderzoek. Een niet uitputtende opsomming:

Het Enschedees Eindonderzoek Groep Acht (EEG8) is een voorbeeld van afspraken tussen het openbaar primair onderwijs en Het Stedelijk Lyceum in Enschede in verband met het schoolkeuzeadvies en het onderwijskundig rapport. Het EEG8 is een verzamelnaam voor een intelligentietest, schoolvorderingen op het gebied van taal-spelling, rekenen, begrijpend lezen en technisch lezen en de sociaal-emotionele ontwikkeling van het kind.

In Drenthe ondersteunt Timpaan Onderwijs scholen in basisonderwijs en voortgezet onderwijs bij gebruik van de Plaatsingswijzer. De Plaatsingswijzer werkt met twee belangrijke indicatoren: de leervorderingen en de leerhouding. Voor de leervorderingen wordt er gebruik gemaakt van reeds verzamelde gegevens uit het leerlingvolgsysteem van Cito van groep 6, 7 en 8 (begrijpend lezen, technisch lezen, spelling en rekenen/wiskunde). Voor de leerhouding spelen de dagelijkse indrukken van de leerkracht een belangrijke rol, eventueel aangevuld met gegevens uit bijvoorbeeld de

schoolvragenlijst (SVL) of de Apeldoornse vragenlijst (AVL). Ook de mening van de ouders en het kind worden meegenomen in het uiteindelijke advies.

In Limburg kiest een aantal basisscholen voor het Groep 8-onderzoek van het Loopbaan Adviesbureau Limburg. Het betreffende onderzoek richt zich op intelligentie, schoolse kennis (tekstbegrip, spelling en rekenen), werkhouding en kenmerken die het studiegedrag kunnen beïnvloeden, zoals concentratie werktempo en zelfvertrouwen.

Ook zijn er steden en gemeentes die de instrumenten die ingezet worden bij overgangsbepalingen naar het vo inzetten in het kader van het verbeteren van de prestaties van po-scholen. In Amsterdam is bijvoorbeeld sinds 1995 het gemeentelijk onderwijsbeleid expliciet gericht op het verbeteren van de leerprestaties, zoals gemeten met de Eindtoets Basisonderwijs van Cito. Reden hiervan was dat in het begin van de jaren negentig in Amsterdam grote zorg ontstond over de kwaliteit van zowel het basis- als het voortgezet onderwijs omdat de uitval in het voortgezet onderwijs veel hoger was dan in de rest van Nederland. Na de introductie van het zogenoemde NBRA-beleid zijn de prestaties van leerlingen in Amsterdam verbeterd ten opzichte van de prestaties van leerlingen in de rest van Nederland. De verbetering kan niet worden toegeschreven aan strategisch gedrag van scholen (Chorny & Webbink, 2010). En in Rotterdam gebruikt men sinds 1997 de Eindtoets Basisonderwijs in een strak vormgegeven doorstroomprocedure (Kassenberg, 2005).

Alle genoemde procedures baseren zich op het in hoofdstuk 2 beschreven principe; het in kaart brengen van informatie over leerlingen die in statistische zin een samenhang vertoont met toekomstige prestaties in de schoolloopbaan. Vervolgens worden dan besluiten genomen op basis van deze informatie. De mate waarin evidentie wordt aangedragen voor de betrouwbaarheid en validiteit van instrumenten en procedures varieert. Enkele voorbeelden zijn te vinden in verantwoordingen van Van Boxtel, Engelen en De Wijs (2010), Van Dijk en Tellegen (2004) en Kapinga (2006) en onderzoeksverslagen van Eichorn, R., Jansson, R.J., & Van Steenberg (2004), de Inspectie van het Onderwijs (2011a), Kapinga (2008), Stroucken, Takkenberg en Béguin (2008) en Hakkenes en De Wijs (2012).

Verder zijn er de nodige 'tools' ontwikkeld om de communicatie tussen po en vo te verbeteren, en de administratieve lasten te verlichten. Dit vanuit de gedachte dat er meer gefocused kan worden op de problematiek zelf, zodat mogelijk een groter percentage van de leerlingen correct geclassificeerd wordt. Een eerste voorbeeld is de Overstapservice Onderwijs (zie <http://www.oso-od.nl/>). Het betreft hier een gezamenlijke initiatief van PO-Raad en VO-raad, dat bedoeld is om uiteindelijk bij alle po- en vo-scholen een zogenoemd overstapdossier te implementeren. De website Overgangpovo.nl (zie <http://www.Overgangpovo.nl>) is een tweede voorbeeld. Deze omvangrijke, door KPC Groep ontwikkelde, website bevat informatie die voor scholen, ouders en samenwerkingsverbanden informatie bevat om de overgang van basisschool naar voortgezet onderwijs te vergemakkelijken. Voor ouders en leraren is inzichtelijk gemaakt aan welke zaken aandacht besteed moet worden vanaf groep 7 van de basisschool tot en met de brugklas. Voor samenwerkingsverbanden en scholen zijn een kijkwijzer en standaarden voor een goed verlopende overgang ontwikkeld. Een derde en laatste voorbeeld is het programma 'Effectief schakelen' (PO-raad, AVS en VO-raad, 2011) dat eveneens tot doel heeft om de aansluiting van primair naar voortgezet onderwijs te versoepelen. Dit programma richt zich onder meer op het verbeteren van

de informatie-uitwisseling tussen PO en VO en het verbeteren van het instroomniveau van leerlingen in het VO.

3.2. Recente ontwikkelingen

Enkele jaren geleden constateerde de OESO dat door de vroege selectie in het Nederlandse onderwijs leerlingen uit achterstandsgroepen zouden worden benadeeld. Ook zou hierdoor de doorstroom naar het hoger onderwijs achterblijven. Landen waar selectie later plaatsvindt zouden het volgens de OESO in beide opzichten beter doen. Naar aanleiding hiervan heeft de Onderwijsraad onderzoek laten doen en een advies uitgebracht (Onderwijsraad, 2010). De Onderwijsraad geeft in dat advies aan dat naar zijn mening er onvoldoende aanwijzingen zijn dat een breed en voor alle scholen verplicht uitstellen van het selectiemoment leidt tot een verbetering van schoolprestaties.

Begin 2012 is een voorstel tot wetwijziging ingediend bij de Tweede Kamer in verband met onder andere de invoering van een verplichte centrale eindtoets (Rijksoverheid, 2012). Deze toets zou moeten gaan fungeren als het tweede objectieve gegeven bij de overdracht van het primair naar het voortgezet onderwijs. Daarnaast zou deze verplichte eindtoets moeten gaan fungeren als instrument voor de externe beoordeling van de leerresultaten van scholen. De derde functie zou dan het fungeren als hulpmiddel voor zelfevaluatie van scholen zijn, in het kader van opbrengstgericht werken.

Het wetsvoorstel is nog niet behandeld door de Tweede Kamer. Vooralsnog is dus nog onduidelijk of en op welke termijn er een centrale eindtoets zal komen. Bij invoering van één centrale eindtoets maken alle leerlingen in dezelfde periode dezelfde toets, waardoor een eenduidige overdracht en interpretatie van overeenkomstige gegevens van primair naar voortgezet onderwijs mogelijk wordt. Het achterliggende idee bij dit scenario is dat de overheid aldus in staat is om via het College voor Examens de gewenste regie te voeren op de informatie waarmee scholen zich verantwoorden voor hun onderwijskwaliteit. Dit zorgt voor transparantie bij de waardering van de resultaten van leerlingen en scholen. Verder is de gedachte dat de overheid met één centrale eindtoets ook een instrument in handen heeft om de kwaliteit van het onderwijsstelsel systematisch te volgen. Voor het thema overgangsbepalingen is echter voornamelijk van belang dat op deze wijze de overgangsbepaling van PO naar VO verder gestandaardiseerd wordt.

Bij dit scenario kunnen zich echter verschillende systeemeffecten voordoen. Zo kunnen scholen zich in hun onderwijs te sterk gaan richten op de inhoud van de centrale eindtoets; het eerder gememoreerde backwash-effect. Ook bij (ouders van) leerlingen zal de focus meer komen te liggen op de wens om toch vooral de inhoud van die centrale eindtoets te beheersen. Een tweede probleem is dat in dit scenario met één en hetzelfde instrument verschillende functies gerealiseerd moeten worden. Een toets die slechts één functie heeft, kan zodanig ontwikkeld worden dat deze functie optimaal vervuld wordt. Het tezelfdertijd goed vervullen van verschillende functies wordt echter lastiger, naarmate die functies sterker van karakter verschillen. In het geval van een landelijke centrale eindtoets vormt het moment van afname op dit vlak een probleem. In het kader van de externe verantwoording en het vaststellen van het eindniveau van leerlingen voor de overdracht naar het vo, is een zo laat mogelijke afname in het schooljaar gewenst. Voor de adviesfunctie, daarentegen, wensen veel scholen juist een afname vroeger in het schooljaar. Het primair onderwijs kan de resultaten ervan dan beter betrekken in de uiteindelijke advisering én de

scholen voor voortgezet onderwijs beschikken tijdig over deze informatie in verband met de plaatsing van leerlingen voor het volgende schooljaar.

Verder is van belang bij het bepalen van het afnametijdstip dat zich ook hier een systeemeffect kan voordoen. Een vroege afname – zoals nu het geval is bij vrijwel alle tests en toetsen – kan tot effect hebben dat scholen in de daaropvolgende maanden geen effectief onderwijs meer geven en dat het leerlingen aan motivatie ontbreekt om dat onderwijs te volgen. Bovendien beoogt de overheid door een latere afname om het gewicht van het advies van de school voor primair onderwijs te vergroten. Sommige vo-scholen hanteren de resultaten op de Eindtoets Basisonderwijs immers als ingangseis voor plaatsing in een bepaalde leerweg en/of brugklatype. En dit zou onvoldoende recht doen aan het professionele oordeel van de basisschool. Door een latere afname komt het tweede onafhankelijke gegeven later beschikbaar, waardoor scholen voor voortgezet onderwijs zich meer moeten verlaten op het schooladvies.

Het is lastig, zo niet onmogelijk, om hier een balans te vinden. Hoe later de centrale eindtoets in het laatste schooljaar wordt afgenomen, hoe minder invloed de uitkomst ervan zal kunnen hebben op de plaatsing van de leerling in het vo. Vo-scholen moeten immers hun formatie en rooster rond hebben, boeken moeten tijdig kunnen worden besteld. Wanneer de toets geen concrete rol meer speelt bij het bepalen van het meest geschikte type van voortgezet onderwijs, zullen leerlingen dit snel bemerken. Daarmee is de motivatie om goed te presteren op de toets verdwenen. Het lijkt plausibel te veronderstellen dat leerlingen in een dergelijke situatie de toets aanmerkelijk minder goed zullen maken. Hiervoor is ook al empirisch bewijs gevonden (Hemker & Mittelhaëuser, 2012). De scores op de toets zullen in zo'n geval geen reëel beeld meer geven van wat leerlingen daadwerkelijk kennen en kunnen aan het einde van het basisonderwijs. En daarmee verliest de toets ook zijn waarde voor de functies zelfevaluatie, externe verantwoording en beoordeling. Hoe eerder de centrale eindtoets wordt afgenomen, hoe minder waarde de toets zal hebben voor het vaststellen van het eindniveau van leerlingen bij de overgang van po naar vo. Merk in dit kader overigens op dat het ontwikkelen van twee aparte instrumenten niet tot een volledige oplossing zal leiden. Een combinatie van een vroege centrale eindtoets ten behoeve van overgangsbepalingen met een late eindtoets voor niveaubepaling zal eveneens te leiden hebben van het feit dat er voor leerlingen bij de tweede toets niets op het spel zal staan, waardoor zij suboptimaal zullen presteren.

Een ander systeemeffect is dat de kans op strategisch gedrag bij scholen toeneemt naarmate zij ervaren dat er meer voor ze op het spel staat bij de beoordeling van die eindopbrengsten. In de eerste plaats zullen dan de andere functies – verwijzing, niveaubepaling van leerlingen en zelfevaluatie- minder goed vervuld kunnen worden. En in de tweede plaats zullen inhoudsgebieden die niet in de toets voorkomen op de achtergrond raken, zodat het curriculum zal versralen. Directie en leerkrachten zullen zich in extreme gevallen slechts richten op het zo hoog mogelijk scoren op de centrale eindtoets. Indien de invoering van een centrale eindtoets tot dit soort effecten leidt, zal eerder sprake zijn van een verlaging dan van een verhoging van de kwaliteit van het primair onderwijs.

Veel scholen voor voortgezet onderwijs starten met een brugperiode waarin nog geen categoriale vormen van voortgezet onderwijs worden aangeboden, maar waarin leerlingen in klassen van gemengde samenstelling zijn ondergebracht. Aan het eind van de brugperiode moet dan alsnog classificatie plaatsvinden. Tot voor enkele jaren geleden was het gebruik van toetsen hierbij ongebruikelijk. Een derde recente ontwikkeling is dat de laatste jaren sprake is van een duidelijke

stijging in het percentage VO-scholen dat met dergelijke toetsen werkt. Het betreft hier in concreto de toetsen uit het Volgsysteem VO van Cito dat een sterk groeiende populariteit kent (zie: http://www.cito.nl/nl/onderwijs/voortgezet%20onderwijs/cito_volgsysteem_vo.aspx). De combinatie van een brugperiode met heterogene klassen en het inzetten van toetsen die daadwerkelijk in staat zijn om de ontwikkeling van leerlingen te volgen, biedt scholen de mogelijkheid om extra informatie te vergaren over de geschiktheid van leerlingen voor de verschillende soorten van voortgezet onderwijs die de school aanbiedt. Daarmee neemt de kans op misclassificaties verder af, mede omdat scholen ook de gelegenheid krijgen om te remediëren op basis van deze informatie.

3.3. Evaluatie

In de voorafgaande twee paragrafen is een niet uitputtende beschrijving gegeven van de stand van zaken bij overgangsbepalingen naar het vo en is stil gestaan bij negatieve effecten die zich op stelselniveau kunnen voordoen vanwege de invoer van een centrale eindtoets. Het afzetten van de stand van zaken op dit vlak tegen hetgeen in hoofdstuk 2 aan de orde is gekomen, leidt tot de volgende observaties.

In de eerste plaats valt op dat er veel aandacht is voor de kwaliteit van instrumenten die een rol spelen bij het nemen van overgangsbepalingen. De nodige instrumenten zijn door een onafhankelijke instantie –i.c. de COTAN- positief beoordeeld op een reeks kwaliteitseisen. Er kan dus vertrouwd worden op hun betrouwbaarheid en validiteit. Dit betekent dat deze instrumenten daadwerkelijk in te zetten zijn bij het nemen van overgangsbepalingen. Een recente ontwikkeling rond het gebruik van toetsen is dat er in de eerste jaren van het VO een toenemend gebruik valt te bespeuren van toetsen voor het volgen van de ontwikkeling van leerlingen. Dit biedt een betere basis voor determineren en geeft scholen tevens de mogelijkheid om te remediëren.

In de tweede plaats valt op dat er bepaald geen sprake is van één en dezelfde standaardprocedure. De Eindtoets Basisonderwijs van Cito is een veelgebruikt instrument, maar bepaald niet het enige. Het is niet mogelijk om de gehanteerde instrumenten en procedures op kwaliteit te rangordnen en te stellen dat de een 'beter' is dan de ander. De onderzoeken die in dit kader hebben plaats gevonden zijn alle correlatief of hoogstens 'quasi-experimenteel' (Cook & Campbell, 1979) van aard en maken een dergelijke conclusie onmogelijk. Onder meer vanwege verschillende operationalisaties van tests en criteria, verschillende opzetten, het ontbreken van kruisvalidering en onvolkomenheden in het onderzoeksdesign, of het ontbreken van informatie daarover.

Zoals vermeld is bij wet bepaald dat toelating tot het vo mede gebaseerd moet zijn op een tweede gegeven naast het oordeel van de school zelf. Er zijn echter geen richtlijnen voor de wijze waarop scholen in de praktijk het eigen oordeel met de extra informatie moeten combineren. Deze worden dan ook regionaal en lokaal anders ingevuld. Tot op zekere hoogte zal dit in de praktijk leiden tot willekeur. Indien de resultaten van het wetenschappelijk onderzoek waar al eerder aan gerefereerd is, ook gelden voor overgangsbepalingen van PO naar VO, dan zou hier ruimte voor verbetering liggen. Merk op dat het gebruik door VO-scholen van strikte rekenregels voor het aannemen van leerlingen voor verschillende vormen van voortgezet onderwijs op basis van de wetenschappelijke literatuur te billijken is. Het terugvallen op louter toetsresultaten is echter om twee redenen niet aan te raden. Allereerst, omdat het niet voldoet aan hetgeen bij wet bepaald is. Bovendien zal hiermee

het sturende karakter van de inhoud van de toetsen alleen maar sterker worden en leiden tot verwatering van het curriculum en mogelijk zelfs ondergraving van de validiteit van de betreffende instrumenten. Het opstellen van een landelijk standaardprotocol zou een rol kunnen spelen in het terugdringen van willekeur bij overgangsbepalingen naar het VO en vergroting van de transparantie van het proces.

Door te kiezen voor één centrale eindtoets zouden overgangsbepalingen verder gestandaardiseerd kunnen worden en dat heeft de nodige voordelen. Een probleem hierbij is dat de overheid van plan is om met het beoogde instrument verschillende functies te realiseren. Dit doet afbreuk aan de rol die het instrument zou kunnen spelen bij overgangsbepalingen. Er zijn verschillende systeemeffecten met negatieve consequenties mogelijk, zowel voor de kwaliteit van het onderwijs als voor de kwaliteit van overgangsbepalingen naar het VO.

Een laatste observatie is dat de kans aanwezig is dat implementatie van de referentieniveaus (zie paragraaf 2.5) ertoe zal leiden dat de tests en toetsen die nu een rol spelen bij het nemen van overgangsbepalingen naar het VO herijkt zullen moeten worden, omdat de opleidingen in het VO deels zullen veranderen en daarmee de aard van de te voorspellen criteria.

4

De overgang naar het mbo

4.1. Stand van zaken

Volgens de meest recente cijfers van het Centraal Bureau voor de Statistiek (CBS, 2012) zaten in het schooljaar 2010/'11 ruim 100.000 leerlingen in het examenjaar van het vmbo. Aan het eind van het schooljaar had 90 procent van hen een diploma behaald. Vanuit de basis- en kaderberoepsgerichte leerweg van het vmbo stroomde bijna 95 procent van de gediplomeerden een jaar later naar een opleiding in het mbo. Vanuit de gemengde en theoretische leerweg van het vmbo stroomde een jaar later ongeveer 81 procent van de gediplomeerden naar het mbo. Verder vervolgt een kleine groep van 3 procent van de havo-gediplomeerden de opleiding op het mbo. Van deze groep gaat 79 procent naar een mbo-opleiding op niveau 4. Ongeveer 14 procent gaat een opleiding op niveau 3 doen en zelfs 7 procent op niveau 2.

Bij de overgang naar het mbo is voor een deel van de opleidingen formeel sprake van een gecombineerde beheersings/selectiebeslissing op basis van toetsresultaten. Diploma's worden immers verkregen op basis van examenresultaten (beheersing) en voor verschillende opleidingen is het in het bezit hebben van een diploma een vereiste. Merk op dat de vmbo-examens niet worden ontwikkeld met het doel om zo goed mogelijk te voorspellen. Zij zijn uitsluitend en alleen bedoeld om vast te stellen of leerlingen voldoende hebben opgestoken van het voorafgaande onderwijs. Merk op dat de beslissing gebaseerd wordt op de reeks cijfers voor het centrale examen én het schoolexamen. Merk verder op dat er bij de centrale examens maatregelen zijn genomen die ertoe leiden dat het moeilijkheidsniveau ervan over de jaren constant blijft. (zie Alberts, R.V.J., & Erens, B.J.M, 2012).

De toelatingseisen voor het middelbaar beroepsonderwijs (mbo) verschillen per opleidingsniveau en leerweg. Voor het volgen van een assistentenopleiding (niveau 1) gelden geen toelatingseisen. Deelnemers die de assistentenopleiding hebben afgerond moeten verder leren op niveau 2. Behalve als zij 18 jaar of ouder zijn. Voor de basisberoepsopleiding (niveau 2) gelden eveneens geen toelatingseisen. Maar als een kandidaat al een assistentopleiding heeft gevolgd voor hetzelfde beroep of dezelfde richting, zijn er wél aanvullende eisen. Kandidaten moeten dan bijvoorbeeld een bewijs hebben dat zij de eerste drie leerjaren van havo of vwo met goed gevolg hebben doorlopen, of over een vmbo-diploma beschikken of een ander erkend diploma of bewijsstuk. De toelatingseisen voor de vakopleiding (niveau 3) of de middenkaderopleiding (niveau 4) komen overeen met de aanvullende eisen die hiervoor genoemd zijn. Voor toelating tot de specialistenopleiding (niveau 4) is een diploma vereist van de vakopleiding (niveau 3), voor hetzelfde beroep of dezelfde richting.

Een cohortstudie van het CBS (2012) laat zien dat van de eerstejaars mbo van 2005/'06 10 procent na 1 jaar het mbo zonder diploma heeft verlaten. Na 4 jaar is dit percentage opgelopen naar 18 procent. Vooral vanuit het laagste niveau van het mbo, de assistentopleiding, is de ongediplomeerde uitval hoog. Zo lag die na 1 jaar op 21 procent en na vier jaar op 28 procent.

Elffers (2011) geeft een verklaring voor de lastige overgang van vmbo naar mbo. De inhoudelijke aansluiting die zij in het eerste jaar ervaren in hun mbo-opleiding beïnvloedt hun studiesucces in dat eerste jaar sterk. De verwachtingen van deelnemers spelen ook een belangrijke rol: risicoleerlingen blijken hoge aspiraties te hebben aan het begin van de opleiding. Dit blijkt de kans op uitval in het tweede semester te vergroten. Zij constateert onder meer: *'Emotionele betrokkenheid bij school, en dan met name de waardering voor de opleiding, heeft een sterk direct effect op de studievoortgang in het eerste jaar in het MBO ... Het is daarom belangrijk de emotionele betrokkenheid van MBO-studenten regelmatig te monitoren.'* (p. 173-174). Een andere belangrijke constatering is dat het kiezen van een mbo-opleiding inzicht vereist in de eigen interesses en mogelijkheden. Het blijkt in de praktijk juist hieraan én aan inzicht in de structuur van de arbeidsmarkt en het onderwijsbestel te ontbreken.

Het is daarom niet verwonderlijk dat mbo-instellingen behoefte hebben aan relevante informatie over de instromende deelnemers en dat daar op wordt ingespeeld door het aanbieden van toetsen en tests. Een eerste voorbeeld hiervan is de firma AMN (zie <http://www.amn.nl/amn-voor-scholen>). Met de Intaketest wordt de geschiktheid voor bepaalde opleidingen en bijbehorende competenties in kaart gebracht. Dit stelt de instellingen in staat om een advies op maat te geven en een geschikte begeleidingsvorm te kiezen. De diverse onderdelen van het instrumentarium meten verbale aanleg, numerieke aanleg, ruimtelijk inzicht, redeneervermogen en evaluatievermogen. Met behulp van de persoonlijkheidstest worden de intrinsieke motivatie en de persoonlijkheid van de deelnemer in beeld gebracht. Een tweede voorbeeld is de firma Bureau ICE (zie <http://www.bureau-ice.nl/>) die over een webbased toetssysteem genaamd TOA beschikt waarvan onderdelen ingezet kunnen worden bij de intake van deelnemers. Over de kwaliteit van de gehanteerde instrumenten is binnen het bestek van deze beperkte literatuurstudie geen informatie gevonden. Beide firma's leveren ook instrumenten met behulp waarvan de ontwikkeling van deelnemers in de verschillende opleidingen te volgens zou zijn.

De toets- en testpraktijk bij mbo-instellingen is vergeleken met de situatie in de Verenigde Staten op de vergelijkbare community colleges nog zeer beperkt. Geavanceerde systemen als Compass® van American College Testing (zie <http://www.act.org/products/higher-education-act-compass>) en ACCUPLACER® van de College Board (zie <http://accuplacer.collegeboard.org/>) worden daar gebruikt bij de intake van deelnemers en het volgen van deelnemers om de uitval te minimaliseren. Compass bevat een reeks computergestuurde adaptieve toetsen voor lezen, schrijven, rekenen, wiskunde en Engels als tweede taal om deelnemers optimaal te plaatsen in en voor te bereiden voor hun opleiding. Daarnaast bevat het systeem verschillende services voor deelnemende instellingen, zodat deze over actuele informatie over de samenhang tussen de scores en prestaties in de opleidingen kunnen beschikken en eventueel hun intakeprocedures kunnen aanpassen. ACCUPLACER is een vergelijkbaar systeem.

Ook op andere manieren wordt getracht om informatie te verzamelen met behulp waarvan instellingen hun rendement kunnen verhogen. Een voorbeeld hiervan is een notitie van het Kenniscentrum Beroepsonderwijs Arbeidsmarkt in opdracht van de MBO-raad (2010) die tot doel heeft de instellingen informatie aan te reiken over studiesucces van deelnemers met behulp waarvan deze kunnen zien waar zij hun prestaties op dat terrein kunnen verbeteren.

4.2. Recente ontwikkelingen

De afgelopen jaren is er de nodige discussie geweest over de kwaliteit van de examinering in het mbo. Dit heeft ertoe geleid dat er een reeks maatregelen genomen is om de examenkwaliteit te verbeteren. Zo heeft de Inspectie enkele jaren geleden de beoordeling van de examens weer overgenomen van het KwaliteitsCentrum Examinering en aangescherpt (zie Inspectie van het Onderwijs, 2011b). Bovendien is er binnen het mbo zelf ook veel meer aandacht voor dit thema gekomen. De MBO Raad en Colo hebben daarom begin 2008 het gemeenschappelijk voorstel 'standaardisering van examinering' opgesteld. Hierin stellen ze voor om te gaan werken met examenprofielen. In deze profielen maken onderwijs en bedrijfsleven afspraken over de randvoorwaarden voor examinering. Zo heeft het bedrijfsleven een stem in het examenprofiel, waardoor het vertrouwen in de mbo-examens zal toenemen. De implementatie van het voorstel is in het leerjaar 201-2011 afgerond en heeft onder meer geleid tot een zogenoemde procesarchitectuur examinering en een eenduidig begrippenkader rond examinering. Meer informatie hierover is te vinden op de website www.examinering.mbo.

In het actieplan Focus op vakmanschap (Ministerie van OCW, 2011) staan plannen om de drempelloze instroom voor mbo niveau 2 af te schaffen. Voor jongeren zonder vooropleiding zou er een entreeopleiding moeten komen ter vervanging van de huidige assistentopleiding. Deze zogenoemde entreeopleiding geeft toegang tot het mbo of bereidt jongeren die geen startkwalificatie kunnen behalen voor op de arbeidsmarkt. Om deze plannen te realiseren is wetgeving in voorbereiding. Het is de bedoeling dat de entreeopleiding op 1 januari 2014 ingaat. Enkele andere belangrijke –en voor deze notitie relevante- doelstellingen van dit plan zijn het verbeteren van de kwaliteit van de examinering, het invoeren van de centraal ontwikkelde examens voor taal en rekenen, het versterken van de oriëntatie op de loopbaan en het verbeteren van de beroepskeuzevoorlichting. Dit rechtvaardigt de conclusie dat het onderwijs in het mbo op het moment een sterke verandering aan het ondergaan is.

Tot slot mag niet onvermeld blijven dat het streven is om op 1 augustus 2014 een wettelijk kader in werking te laten treden, dat vmbo-scholen en bve-instellingen de ruimte geeft om binnen een structurele samenwerking het onderwijs zodanig in te richten dat sprake is van een doorlopende leerlijn die vanuit het vmbo zonder onderbreking leidt naar een mbo-diploma.

4.3. Evaluatie

De overgang naar het MBO is te kenschetsen als minder volwassen dan het proces zoals dat zich voltrekt bij de overgang naar het VO. Er is sprake van minder standaardisatie. Het is onduidelijk in hoeverre de instrumenten die instellingen nu inzetten bij de intake en het volgen van deelnemers aan relevante kwaliteitseisen zoals bijvoorbeeld die van de COTAN voldoen. Bovendien wordt er het nodige aan de instellingen zelf overgelaten, waardoor er verschillen kunnen ontstaan die vanuit het oogpunt van billijkheid naar kandidaten minder gewenst zijn. Dat is zeker ook te wijten aan het feit dat bij een deel van de instroom geen eisen gesteld worden. Bovendien is de sector de laatste jaren sterk aan verandering onderhevig. Door de komst van de centraal ontwikkelde examens voor taal en rekenen, de aanpassingen van de kwalificatiedossiers en de grotere nadruk op de kwaliteit van de examens is het ook lastig om overgangsbeslissingen goed vorm te geven. De opleidingen waar de overgangsbeslissingen betrekking op hebben zijn immers aan verandering onderhevig en dat

geldt ook voor de criteria op basis waarvan het welslagen van de opleidingen bepaald wordt. Het inzetten van tests van aantoonbare kwaliteit bij de intake én bij het volgen van deelnemers zou in ieder geval tot meer transparantie kunnen leiden en wellicht ook bijdragen aan het verhogen van het rendement.

De ontwikkelingen die in gang gezet zijn via het actieplan Focus op Vakmanschap, moeten bij een geslaagde implementatie kunnen leiden tot een hoger rendement dat gekoppeld gaat aan een toegenomen kwaliteit. In hoeverre dit in de praktijk het geval zal zijn, valt niet te voorspellen. Een toename in de kwaliteit van het onderwijs zou tot een sterke druk op de entree-opleiding kunnen leiden. Deze moet dan zodanig ingericht zijn dat de kandidaten die hier instromen voldoende competenties krijgen bijgebracht om –al dan niet gekwalificeerd - een positie op de arbeidsmarkt te verwerven. De druk op de entree-opleiding zal groter zijn, naarmate het in de praktijk minder goed lukt om de kwaliteit te verbeteren van het voorafgaande onderwijs voor wat betreft taal en rekenen. En dat is voor een deel weer afhankelijk van de mate waarin implementatie van de plannen rond de referentieniveaus voor taal en rekenen succesvol zal zijn. Versterken van de oriëntatie op loopbaan en verbeteren van de beroepskeuzevoorlichting zou verder tot effect kunnen hebben dat de problematiek die Elffers (2011) benoemt zal afnemen, waardoor het rendement van opleidingen toe zou kunnen nemen.

5

De overgang naar het HBO en WO

5.1. Stand van zaken

Het CBS (2012) rapporteert dat ruim 51.000 leerlingen in het schooljaar 2010/'11 in het examenjaar van de havo zaten. Aan het eind van het schooljaar had 83 procent van hen een diploma gehaald. Van de havo-geëdiplomeerden ging 78 procent in dit leerjaar naar het hbo. Van de bijna 38.000 leerlingen die in het leerjaar 2010/'11 in het zesde leerjaar van het vwo zaten heeft, bijna 87 procent een diploma gehaald. Van de vwo-geëdiplomeerden stroomde 12 procent door naar het hbo; 71 procent naar het wo en 17 procent (tijdelijk) uit het onderwijs. Vanuit de geslaagden voor een mbo-4 opleiding blijkt ongeveer 40 procent aan een hbo-studie te beginnen.

Op 3 april 1999 werd de veranderde Wet op het hoger onderwijs en wetenschappelijk onderzoek (Rijksoverheid, 1999) van kracht. De aanpassing houdt in dat onderwijsinstellingen met een toelatingsbeperking zelf voor maximaal de helft van hun plaatsen mogen bepalen op welke kwalificaties zij studenten selecteren. Voor het overblijvende gedeelte van de plaatsen is sprake van een centrale, landelijke selectie met een gewogen lotingsysteem op eindexamencijfers. Daarom mag bij de decentrale selectie het eindexamencijfer geen rol spelen. Overigens zijn studenten met een gemiddeld eindexamencijfer van een acht of hoger zeker van een plaats.

Bij de overgang van het mbo, havo en vwo naar het hbo en wo is formeel sprake van een gecombineerde beheersings/selectiebeslissing op basis van toetsresultaten. Zo bestaat een havo-examen uit een reeks toetsen voor verschillende vakken, waar een leerling cijfers voor behaalt. Dat cijfer wordt bepaald op basis van de score op iedere afzonderlijke toets en de positie van de cesuur. Of de leerling slaagt of zakt voor het examen, wordt bepaald via een beslisregel die uitgaat van de cijfers op alle toetsen tezamen. Merk op dat de beslissing gebaseerd wordt op de reeks cijfers voor het centrale examen én het schoolexamen. Merk verder op dat er bij de centrale examens maatregelen zijn genomen die ertoe leiden dat het moeilijkheidsniveau ervan over de jaren constant blijft. (zie Alberts, R.V.J., & Erens, B.J.M., 2012).

De formele toelatingseis voor het hbo is dat een kandidaat moet beschikken over een diploma mbo-4, havo of vwo. De formele toelatingseis tot het wo is het met succes afgerond hebben van een vwo opleiding, een volledige hbo-opleiding, of een bachelor-opleiding in het hbo. In het laatste geval kunnen aanvullende eisen gesteld worden. Zoals ook al in het kader van de overgang van vmbo naar mbo opgemerkt: examens zijn niet bedoeld om toekomstig studiesucces te voorspellen, maar om te bepalen of leerlingen voldoende hebben opgestoken van het voorafgaande onderwijs. Bij sommige opleidingen gelden extra eisen, zoals het gevolgd hebben van een bepaald vakkenpakket, of het bezitten van specifieke vaardigheden, kennis of eigenschappen. De extra toelatingseisen voor deze opleidingen zijn wettelijk vastgesteld. Instellingen mogen zelf bepalen of kandidaten aan de aanvullende eisen voldoen. Bij een aantal opleidingen is ook sprake van een numerus fixus. Bepaalde categorieën studenten, onder wie kandidaten met een gemiddeld eindexamencijfer van een 8 of hoger, zijn direct toelaatbaar. De overige plekken worden verdeeld door loting, selectie van studenten of een combinatie van beide.

Het bezitten van de voornoemde diploma's is in sommige gevallen dus geen voldoende voorwaarde (meer) voor toelating tot het hbo en wo. Bovendien wint het bindend studie-advies (BSA) de laatste jaren sterk aan populariteit. Van den Broek, Ribberink, Wartenbergh-Cras en 't Hart (2009) rapporteren dat maar liefst 98 procent van de hbo-opleidingen en 43 procent van de wo-opleidingen het BSA inzet. Voor het overgrote deel wordt het BSA verstrekt op basis van een minimaal door een student te behalen aantal studiepunten. Ook selectie aan de poort doet zich meer en meer voor.

Opmerkelijk is dat de pabo's een andere benadering hanteren en al een reeks van jaren gebruik maken van door het Cito ontwikkelde entreetoetsen voor taal en rekenen voor het geven van een negatief BSA. De HBO-raad heeft al in 2006 besloten tot de ontwikkeling en verplichte afname van deze reken- en een taaltoets naar aanleiding van een publicatie (Straetmans en Eggen, 2005) waaruit bleek dat meer dan de helft van de eerstejaars pabo-studenten onvoldoende rekenvaardig was. De cesuur op de toetsen is zodanig gekozen dat studenten minimaal moeten beschikken over dezelfde vaardigheid als de beste 20 procent leerlingen aan het einde van het basisonderwijs (Straetmans & Eggen, 2011).

De MBO Raad en de HBO-raad hebben afspraken gemaakt om ervoor te zorgen dat mbo'ers die instromen in het hbo voldoende goed voorbereid zijn. De doorstroom van mbo naar hbo wordt regionaal vormgegeven in samenwerkingsverbanden tussen de verschillende instellingen. Dankzij de komst van het onderwijsnummer in het mbo kunnen de instellingen voorzien worden van gedetailleerde doorstroomgegevens. Daarbij wordt in eerste instantie zichtbaar welk deel van de geslaagden per mbo-instelling en -opleiding doorstromen naar welke hogeschool en hbo-opleiding. Hieraan kan dan in een latere fase informatie over uitval van deelnemers, overstappen naar andere opleidingen en rendement worden toegevoegd. In dit kader wordt een doorstroomkaart onderhouden (HBO-raad & MBO-raad, 2009). Voor detailinformatie over het thema examinering wordt verwezen naar <http://www.examineringmbo.nl/>.

Ook in het hoger onderwijs wordt op verschillende manieren getracht om betere voorlichting aan (aankomende) studenten te verstrekken. Een onafhankelijke bron van informatie over opleidingen de website Studiekeuze 123 (zie <http://www.studiekeuze123.nl/>). Deze site is een initiatief van studenten, universiteiten en hogescholen en de Rijksoverheid. De website brengt onder meer prestaties, faciliteiten, toekomstperspectieven en studentensteden in kaart. Er is expliciete aandacht voor de kwaliteit van opleidingen en de arbeidsmarktprospectieven.

Er is veel literatuur beschikbaar die relevant is voor overgangsbeslissingen naar het hbo en het wo. Zo hebben van Weeren, Hermans en Sluijter (2009) zich in opdracht van de Onderwijsraad gebogen over de vraag op welke wijze de overgang van het vo en het mbo naar het ho soepeler kan verlopen. Zij geven een overzicht van een reeks toetsen die een rol spelen bij initiatieven om tekorten weg te werken voorafgaand aan of na de start van een ho-opleiding en doen op basis hiervan enkele aanbevelingen. Verder constateren zij dat er in de regel door instellingen geen speciale voorzieningen getroffen worden voor het bijspijkeren of opvijzelen van basale leerinhouden, wanneer kandidaten de vakken hebben gevolgd die relevant zijn voor de ho-opleiding. Er zijn echter wel veel voorzieningen voor kandidaten met een deficiënte vooropleiding. Ook geven zij aan dat naarmate de instroom in het ho heterogener wordt de kans toeneemt dat er (digitale) toetsen op brede schaal een rol zullen gaan spelen bij de intake. En dat het daarbij dan vooral zal moeten gaan om situaties van zelfselectie en/of zelfdiagnose.

Verder constateert Torenbeek (2011) ten aanzien van de aansluiting tussen vo en het universitair onderwijs dat studenten binnen opleidingen die qua opzet veel lijken op wat ze gewend waren in het vo significant meer studiepunten in het eerste jaar halen. Het betreft dan onderwijs dat net iets meer studentgecentreerd is dan in het vo. Ten opzichte van het onderwijs dat de minste studiepunten oplevert –minder studentgecentreerd dan in het vo- is het verschil maar liefs 14 punten. De aard van het type onderwijs dat gegeven wordt, houdt dus verband met de mate van studiesucces. Voorlichting over het type onderwijs dat gegeven wordt kan dus van belang zijn.

Uit Amerikaans onderzoek -een meta-analyse van 109 studies- constateren Robbins, Lauver, Le, Davis en Langley (2004) dat de beste voorspeller van studieprestaties in het hoger onderwijs het gemiddelde eindexamencijfer is of een afgeleide daarvan, zoals een centraal toelatingsexamen. Daarnaast voorspelden twee niet-cognitieve factoren goed, namelijk prestatiegerichtheid en de inschatting van de student zelf van de kans om de academische studie met succes af te ronden. Dat ook in Nederland studenten met een hoog gemiddeld vooropleidingcijfer beter presteren dan studenten met lage vooropleidingcijfers is door onderzoek aannemelijk gemaakt. Bruinsma (2003) concludeert in haar proefschrift dat het gemiddelde vooropleidingcijfer een belangrijke indicator is van studievoortgang. Uit een onderzoek van de Gruijter, Yildiz, 't Hart en & Béguin, A.A. (2005) blijkt dat de relatie tussen VWO-cijfers en studiesucces verschilt van opleiding tot opleiding. En in alle opleidingen blijken studenten met een relatief laag gemiddeld eindexamencijfer minder succesvol in de studie te zijn.

Kappe (2011) doet verslag van een onderzoek onder hbo-studenten waaruit blijkt dat kenmerken als conciëntieusheid en intrinsieke motivatie belangrijker zijn voor studieduur en studieprestaties dan intelligentie. In het onderwijs in de Verenigde Staten is het gebruikelijk toetsen in te zetten bij het selecteren van studenten voor het hoger onderwijs. Zwick, (2006) maakt duidelijk dat de betreffende toetsen het als predictor iets minder goed doen dan het gemiddelde cijfer in de vooropleiding.

Kamphorst, Hofman, Jansen en Terlouw (2009) geven aan dat ook de tevredenheid van studenten over de opleiding een voorspellende waarde blijkt te hebben, nádat is gecontroleerd op de invloed van andere relevante factoren op studiesucces. Torenbeek, Jansen en Hofman (2009) komen tot dezelfde conclusie.

Torenbeek en Kamphorst (2012) staan in algemene zin stil bij de voorspellende waarde van verschillende studentkenmerken. Over het effect van de leeftijd van studenten durven zij geen conclusie te trekken, omdat verschillende onderzoeken andere uitkomsten hebben. Vrouwen blijken beter te presteren dan mannen. Naarmate hun gemiddeld eindexamencijfer hoger ligt, blijken studenten meer studiepunten in het eerste leerjaar te behalen en sneller af te studeren. Ook de vooropleiding doet ertoe. Zo verlaten relatief veel studenten met een mbo-opleiding de opleiding voortijdig en studenten met een havo-opleiding blijken vaker over te stappen en langer te doen over de studie. Ouderlijk milieu doet er eveneens toe: studenten met hoogopgeleide ouders hebben een hogere kans op het behalen van het diploma dan de overige studenten. Ook is de deelname aan het hoger onderwijs van jongeren met hoogopgeleide ouders groter dan jongeren met laagopgeleide ouders. Tot slot melden zij dat autochtone studenten wat sneller afstuderen dan allochtone en dat autochtone studenten hogere cijfers behalen.

Ten aanzien van het thema selectie aan de poort, waar in deze notitie al eerder aandacht is besteed, zijn de meningen verdeeld. In de media verschijnen veelal positieve berichten. Zoals al eerder in

deze notitie aangegeven rapporteren de media dat selectie werkt (zie bijvoorbeeld NRC, 22 februari 2011). Verschillende onderzoeken en artikels nuanceren de positieve berichtgeving echter. Studiesucces bevorderen kan (van Berkel, Jansen & Bax, 2012), maar of selectie aan de poort en het voeren van intakegesprekken of het laten schrijven van motivatiebrieven daar nu de beste middelen voor zijn, wagen verschillende auteurs (Van Berkel, 2011; Sluijter, Straetmans & Treep, 2011) te betwijfelen. Overigens constateert Mellenbergh (in Cito, 2005) dat in zijn algemeenheid de voorspellende waarde van toelatingstoetsen relatief gering is. De minst slechte correlaties tussen toetsen en studiesucces worden gevonden bij instrumenten die 'achter de poort' worden ingezet. Resultaten op het eerste tentamen geven een minder slechte voorspelling dan toetsen aan de poort (Mellenbergh in Cito 2005, p. 26). Op zich is de geringe voorspellende waarde van dergelijke instrumenten geen verrassing, omdat er in feite al selectie heeft plaatsgevonden. Aankomende studenten hebben immers het mbo, de havo of het vwo doorlopen en hebben het eindexamen van deze opleidingen met goed gevolg afgelegd. Zij vormen dus al een homogene groep. Deze 'restriction of range' leidt er per definitie toe dat toelatingstests een lagere voorspellende waarde hebben dan bij een meer heterogeen samengestelde groep.

Een bijzondere vermelding verdient het onderzoek naar de effecten van decentrale selectie aan het Erasmus MC. (Urlings-Strop, Stijnen, Themmen & Splinter, 2009; 2011) dat een aantal jaren geleden is uitgevoerd. Binnen dat onderzoek was namelijk sprake van een unieke situatie waarin geselecteerde studenten vergeleken konden worden met ingelote studenten. Hierdoor was een onderzoeksopzet mogelijk waarin sprake was van een experimentele (de geselecteerde studenten) en een controlegroep (de ingelote studenten). De geselecteerde studenten bleken veel minder vaak uit te vallen. In dit onderzoek bleek selectie studie-uitval in de eerste vier jaar met zeventig procent te reduceren.

In het kader van selectie aan de poort worden ook regelmatig procedures genoemd als het voeren van intake-, motivatie- of sollicitatiegesprekken, interviews, het schrijven van motivatie- of sollicitatiebrieven, het laten opstellen van essays of het laten maken van voortentamens. Zonder in te gaan op de onderzoeken en verslagen die rond dit soort procedures zijn verschenen, kan in algemene zin gesteld worden dat het veel inspanning zal kosten om deze procedures zodanig in te richten dat ze bij onderzoek valide zullen blijken te zijn. Als al eerder aangegeven: selectie tot het hoger onderwijs is een high-stakes beslissing. Van iedere procedure die hierbij gebruikt wordt, zal de validiteit aangetoond moeten worden (Mellenbergh, 2011, p. 455-456). De vraag is dan ook of de moeite en kosten die het inrichten van dergelijke processen met zich meebrengt opweegt tegen een eventuele verhoging van de voorspellende waarde van het totaal aan informatie dat een rol speelt bij het nemen van de beslissing. Bovendien kunnen studenten zich concreet gaan voorbereiden op dergelijke procedures, waardoor in gevallen waar aanvankelijk sprake is van toegevoegde voorspellende waarde deze al snel weer zal verdwijnen.

5.2. Recente ontwikkelingen

Recentelijk is besloten de eindexameneisen vo te verzwaren (Rijksoverheid, 2010). Deze aanscherping van de eisen zou een negatief effect kunnen hebben op het slagingspercentage in havo en vwo, maar een positief effect voor het studiesucces in het ho. (Zie ook de eerder in deze notitie aangehaalde onderzoeken van Van Rijn, 2009 en van Van Rijn, Beguin & Verstralen, 2009). In de oude situatie was het mogelijk dat een leerling met onvoldoendes voor alle centrale

examenvakken met hoge cijfers voor het schoolexamen kon slagen. Dit leidde teveel tot calculerend gedrag van jongeren. Daarom is de zak/slaagregeling zodanig aangescherpt dat het gemiddeld centraal examencijfer onafgerond een 5,5 of hoger moet zijn. Bovendien mag voor de vakken Nederlands, Engels en wiskunde maximaal maar één 5 behaald worden op havo en vwo. De eerstgenoemde regel is op 1 augustus 2011 van kracht geworden. De tweede regel is met ingang van 1 augustus 2012 ingegaan. Een eerste analyse (DUO, 2012) laat zien dat het inwerkingtreden van de eerste maatregel geen negatief effect lijkt te hebben op het presteren van leerlingen op het centraal examen. Ook zijn er niet veel meer leerlingen gezakt en zijn er nauwelijks wijzigingen in de profielkeuze, zoals de vrees was in de discussie rond deze maatregel.

De Inspectie van het onderwijs heeft in haar rapport 'Alternatieve afstudeertrajecten en de bewaking van het eindniveau in het hoger onderwijs' (2011c) kritiek geuit op de borging van de kwaliteit van de examens en diploma's in het hbo. De Staatsecretaris van onderwijs heeft onder meer op basis hiervan in zijn strategische agenda '*Kwaliteit in verscheidenheid*' aangegeven dat het wenselijk is zowel de externe legitimering van examens beter te borgen als de toetsingspraktijk te versterken. Hij spreekt daarbij zijn voorkeur uit voor 'landelijke toetsing van ...kernvakken' en pleit voor het gebruik maken van externe examinatoren indien landelijke toetsing niet mogelijk is. Als reactie hierop heeft de HBO-raad aangegeven begrip te hebben voor het streven van de Staatssecretaris, maar beducht te zijn voor het ontstaan van kostbare, starre en bewerkelijke instrumenten. De HBO-raad stelt –terecht- dat er een breder spectrum aan vormen van externe validering mogelijk is en heeft een commissie van deskundigen ingesteld om hierover een advies op te stellen. Deze Commissie "Externe validering examenkwaliteit hoger beroepsonderwijs", beter bekend als de commissie Bruijn, heeft vervolgens het advies 'Vreemde ogen dwingen' (2012) opgesteld.

De belangrijkste adviezen uit dit rapport zijn:

- Externe validering van toetsen: waar mogelijk zouden overeenkomstige opleidingen op verschillende instellingen gezamenlijk voortgangstoetsen moeten ontwikkelen voor kennis- en vaardigheidsdomeinen die zij gemeenschappelijk hebben;
- Externe validering van eindwerkstukken door invoering van een protocol;
- Externe validering via wettelijk vastleggen van de verplichting toetsbeleid te voeren.
- Externe validering via certificering van examinatoren en opleiding van docenten als examendeskundige;
- Externe validering via toetskwaliteit. In het hbo dient systematischer onderzoek naar toetsbeleid, toetsontwikkeling en toetskwaliteit plaats te vinden.

In het rapport 'Differentiëren in Drievoud' van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel (2010), beter bekend als de Commissie Veerman, is geconstateerd dat de studie-uitval in het HO te hoog is en dat studenten te weinig uitgedaagd werden. Een derde van de studenten stopt al in het 1e jaar met de studie of switcht naar een andere studie. Eén van de belangrijkste redenen van de studie-uitval is dat zij de verkeerde studiekeuze hebben gemaakt. Dat is voor de overheid aanleiding geweest om te besluiten dat de studiekeuze in het HO beter moet worden vormgegeven; dat sprake moet zijn van een uitdagender studiecultuur en dat selectie, zowel voor als na de poort, strenger moet worden.

Met ingang van studiejaar 2012-2013 is vanwege de inwerkingtreding van de wet 'Ruim baan voor talent' (Staatsblad, 2011, nr 369) de regeling met betrekking tot de decentrale selectie aangepast.

Daardoor kan de toelating tot universiteiten en hogescholen met ingang van het studiejaar 2012-2013 selectiever van aard zijn. De selectie moet erop gericht zijn de motivatie en inzet van de student te bevorderen. Opleidingen waar de onderwijskwaliteit door te grote groepen studenten onder druk staat, kunnen hun capaciteit reguleren op basis van selectie aan de poort. Daarnaast kunnen deze opleidingen studenten nog toelaten door loting. Ook opleidingen die te typeren zijn als 'kleinschalig, residentieel onderwijs' (zoals de university colleges), mogen studenten aan de poort selecteren. Verder is ook selectie ná de poort mogelijk voor specifieke programma's binnen een opleiding. Daarnaast wordt selectie ook mogelijk gemaakt bij opleidingen die door de NVAO als 'goed' en 'excellent' zijn gekwalificeerd, of opleidingen met een scherp onderwijsprofiel, bijvoorbeeld opleidingen met een hoog eindniveau, een sterke internationale oriëntatie of met een specifiek onderwijsconcept.

De selectie van studenten moet plaatsvinden aan de hand van tenminste twee criteria die door de instellingen zijn geformuleerd. Een instelling mag dus niet alleen selecteren op eindexamencijfers. In aanvulling hierop is ook sprake van een strengere toelating van studenten met een hbo-propedeuse tot de universiteit en van mbo₄-studenten tot een hbo-opleiding. De studie-uitval onder studenten die met een hbo-propedeuse op de universiteit instromen, blijkt namelijk hoog te zijn. Daarom mogen universiteiten voortaan aan deze groep studenten ook toelatingseisen stellen. Dat geldt ook voor mbo₄-studenten. Vooral om te voorkomen dat studenten kiezen voor een opleiding die niet aansluit bij hun vooropleiding.

Verder gaat de Rijksoverheid het in pilots mogelijk maken dat instellingen ook negatief bindende studieadviezen kunnen afgeven ná het eerste jaar, bijvoorbeeld aan de hand van een vastgestelde voortgangsnorm of wanneer –in het hbo- een praktijkonderdeel onvoldoende wordt afgesloten. Bij de pilots moet echter ook sprake zijn van flankerende maatregelen die de studievoortgang van studenten bevorderen. Het betreft hier dan bijvoorbeeld intensievere studiebegeleiding en verbetering van de docentkwaliteit. Er komen meer brede bacheloropleidingen op universiteiten. Het idee hierachter is dat studenten die starten met een brede opleiding en daarbinnen zich steeds meer specialiseren, minder kans hebben om uit te vallen. Ook krijgen studenten het recht om een studiekeuzegesprek aan te vragen. Het recht is tweezijdig, dus ook de opleiding kan het van de student vragen. Als dit gebeurt, is de ander verplicht deel te nemen aan het gesprek. Om de studiekeuzegesprekken te kunnen plannen en uitvoeren hebben hogescholen en universiteiten tijd nodig. Daarvoor is het belangrijk dat studenten zich tijdig aanmelden voor een studie. De aanmelddatum voor studies wordt daarom vervroegd naar 1 mei.

5.3. Evaluatie

De overgang naar het ho is net als in het mbo minder volwassen dan het proces zoals dat zich voltrekt bij de overgang naar het VO. Er is sprake van minder standaardisatie. Sinds 2000 is meer en meer sprake van aanvullende eisen voor instromende studenten. En ook de eindexameneisen havo en vwo zijn recentelijk verscherpt. Het is onduidelijk welke effecten de aanvullende eisen die instellingen nu hanteren zullen hebben. Instellingen zijn relatief autonoom op dit vlak, waardoor er verschillen kunnen ontstaan die vanuit het oogpunt van billijkheid naar kandidaten minder gewenst zijn. Door de ontwikkelingen rond de referentieniveaus voor taal en rekenen zal het niveau van de instroom van kandidaten veranderen. Voor het hbo wordt gewerkt aan de implementatie van de adviezen van de Commissie Bruijn die tot doel hebben de kwaliteit van de examinering beter te

borgen. Veel opleidingen hanteren bindende studie-adviezen die veelal uitgaan van een minimaal behaald aantal studiepunten.

Dit zijn allemaal maatregelen die passen binnen het streven de kwaliteit van het ho te verbeteren en studenten meer uit te dagen. De overgang naar het ho en het behalen van een bachelor- of masterdiploma lijkt er de laatste jaren voor kandidaten al met al niet eenvoudiger op te zijn geworden. En de toekomstige maatregelen zullen de drempel verder verhogen. De keerzijde van deze ontwikkeling is dat het percentage studenten dat uitvalt toe zou kunnen nemen. Bemoedigend in dat kader is dat –in eerste instantie- het aanscherpen van de exameneisen bij havo en vwo geen desastreuze gevolgen blijkt te hebben gehad voor het percentage geslaagden. Desalniettemin is het van groot belang dat er flankerende maatregelen worden genomen om te voorkomen dat de uitval zal toenemen.

Dergelijke maatregelen zijn of worden ook getroffen. Het geven van betere voorlichting over de concrete inhoud van opleidingen, het verbeteren van de communicatie en aansluiting tussen aanleverende en ontvangende opleidingen en het ontwikkelen van bredere bachelors kan instrumenteel zijn in het verhogen van studiesucces.. Betere voorlichting voorkomt teleurstelling en demotivatie bij- en vervolgens afhaken van beginnende studenten voor wie de gekozen opleiding toch iets anders bleek dan verwacht. Bredere bachelors bieden studenten de mogelijkheid zich beter te oriënteren op vervolgstudies die hen werkelijk aanspreken en dus motiveren. Onderzoek zou kunnen uitwijzen dat dergelijke maatregelen een positief effect hebben op studiesucces.

Onduidelijk is in hoeverre er evidentie is voor de toegevoegde voorspellende waarde van gehanteerde aanvullende eisen, c.q. de procedures die ingezet worden bij intake en selectie van kandidaten. Van instrumenten als het intake-, of sollicitatiegesprek, het laten schrijven van een motivatiebrief en dergelijk, mag hoogstens verwacht worden dat deze een afschrikwekkende werking zullen hebben voor weinig gemotiveerde kandidaten. De inspanningen die men zich moet getroosten om een procedure te ontwikkelen die aantoonbaar voldoende kwaliteit heeft, is groot. De vraag is in hoeverre sprake zal zijn van een groei in de voorspellende waarde van de procedure bij inzet van een dergelijk instrument.

6

Overgang van bachelor naar masterfase

Met ingang van het studiejaar 2002/'03 is in het hoger onderwijs het bachelor-masterstelsel ingevoerd. Als gevolg daarvan zijn de meeste reguliere hbo-opleidingen omgezet in bacheloropleidingen van vier jaar en zijn de wetenschappelijke opleidingen opgedeeld in een bachelorfase van drie jaar en een daarop aansluitende masterfase van één jaar of langer. Om de instroom in het hoger onderwijs te vergroten kunnen studenten in het hoger beroepsonderwijs vanaf 2006/'07 ook voor een tweejarig associate degree-programma kiezen.

Een afgeronde bacheloropleiding geeft niet automatisch toelating tot een masteropleiding. Uiteraard moet er sprake zijn van voldoende aansluiting. Daarnaast worden aanvullende eisen gesteld. Deze moeten beschreven staan in de onderwijs- en examenregeling van de desbetreffende opleiding. Deze eisen mogen uitsluitend betrekking (hebben) op kennis, inzicht en vaardigheden die kunnen zijn verworven bij beëindiging van een bacheloropleiding' (Rijksoverheid, 2008). Merk op dat dit het spectrum van te gebruiken informatiebronnen beperkt. Formeel zouden persoonlijkheidskenmerken niet in ogenschouw mogen worden genomen.

De literatuursearch heeft weinig hits opgeleverd voor de overgangsbeslissing van bachelor- naar masterstudie. Hopelijk zal in veel gevallen bij overgangsbeslissingen ten behoeve van masteropleidingen sprake zijn van een benadering zoals beschreven in een artikel van Van Os (2009). Deze concludeert op basis van onderzoek binnen de Vrije Universiteit dat weliswaar substantiële verschillen bestaan tussen wo- en hbo-bachelors als het gaat om hun studiesucces en studie tempo in de masterfase, maar dat die verschillen uiteindelijk niet leiden tot masters die qua niveau verschillen. Het voornoemde artikel beschrijft verder de gevolgde werkwijze bij selectie als ook het gehanteerde selectie-instrumentarium. Op basis van een zogenoemd premasterassessment besluit zowel de opleiding als de student of deze geschikt is voor de betreffende master. Hierna volgt een premasterjaar waarin bestaande lacunes worden weggewerkt.

Het assessment bestaat uit de volgende onderdelen:

- een Cognitieve Capaciteiten Test, vergelijkbaar met een algemene intelligentietest;
- een Wiskunde & Statistiek Test;
- een Engels Tekstbegrip Test;
- een Wetenschappelijke Casus;
- een Motivatie & Leerstijl Vragenlijst, bestaande uit de subschalen 'Leerstijlen' (betekenisgericht, reproductiegericht, toepassingsgericht), 'Prestatiemotivatie', 'Extrinsieke Motivatie', 'Intrinsieke Motivatie' en 'Zekerheid Studiekeuze'.

In het betreffende artikel wordt ook verwezen naar onderzoek naar de validiteit van het instrumentarium (Van Os, 2005, 2007; Van Os & Reumer, 2006) en het artikel bevat zelf ook de nodige informatie in dit kader. De combinatie van het inzetten van een assessment van bewezen kwaliteit met een overgangsperiode waarin eventuele deficiënties weg te werken zijn, is een goede praktische manier om de kwaliteit van overgangsbeslissingen van bachelor- naar masteropleiding te borgen.

Op de website <http://universitairebachelors.nl/bachelor-naar-master/> kunnen studenten onderzoeken voor welke master-opleiding zij met een specifieke bacheloropleiding in aanmerking komen en welke aanvullende eisen er eventueel zijn. Er kan bijvoorbeeld sprake zijn van een toelatingstoets, een interview, eisen ten aanzien van de vooropleiding, een schakelprogramma, kandidaten moeten een minimaal aantal studiepunten behaald hebben voor een bepaald thema en soms staat er alleen maar vermeld dat sprake is van een 'strenge selectieprocedure'. Er zijn binnen het bestek van de literatuurstudie verder geen concrete voorbeelden gevonden van onderzoeken die evidentie opleveren over de validiteit van de gehanteerde toelatingsprocedure. Slechts in die situaties waar sprake is van voldoende aantallen studenten is het overigens mogelijk om onderzoek uit te voeren op basis waarvan goed onderbouwde conclusies te trekken zijn.

7

Slotbeschouwing en aanbevelingen

In deze notitie zijn verschillende overgangbeslissingen in het onderwijs beschouwd vanuit een methodologisch referentiekader. De literatuurstudie maakt duidelijk dat stakeholders er niet over van mening verschillen dat tests –in de meest brede zin des woords- een rol kunnen spelen om dergelijke beslissingen te verbeteren. Er wordt namelijk in het onderwijsstelsel volop gebruik van gemaakt. Bij de overgang naar het vo is sprake van een scala aan gehanteerde procedures, waar in veel gevallen de Eindtoets Basisonderwijs van Cito een belangrijke rol speelt. Er zijn echter ook verschillende andere instrumenten die in dit kader ingezet worden. Veelal zijn deze instrumenten ook van aantoonbare kwaliteit en heeft de COTAN ze op de relevante kwaliteitscriteria als voldoende of goed beoordeeld. Maar ook bij de overgang van vo naar mbo en van vo naar hbo en wo is formeel sprake van de inzet van tests, namelijk in de vorm van examens. Die zijn vanuit het oogpunt van de psychometrische besliskunde niet alleen te zien als eindcriteria voor studiesucces in een opleiding, maar ook als tests waar men de overgangbeslissing (mede) op baseert. Voor de overgang van de bachelor- naar de masteropleiding heeft deze literatuurstudie weinig informatie opgeleverd. Het valt te hopen dat bij de overgangsbepalingen procedures vergelijkbaar zijn ingericht met het beschreven onderzoek.

In deze notitie is er op gewezen dat het gehanteerde principe van het verzamelen van informatie die zo hoog mogelijk samenhangt met een criterium dat het welslagen van een behandeling weerspiegelt een statistisch probleem met zich meebrengt. Het is de kunst te werken met een set predictoren die onderling zo weinig mogelijk samenhangen en elk zo sterk mogelijk met het criterium. Dit is in de praktijk lastig te realiseren, waardoor de toegevoegde waarde van een extra predictor aan de set van predictoren snel afneemt. Zeker ook als de kosten van het ontwikkelen van een goede predictor daarbij verdisconteerd worden.

Selectie op zich kan zeker niet als zaligmakend gezien worden. Het is niet moeilijk om een hoog rendement te bereiken, maar de proportie kandidaten die afgewezen wordt, maar bij aanname succesvol geweest zou zijn, neemt toe naarmate de selectie strenger is. Het is dan ook de kunst om een goede balans te vinden tussen het bieden en het ontzeggen van kansen. De psychometrische besliskunde maakt het mogelijk om een balans te vinden die uit rationeel oogpunt optimaal is. Het probleem is echter dat beslissers hiermee in de regel niet bekend zijn en dat het lastig is om de methode in de praktijk in te zetten. Ook kunnen verschillende leden van de beslissende instantie er een andere waardenstructuur op nahouden. Bovendien hebben de actuarische procedures die daarbij vanuit de theorie ingezet zouden moeten worden in de praktijk onwenselijke systeemeffecten. Ze kunnen bijvoorbeeld leiden tot verwatering van het curriculum en de validiteit van de gehanteerde procedures ondergraven.

Het voornemen om een centrale eindtoets in te gaan voeren voor het PO zou een effect kunnen veroorzaken dat tegenovergesteld is aan het beoogde. De huidige invulling van de overgang naar VO is echter vanuit methodologisch oogpunt alleszins acceptabel te noemen. De instrumenten die ingezet worden voldoen aantoonbaar aan de eisen van betrouwbaarheid en validiteit en ongewenste systemische effecten worden tegengegaan door de beslissingen mede aan het

menselijk oordeel over te laten. Dit kan gezien worden als een balans tussen het streven naar besliskundig optimale beslissingen en de wens om de oneigenlijke effecten van een focus op testresultaten te mitigeren.

Formeel beschouwd was bij de overgang naar mbo, hbo en wo tot voor enkele jaren sprake van een vanuit besliskundig oogpunt optimale procedure. Examens bestaan immers uit een reeks tests en op basis hiervan wordt - door gebruik te maken van rekenregels- besloten of een kandidaat al dan niet geslaagd is. De resultaten van kandidaten die op grond van het in bezit hebben van hun diploma toelaatbaar zouden moeten zijn, lieten in de praktijk blijkbaar echter te wensen over. Dat is de reden dat er de laatste jaren in toenemende mate sprake is van aanvullende eisen waar kandidaten aan moeten voldoen. Het is aan de instellingen zelf om hier vorm aan te geven. In de praktijk is dan ook een veelheid van procedures aan te treffen. De vraag is of dit uit het oogpunt van billijkheid naar kandidaten een gewenste situatie is.

De manier waarop het onderwijsstelsel is ingericht maakt duidelijk dat stakeholders zich er –in ieder geval impliciet- bewust van zijn dat er bij overgangsbepalingen misclassificaties mogelijk zijn. Op verschillende momenten is al sprake van een overbruggingsperiode, of is deze in ontwikkeling. Zo is er in het vo sprake van een brugperiode die de mogelijkheid biedt om te constateren of leerlingen op hun plaats zijn. Dit geeft de mogelijkheid om het potentiële aantal misclassificaties te verkleinen. Ook in het mbo, hbo en zijn de contouren van een dergelijke vormgeving van opleidingen te ontwaren. Met name in het ho, waar niet alleen sprake is van reductie van het aantal opleidingen, maar ook van het inrichten van brede bacheloropleidingen. Dat biedt studenten de gelegenheid om een definitieve keuze uit te stellen en een gerichtere en meer gemotiveerde keuze te maken. Bovendien is in het ho het bindend studieadvies op veel plaatsen ingevoerd en vinden de nodige experimenten met selectie aan de poort plaats. Voor het nut van het inzetten van intakegesprekken of motivatiebrieven is weinig evidentie te vinden. Instellingen die dergelijke instrumenten als predictor inzetten, moeten zich realiseren dat zij bewijs zullen moeten aandragen voor de validiteit ervan. Bovendien zullen er al snel hulpmiddelen op de markt komen voor kandidaten om zich beter voor te bereiden. Dergelijke procedures kunnen hoogstens een rol spelen bij het afschrikken van minder gemotiveerde kandidaten. Het is dan echter de vraag of de kosten de implementatie van de procedure rechtvaardigen.

Het geven van goede voorlichting over de concrete inhoud van een opleiding is, naast het inbouwen van mogelijkheden om te checken of kandidaten nog op de juiste plaats zitten en remediërende programma's, een goede maatregel om een aantal onnodige misclassificaties te ondervangen.. Datzelfde geldt voor verbetering van de communicatie tussen de verschillende onderwijssectoren en het verbeteren van de aansluiting ertussen door onder meer de introductie van de referentieniveaus voor taal en rekenen. Het uitvoeren van deze activiteiten moet tezamen met de inzet van tests die aantoonbaar aan kwaliteitseisen voldoen en regelgeving die ervoor zorgt dat landelijk gezien gelijke gevallen ook gelijk behandeld worden, kunnen leiden tot verbetering van het rendement van overgangsbepalingen. Niet in de zin dat sprake is van strengere selectie, maar in de zin dat er een goede balans gevonden wordt tussen het geven en ontzeggen van kansen. Daarbij is het aan te bevelen om de samenhang tussen de informatie die gebruikt wordt bij overgangsbepalingen continu te relateren aan prestaties die leerlingen, deelnemers en studenten vervolgens leveren. Het Nederlands onderwijsstelsel is immers een dynamisch systeem dat aan verandering onderhevig is en zal blijven. Naarmate er bij een overgangsbepaling sprake is van

grotere volumes kunnen daarbij meer betrouwbare gegevens verzameld worden die in een cyclisch proces leiden tot het verder optimaliseren van landelijk of regionaal gehanteerde procedures.

Het combineren van een stelsel dat is ingericht om een kwalitatief hoogwaardig onderwijsaanbod met een hoge mate van flexibiliteit af te stemmen op de ontwikkeling van leerlingen, deelnemers en studenten met een stelsel waarin op gezette momenten sprake is van transparante en goed onderbouwde overgangsbepalingen leidt tot een goede balans tussen het geven en onttrekken van kansen en kan ertoe leiden dat het aanwezige potentieel in het Nederlands onderwijsstelsel ten volle benut wordt.

Literatuur

- 678 Onderwijs Advisering. Gevonden op 7-1-2013 op <http://www.drempelonderzoek.nl>
- American Psychological Association, American Educational Research Association & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Alberts, R.V.J., & Erens, B.J.M. (2012) *Verslag van de examencampagne 2012 voortgezet onderwijs*. Arnhem: Cito
- Angoff, W.H. (1988). Validity: an evolving concept. In: H. Wainer, & H.I. Braun (Eds.). *Test validity* (pp. 19-32). Hillsdale: Lawrence Erlbaum.
- Berkel, H.J.M. van. (1984). *De diagnose van toetsvragen*. Proefschrift. Amsterdam: Universiteit van Amsterdam.
- Berkel, H. van (2011). Selecteren aan de poort? Niet doen. *Thema*, 18, 5, 43-47
- Berkel, H. van, Jansen, E., & Bax, A. (Red.) (2012). *Studiesucces bevorderen: het kan en is niet moeilijk : bewezen rendementsverbeteringen in het hoger onderwijs ?* Den Haag : Boom Lemma
- Born, M.P. (2007). *Vruchten van selectie*. Oratie. Rotterdam: Erasmus Universiteit Rotterdam.
- Borsboom, D., Mellenbergh, G.J., & Van Heerden, J. (2004) The concept of validity. *Psychological Review*, 111, 1061-1071
- Boxtel, H. van, Engelen, R., & Wijs A. de. (2010). *Wetenschappelijke verantwoording van de Eindtoets 2010*. Arnhem: Cito.
- Centraal Bureau voor de Statistiek (2012). *Jaarboek Onderwijs in cijfers 2012*. Den Haag/Heerlen: Centraal Bureau voor de Statistiek
- Chorny, V., & Webbink, D. (2010). *The effect of accountability policies in primary education in Amsterdam*. CPB discussion Paper No 144. Den Haag: CPB.
- Cito (2005). *Schipper mag ik overvaren? Over selectie van aankomende studenten door universiteiten*. Arnhem: Cito.
- Commissie externe validering examenkwaliteit hoger beroepsonderwijs. *Vreemde ogen dwingen. Eindrapport commissie externe validering examenkwaliteit hoger beroepsonderwijs*. Den Haag: HBO-raad.
- Cook, T., & Campbell, D. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cooksey, R.W. (1988). Social judgment theory in education: current and potential applications. In : B. Brehmer & C.R.B. Joyce (Eds.). *Human Judgment: the SJT View* (Advances in Psychology volume 54) (pp 273-315). Amsterdam: Elsevier.
- COTAN. (2010). *Beoordelingssysteem voor de Kwaliteit van Tests*. Amsterdam: Nederlands Instituut van Psychologen
- Cronbach, L.J. (1988). Five perspectives on validity argument. In: H. Wainer & H.I. Braun (Eds.). *Test validity* (pp. 3-17). Hillsdale: Lawrence Erlbaum.
- Cronbach, L.J., & Gleser, G.C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana: University of Illinois press.
- Dijk, H. van & Tellegen, P.J. (2004). *De Nederlandse intelligentietest voor onderwijsniveau. Handleiding en verantwoording*. Amsterdam: Boom.

- Dienst Uitvoering Onderwijs Ministerie van Onderwijs, Cultuur en Wetenschap (2012). *Examenmonitor VO*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap
- Drenth, P.J.D., & Sijtsma, K. (2006) *Testtheorie. Inleiding in de theorie van de psychologische test en zijn toepassingen* (Vierde druk). Houten: Bohn Stafleu Van Loghum
- Eichorn, R., Janson, R.J., & van Steenberg, E.T. (2004). *Advies Voortgezet Onderwijs*. Amsterdam: Universiteit van Amsterdam.
- Elffers, L. (2011). *The transition to post-secondary vocational education: students' entrance, experiences, and attainment*. Amsterdam: Proefschrift. Universiteit van Amsterdam.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2007). *Over de drempels met taal en rekenen*. Hoofdrapport van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen. Enschede: Expertgroep Doorlopende Leerlijnen Taal en Rekenen.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18 (9), 27-32.
- Gaag, N.L. van der (1990). *Empirische utiliteiten voor psychometrische beslissingen*. Proefschrift. Amsterdam: Universiteit van Amsterdam.
- Garret, H.E. (1937). *Statistics in psychology and education*. New York: Longmans, Green.
- Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). *Clinical versus mechanical prediction: A meta-analysis*. *Psychological Assessment*, 12, 19-30.
- Gruijter, D.N.M. de, Yildiz, M., 't Hart, J., & Béguin, A.A. (2005). *Presteren in het VWO en het HO*. Deelonderzoek van experimenten met selectie: selectie op basis van vooropleidingsgegevens. Rapport 148. Leiden, ICLON
- Haertel, E.H. (2006). Reliability. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 65-110). Westport, CT: American Council on Education and Praeger Publishers.
- Hakkenes, A. & Wijs, A. de. (2012). Van Citotoets naar brugklas en door naar diploma. *Socialeconomische trends, Eerste kwartaal 2012*, 65-79.
- Hambleton, R.K., & Novick, M.R. (1973). Towards an Integration of Theory and Method for Criterion-Referenced Tests, *Journal of Educational Measurement*, 10, 3, 159-170
- HBO-raad & MBO-raad. *Doorstroom MBO-HBO*. Gevonden op http://www.hbo-raad.nl/hbo-raad/feiten-en-cijfers/cat_view/60-feiten-en-cijfers/63-onderwijs/117-doorstroom in januari 2013.
- Hemker, B.T., & Mittelhaeuser, M.A. (2012) *Optimal performance and typical performance: "What's in it for me?"*. Presentatie tijdens de achtste conferentie van de International Test Commission. Amsterdam: 3-5 juli 2012.
- Inspectie van het Onderwijs (2011a): *Onder advisering van allochtone leerlingen?* Onderzoek uitgevoerd door ITS in opdracht van de Inspectie van het Onderwijs. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2011b). *Toezichtskader BVE 2012*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het onderwijs (2011c) *Alternatieve afstudeertrajecten en de bewaking van het eindniveau in het hoger onderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2012) *Onderwijsverslag 2010/2011*. Utrecht: Inspectie van het Onderwijs.
- Joosten-Ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen*. Tilburg: Fontys Hogescholen.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge university press.
- Kane, M. (1992) An argument-based approach to validity. *Psychological Bulletin*, 112, 527-535.
- Kane, M.T. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education and Praeger Publishers.

- Kapinga, T. J. (2006). *Drempelonderzoek. Didactische plaatsbepaling binnen vmbo en praktijkonderwijs*. Ridderkerk: 678 Onderwijs Advisering.
- Kapinga, T. J. (2008). *Glashelder toelaten*. Ridderkerk: 678 Onderwijs Advisering.
- Kassenberg, A. (2005) *Reader behorende bij de FOKOR-conferentie Doorstroom po–vo*. Toetsen en observatie-instrumenten bij de doorstroom van primair naar voortgezet onderwijs. Rotterdam: FOKOR.
- Linden, W.J. van der (1985). Decision theory in educational research and testing. In: T. Husen, & T.N. Postlethwaite (Eds.). *International encyclopedia of education: Research and studies*. Oxford: Pergamon Press.
- Linden, W.J. van der (1990). Applications of decision theory to test based decision making. In: R.K. Hambleton, & J.N. Zaal (1990). *Advances in educational and psychological testing* (pp. 129-156). Boston: Kluwer.
- Linden, W.J. van der, & Mellenbergh, G.J. (1977). Optimal cutting scores using a linear loss function. *Applied Psychological Measurement*, 1, 4, 593-597.
- Linn, R.L. (1997). Evaluating the validity of assessments: the consequences of use. *Educational Measurement: Issues and Practice*, 16, 2, 14-16.
- Kamphorst, J.C., Hofman, W.H.A., Jansen, E.P.W.A. & Terlouw, C. (2009). Tevredenheid over de aansluiting als voorspeller van studiesucces in het eerste jaar hbo. *Tijdschrift voor Hoger Onderwijs*, 29, 79-92.
- Kenniscentrum Beroepsonderwijs Arbeidsmarkt (2010). *4e benchmark middelbaar beroepsonderwijs. Bouwsteen studiesucces*. Sectorrapportage. Nijmegen: Kenniscentrum Beroepsonderwijs Arbeidsmarkt in opdracht van de MBO-raad.
- Meehl, P.E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minneapolis Press.
- Mehrens, W.A. (1997). The consequences of consequential validity. *Educational Measurement: Issues and Practice*, 16, 2, 16-18.
- Mellenbergh G.J. (2011). *A conceptual introduction to psychometrics. Development, analysis and application of psychological and educational tests*. Den Haag; Eleven International Publishing.
- Messick, S. (1989). Validity. In: R.L. Linn (Ed.). *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Messick, S. (1990). *Validity of test interpretation and use*. Princeton: Educational Testing Service; Research Report 90-11.
- Ministerie van OCW (2011). *Toelatingscode overstap van vmbo naar havo*. Den Haag: Ministerie van OCW.
- Miyazaki, I. (1981). *China's Examination Hell: The Civil Service Examinations of Imperial China*. (Translated by Conrad Schirokauer). New York: Weatherhill.
- Neumann, J. von & Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. (3rd. Ed.). Princeton: Princeton University Press.
- Neuvel, J., Esch, W. van, & Westerhuis, A. (2011). *Eindexamencijfers vmbo en studiesucces op het havo*. Den Bosch: ECBO
- Novick, M.R. @ Lindley, D.V. (1978). The use of more realistic utility functions in educational applications. *Journal of Educational Measurement*, 15, 3, 181-191.
- Onderwijsraad (2010). *Vroeg of laat. Advies over de vroege selectie in het Nederlands Onderwijs*. Den Haag: Onderwijsraad.
- Os, W. van (2005). De voorspellende waarde van het VU pre-master-assessment. *Onderzoek van Onderwijs*, 34, 64-67. Genoemd in Os, W. Van (2009). Studievoortgang en studiesucces van hbo-bachelors in universitaire masterprogramma's. *Tijdschrift voor Hoger Onderwijs*, 27, 2, 127- 137

- Os, W. van (2007). Selection to the master's phase at the binary divide, a Dutch case study. *Tertiary Education and Management*, 13, 127-140. Genoemd in Os, W. Van (2009). Studievoortgang en studiesucces van hbo-bachelors in universitaire masterprogramma's. *Tijdschrift voor Hoger Onderwijs*, 27, 2, 127- 137
- Os, W. Van (2009). Studievoortgang en studiesucces van hbo-bachelors in universitaire masterprogramma's. *Tijdschrift voor Hoger Onderwijs*, 27, 2, 127- 137
- Os, W. van & Reumer, G.C. (2006). De voorspellende en diagnostische waarde van het premaster-assessment voor hbo-studenten. Paper ter gelegenheid van de Onderwijs Research Dagen 2006; Amsterdam. Genoemd in Os, W. Van (2009). Studievoortgang en studiesucces van hbo-bachelors in universitaire masterprogramma's. *Tijdschrift voor Hoger Onderwijs*, 27, 2, 127- 137
- PO-raad, AVS en VO-raad (2011). *Effectief Schakelen: verbeteren van de informatieoverdracht tussen PO en VO*. Utrecht: PO-raad, AVS en VO-raad
- Popham, W.J. (1997). Consequential validity: right concern - wrong concept. *Educational Measurement: Issues and Practice*, 16, 2, 9-13.
- RCEC (2011). *Beoordelingssysteem voor de kwaliteit van toetsen en examens*. Enschede: Research Center voor Examinering en Certificering.
- Rijksoverheid (1993). *Inrichtingsbesluit WVO*. Gevonden op: <http://wetten.overheid.nl/BWBR0005946> in december 2012.
- Rijksoverheid (1999). *Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek*. Gevonden op <http://wetten.overheid.nl/BWBR0005682/> in december 2012.
- Rijksoverheid (2008). *Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek, artikel 7.30b lid 1*. Gevonden op <http://wetten.overheid.nl/BWBR0005682/> in december 2012.
- Rijksoverheid (2010). *Wijzigingsbesluit Staatsexamenbesluit VO, enz. (aanscherping slaag/zakregeling voortgezet onderwijs*. Gevonden op <http://wetten.overheid.nl/BWBR0028184> in januari 2013.
- Rijksoverheid (2011). *Wet referentieniveaus Nederlandse taal en rekenen*. Gevonden op <http://wetten.overheid.nl/BWBR0027679> in december 2012.
- Rijksoverheid (2012). *Voorstel van wet in verband met de invoering van een centrale eindtoets*. Gevonden op <http://www.rijksoverheid.nl/ministeries/ocw/documenten-en-publicaties/kamerstukken/2012/01/31/> in december 2012.
- Rijn, P. W. van (2009). Mogelijke effecten van verschillende uitslagregels. *Examens*, 6, 3, 5-8.
- Rijn, P.W. van, Béguin, A.A., & Verstralen, H.H.F.M. (2009). Zakken of slagen? De nauwkeurigheid van examenuitslagen in het voortgezet onderwijs. *Pedagogische Studiën*, 86, 185-195.
- Robbins, S.B., Lauver, K., Le, H., Davis, D., & Langley, R. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130 (2), 261-288.
- Sanders, P.F. (2011) De betrouwbaarheid van toetsscores. In: P.F. Sanders (Red.) *Toetsen op School*. Arnhem: Cito
- Shepard, L.A. (1993). Evaluating test validity. In: L. Darling-Hammond (Ed.). *Review of research in education: Vol 19* (pp.405-450). Washington, DC: American Educational Research Association
- Shepard, L.A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16, 2, 5-8.
- Sluijter, C (1998). *Toetsen en beslissen. Toetsing bij doorstroombeslissingen in het voortgezet onderwijs*. Proefschrift. Amsterdam: Universiteit van Amsterdam.
- Sluijter, C., Straetmans, G.J.M. & Treep, J. (2012) 'Selectie aan de poort; zinvol of onzin?' *Tijdschrift voor Onderwijsinnovatie*, 13, 4, 27-29
- Staatblad, nr. 369 (2011) *Wet van 8 juli 2011 tot wijziging van de Wet op het hoger onderwijs en wetenschappelijk onderzoek in verband met het uitbreiden van de mogelijkheden van selectie van studenten en van verhoging van het collegegeld alsmede in verband met het aanscherpen van de*

toelatingsvereisten voor aansluitende masteropleidingen in het wetenschappelijk onderwijs (ruim baan voor talent).

- Staatscourant 9 mei 2012, nr. 9019. *Regeling van de Minister van Onderwijs, Cultuur en Wetenschap van 25 april 2012, nr. WJZ/390987*10201.,
- 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. (2011). WISCAT-pabo: Ontwerp, kwaliteit en resultaten van een geruchtmakende toets. In: Schramade e.a. (Red.) *Handboek Effectief Opleiden, sectie 9-2.2.* 's Gravenhage: Reed Business Information BV.
- Stroucken, Takkenberg & Beguin, A. (2008). Citotoets en de overgang van basisonderwijs naar voortgezet onderwijs. *Sociaaleconomische Trends, tweede kwartaal 2008*, 7-16.
- Suen, H.K., & Wu, Q. (2006). The Keju exam system from a psychometric perspective. In Liu, H.F. (Ed.) *Selected anthology of the International Conference on Imperial Examinations*. Xiamen, China.
- Têng, S. (1943). Chinese Influence on the Western examination system. *Harvard Journal of Asiatic Studies*, 7, 267-312.
- Torenbek, M., Jansen, E.P.W.A., & Hofman, W.H.A. (2009). Aansluiting vwo-universiteit: het effect op studiesucces in het eerste jaar. *Tijdschrift voor Hoger Onderwijs*, 29, 93-102.
- Torenbek, M. (2011) *Hop, skip and jump? : the fit between secondary school and university*. Proefschrift. Enschede: Universiteit Twente.
- Torenbek, M., & Kamphorst, J.C. (2012). Studiesucces in het eerste jaar: de invloed van de aansluiting tussen voortgezet onderwijs en hoger onderwijs en achtergrondkenmerken van studenten. In: H. Van Berkel, E.P.W.A. Jansen & A. Bax. *Studiesucces bevorderen: het kan en is niet moeilijk. Bewezen rendementsverbeteringen in het hoger onderwijs*. Den Haag: Boom Lemma
- Urlings-Strop, L.C., Stijnen, T., Themmen, A.P.N., & Splinter, T.A.W. Selection of medical students: a controlled experiment. *Medical Education* 2009: 43 (2), 175-183
- Urlings-Strop, L.C., Stijnen, T., Themmen, A.P.N., & Splinter, T.A.W. Selected medical students achieve better than lottery-admitted students during clerkships. *Medical Education* 2011; 45 (10), 1032-1040
- Weeren, J van, Hermans, P., & Sluiter C. (2009). *Referentieniveaus en aanvang hoger onderwijs. Bijspijkeren bij de overgang naar hoger onderwijs*. Arnhem: Cito. In opdracht van de Onderwijsraad
- Zwick, R. (2006). Higher Education Admissions Testing. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 647-679). Westport, CT: American Council on Education and Praeger Publishers.