

# Het rendement van selectie aan de poort

Psychologische en cognitieve selectiecriteria zijn een goede basis voor het voorspellen van studiesucces. Het kabinet-Rutte lijkt de juiste beleidskeuze te hebben gemaakt met meer aandacht voor wiskunde in de vooropleiding en de acceptatie van selectie aan de poort in het bekostigd onderwijs.

## ERIC MELSE

Bedrijfsconoom en voormalig programmadirecteur van de masteropleiding aan de Nyenrode Business Universiteit

## KEES VAN MONTFORT

Hoogleraar aan de Erasmus Universiteit Rotterdam en aan de Nyenrode Business Universiteit

## JAAP VAN MUIJEN

Rector van SIOO en hoogleraar aan de Nyenrode Business Universiteit

**H**et kabinet-Rutte besloot vorig jaar om selectie aan de poort te bevorderen in het hoger onderwijs. Hiermee volgt het kabinet de aanbeveling op van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel (Veerman, 2010). Kandidaten die een studie willen volgen, mogen voortaan door hogescholen en universiteiten worden getoetst op geschiktheid en motivatie. De hoop en verwachting is dat hiermee ook opportunistisch studiegedrag – getypeerd als de ‘zesjes-cultuur’ – wordt bestreden in het onderwijs. Iemand die een bachelor- of master-opleiding wil volgen aan een ‘excellent instituut’ moet al eerder op topniveau presteren om überhaupt te worden toegelaten (Coppoolse *et al.*, 2010). Daarnaast moeten studenten fors meer collegegeld gaan betalen als zij een opleiding niet behalen binnen de daarvoor gestelde studietijd. Wil iemand een tweede studie volgen dan wordt voortaan de volle prijs gevraagd. Het kabinet tracht met deze drukmiddelen te voorkomen dat de kwaliteit van het bekostigd onderwijs in Nederland nog verder erodeert (Bronneman-Helmers, 2009; Van der Steeg *et al.*, 2011).

De verwachting is dat door een strengere selectie sprake zal zijn van hoger rendement voor de individuele student, het opleidingsinstituut en meer baten voor de kenniseconomie. Noodzakelijk is dan wel dat selectie aan de poort effectief is. Inzicht daarin kan worden verkregen door de effectiviteit te bekijken van de gehanteerde selectiecriteria.

## BESTAANDE TOELATINGSPROCEDURES

Verschillende Nederlandse universiteiten maken bij hun toelatingsprocedure gebruik van psychometrische testen om eigenschappen te onderzoeken van studenten, zoals persoonlijkheid, professionele belangstelling en motivatie. Bijvoorbeeld voor de geneeskundeopleiding aan de Erasmus Universiteit (Crommentuyn, 2010) en aan de Universiteit Leiden (Urlings-Strop *et al.*, 2009). De Vrije Universiteit Amsterdam (VU) hanteert voor de premaster-assessment van hbo-bachelors een cognitieve capaciteitentest en een motivatie- en leerstijltest (Reumer, 2009). Ook toetst de VU schrijfvaardigheid en rekenkundige vaardigheden.

Het verklarend en voorspellend vermogen van zulke testen is onderwerp van studie voor academisch en economisch succes (Almlund *et al.*, 2011). Om te bepalen in hoeverre psychometrische testen studieresultaten kunnen voorspellen, onderzochten Lievens *et al.* (2009) het effect van een persoonlijkheidstest voor Belgische studenten geneeskunde die in 1997 aan hun studie begonnen. Naarmate de studie vordert, blijkt dat de predictieve validiteit toeneemt van persoonlijke eigenschappen; een belangrijk gegeven bij het volgen van de academische en professionele ontwikkeling. Kappe en Van der Flier (2010) onderzochten de predictieve validiteit van persoonlijkheidskenmerken voor studiesucces van vier cohorten aan de Hogeschool InHolland. Afhankelijk van de leersituatie blijkt dat de bijdrage van persoonlijke eigenschappen verschilt bij de voorspelling van het studieresultaat.

Met name voor masteropleidingen beperken onderwijsinstellingen zich vaak bij selectie tot het meten van cognitieve capaciteiten met de internationaal gehanteerde GMAT (Graduate Management Admission Test). Ofschoon eerder onderzoek het effect aantoonde van persoonlijkheid en motivatie op studieresultaat, zijn veel studies beperkt in omvang en in tijd (Almlund *et al.*, 2011). Aan de hand van regressiemodellen kan het rendement van selectie aan de poort door Nyenrode Business Universiteit worden bekeken met een groot cohort gedurende de afgelopen tien jaar.

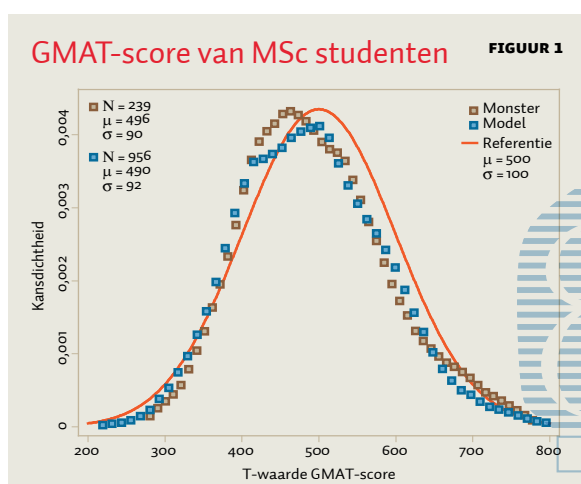
De auteur heeft verklaard dit artikel alleen te publiceren in ESB en niet elders te publiceren in wat voor medium dan ook. Het is wel toegestaan om het artikel voor eigen gebruik en voor publicatie op een intranet van de werkgever van de auteur aan te wenden.

**NYENRODE**

Nyenrode Business Universiteit selecteert aan de poort en hanteert de Nyenrode Admission Test (NAT) voor toegang tot het Master of Science (MSc) in management (De Groot, 2008). Naast onderwijskundige entreevoorwaarden als een hbo- of universitair bachelordiploma is wiskunde in de vooropleiding wenselijk. Als dit ontbreekt moet een kandidaat een zomercursus volgen. Daarnaast wordt elke kandidaat getest op cognitieve capaciteiten, persoonlijkheidskenmerken en werkstijl. Talrijke ondernemingen gebruiken vergelijkbare testen bij het werven en selecteren van hun personeel (Drenth, 2010). De meting van cognitieve capaciteiten geeft een goede indicatie van het analytisch vermogen en dit blijkt een sterke voorspeller van succes tijdens studie en loopbaan; zeker naarmate de leeftijd toeneemt (Resingh en Drenth, 2007). Maar intelligentieverschillen verklaren slechts voor maximaal de helft de verschillen in academische prestaties.

Om mensen te karakteriseren aan de hand van gedrag dat zij vertonen, wordt op Nyenrode een persoonlijkheidstest gebruikt met het Five Factor-model, ook wel de Big 5 genoemd (Van Muijen, 2005). Dit is een wijdverspreid en geaccepteerd persoonlijkheidsmodel waarbinnen mensen verschillen op de kenmerken openheid, consciëntieusheid, extraversie, altruïsme en neuroticisme (Carver en Scheier, 2011). Met de werkstijlkenmerken ondernemingszin, streven naar succes en streven naar kwaliteit wil Nyenrode een beeld krijgen van wat kandidaten belangrijk vinden, welke ambities ze hebben en hoe ze een doel bereiken.

Vermeden moet worden om de testresultaten te interpreteren in termen van categorieën of typen. Afgezien van het gegeven dat een bepaalde score hoog, normaal of laag is: zij zijn in principe allemaal informatief. Ook is er sprake van compenserende effecten tussen de domeinen. Iemand hoeft niet noodzakelijk een hoge GMAT-score te hebben om zeker te zijn van een redelijke kans op een goede studieprestatie. Dergelijke nuances worden onderkend door de predictieve validiteit van alle tijdens de NAT gemeten kenmerken te bestuderen voor het gemiddeld premaster-eindcijfer. Het doel van de premaster is om bachelors de gelegenheid te bieden aanvullende bedrijfseconomische kennis te verwerven en academische vaardigheden te ontwikkelen voordat zij worden toegelaten tot de eigenlijke masteropleiding.

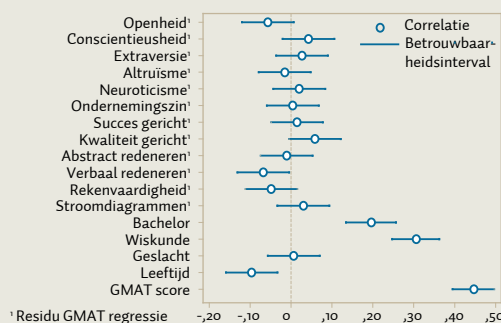
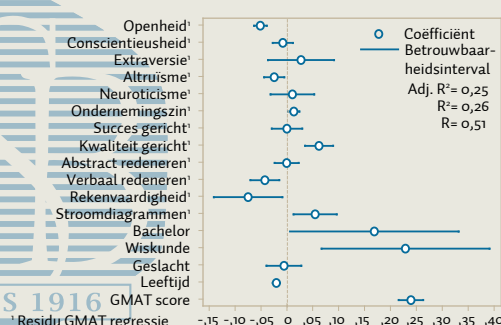
**COHORT**

Het bestudeerde cohort omvat 1195 studenten die in de jaren 2002–2010 werden toegelaten, onder wie 875 mannen en 320 vrouwen. Hiervan slaagde 8,5 procent niet voor de premaster en die verlieten daarom voortijdig de opleiding. 83,5 procent van de hbo-bachelors op Nyenrode vervolgden hun studie op Nyenrode, van wie meer dan de helft wiskunde in hun opleiding kreeg (53,9 procent). Het aantal wo-bachelors met wiskunde is met 84,3 procent beduidend groter. Van de 1094 studenten die de premaster wel voltooiden, behaalden 500 een bovengemiddeld en 594 een benedengemiddeld eindcijfer, respectievelijk 45,7 en 54,3 procent.

Ten behoeve van kruisvalidering is via randomisatie het cohort gesplitst in twee niet overlappende deelsteekproeven. De regressiemodellen zijn gemaakt met tachtig procent van het cohort: het modeldeel (tabel 1; figuur 2). Met de overige twintig procent is de predictieve validiteit getest: het monsterdeel (figuur 3). De psychometrische variabelen zijn bewerkt in enkelvoudige regressiemodellen met de GMAT-score als onafhankelijke variabele zodat collineariteit wordt beperkt (Grimm en Yarnold, 1995; Pindyck en Rubinfeld, 2000). Vervolgens worden de residuscores van de testen voor cognitieve capaciteiten, werkstijl en persoonlijkheidskenmerken in het model beschouwd als onafhankelijk van de GMAT-score, maar niet van leeftijd, geslacht, de gevolgde bachelor en wiskunde in de vooropleiding.

**GMAT**

Juist bij Nederlandse cohorten meent Drenth (2004a; 2004b) dat de correlatie tussen psychologische variabelen en studieprestaties nauwelijks een bruikbaar niveau haalt, aangezien er

**Nyenrode selectievariabelen****FIGUUR 2****2a. Correlaties met met het premaster-eindcijfer****2b. Coëfficiënten van het multiple lineaire regressiemodel**

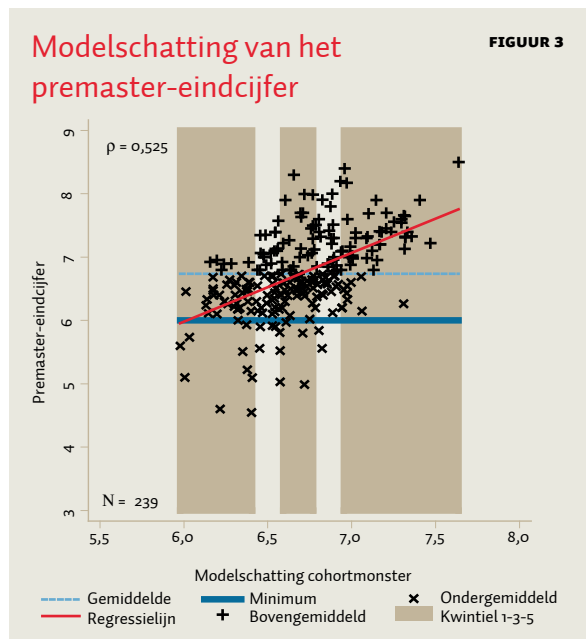
sprake zou zijn van beperking van de variatie tussen de laagste en de hoogste score. Dat wil zeggen dat variatiebeperking ontstaat door eerdere selecties in het gedifferentieerd secundair onderwijs. Als gevolg van dergelijke variatiebeperking loopt men bij onderzoek het risico dat de statistische betrouwbaarheid en validiteit te laag worden geschat (Drenth, 2010).

Bij de NAT wordt een referentiepopulatie gebruikt van academici uit de Nederlandse beroepsbevolking volgens een normale verdeling: de doorgetrokken lijn in figuur 1. Aangezien de kansverdeling van de GMAT-score de kansdichtheidskromme volgt van de referentiepopulatie, is er duidelijk geen sprake van variatiebeperking bij het Nyenrode-cohort.

Het enkelvoudige regressiemodel verklaart twintig procent van de variantie, wat wordt beschouwd als een zwak verband tussen de GMAT-score en het eindcijfer (tabel 1, model 1).

**PREDICTIEVE VALIDITEIT**

De volgende stap is het analyseren van de predictieve validiteit van alle gemeten variabelen om vast te stellen dat selectie rendert voor de student en voor het instituut. Voor het eerste wordt gekeken of de student een boven- of benedengemiddeld eindcijfer ontvangt, voor het laatste of de uitstroom laag is. Predictieve validiteit wordt gebruikelijk vastgesteld door de mate van associatie tussen de gemeten voorspellers en het criterium, hier het eindcijfer. In figuur 2a is de correlatiecoëfficiënt met betrouwbaarheidsinterval weergegeven voor het modeldeel van het cohort. Hieruit blijkt dat wiskunde en type bachelor (hbo of wo) positief associëren met het eindcijfer, terwijl leeftijd negatief associeert. Dat wil zeggen dat de lijn die de betrouwbaarheidsinterval weergeeft van een variabele



niet de nullijn doorkruist. Van de psychologische variabelen associëren openheid en kwaliteit met het eindcijfer, ofschoon de betrouwbaarheidsinterval de nullijn wel raakt. Figuur 2b toont de coëfficiënten van het meervoudige regressiemodel (tabel 1, model 2). Ook andere psychologische variabelen dragen significant bij: altruïsme en ondernemingszin, en verder verbaal redeneren, rekenvaardigheid en stroomdiagrammen. Dat verbaal redeneren en rekenvaardigheid negatief bijdragen in het meervoudige regressiemodel, volgt mogelijk uit het gegeven dat hier het residu is gebruikt van de enkelvoudige regressie op de GMAT-score.

Vergelijken we het meervoudig met het enkelvoudig regressiemodel, dan blijkt er sprake van een verbetering van de correlatiecoëfficiënt R: 0,51 versus 0,45. Daarmee verbetert ook de proportie verklaarde variantie R<sup>2</sup>: 0,26 versus 0,20. Dit meervoudige regressiemodel drukt daarmee een matig verband uit tussen de bij selectie gemeten kenmerken en het eindcijfer. Daarmee is nog niet gezegd dat het risico op uitval ook kleiner wordt. In figuur 3 staan ook punten onder de minimumlijn in het tweede, derde en vierde GMAT-kwintiel van de modelschatting (aangegeven door de verticale blokken). Dat zijn de studenten die ondanks een hoge GMAT-score toch de opleiding verlieten.

De predictieve validiteit van zowel de NAT-testvariabelen individueel als de gepresenteerde modellen is in overeenstemming met onderzoeksresultaten in het buitenland, of beter (Lievens *et al.*, 2009). Echter, zelfs met een meervoudige correlatiecoëfficiënt van 0,51 is nog kritiek denkbaar op nut en noodzaak van selectie aan de poort. Hoewel uit de literatuur blijkt dat de correlatie zelden hoger ligt dan 0,35 (Sackett *et al.*, 2008) is de correlatie voor Nyenrode niet bijzonder hoog met 0,45 voor GMAT. Daarom gaat de aandacht vervolgens uit naar de beneden- en bovengemiddelde spreiding.

**Regressiemodellen voor het gemiddelde premaster-eindcijfer**

**TABEL 1**

N=956	Model 1	Model 2
Openheid**	-	-0,052
Conscientieusheid	-	-0,009
Extraversie	-	0,026
Altruïsme*	-	-0,025
Neuroticisme	-	0,010
Ondernemingszin*	-	0,013
Succes gericht	-	-0,001
Kwaliteit gericht**	-	0,061
Abstract redeneren	-	-0,001
Verbaal redeneren**	-	-0,043
Rekenvaardigheid*	-	-0,075
Stroomdiagrammen*	-	0,054
Bachelor*	-	0,167
Wiskunde**	-	0,227
Geslacht	-	-0,006
Leeftijd**	-	-0,021
GMAT score**	0,282	0,237
Adjusted R <sup>2</sup>	0,199	0,249

\*/\*\* Significant op respectievelijk vijf- en éénprocentniveau

**SPREIDING**

Voor Reumer en Van der Wende (2010) en Drenth (2004a) is een voorspellende validiteit van 0,50 genoeg reden om te twij-

felen aan het gebruik van psychometrische testen bij selectie in het onderwijs. Belangrijker is echter de objectiviteit van de modelschatting ten opzichte van het eindcijfer: figuur 3. Gegeven de correlatie van 0,525 daartussen moet het onderscheid tussen beneden- en bovengemiddelde prestatie overeenkomen met de statistische verwachting. In de puntenwolk zijn diagonaal de punten gelijkmatig verdeeld ten opzichte van de regressielijn. Horizontaal verschilt per kwintiel wel het aantal benedengemiddelde kruisjes en de bovengemiddelde plusjes. In het eerste kwintiel zijn meer kruisjes en in het vijfde kwintiel meer plusjes. Zowel de verwachte afwijkingen diagonaal ten opzichte van de regressielijn, als horizontaal per kwintiel ten opzichte van het gemiddelde, zijn afhankelijk van de correlatiecoëfficiënt.

Objectiviteit en rendement van selectie aan de poort kan worden vastgesteld aan de mate waarin de spreiding van het aantal kruisjes en plusjes per kwintiel overeenkomt met de statistische verwachting (Myors, 1994). Sommige van de studenten die de premaster niet voltooiden, staan ver onder de minimumwaarde van 6 voor toelating tot de masteropleiding: figuur 3. De afwijkingen diagonaal ten opzichte van de regressielijn zijn daarmee zo groot dat de correlatie vanzelfsprekend lager is dan wanneer alle studenten wel hun premaster succesvol zouden afronden. Maar ook onvoorspelbare zaken hebben invloed in het onderwijs en daarom is succes niet gegarandeerd, ook niet na selectie aan de poort. Wel kan objectief worden bepaald in welke mate de beneden- en bovengemiddelde spreiding afwijkend is. Uit een meervoudige t-test blijkt dat deze afwijking niet significant is.

De modelmatige analyse van de studieprestatie, van de maatstaf van individueel rendement, en ook van uitval als maatstaf van het institutioneel rendement, weerlegt in zekere mate de kritiek dat psychometrische testen geen basis kunnen zijn voor het voorspellen van studiesucces. Dergelijk inzicht draagt bij aan het verbeteren van zowel de selectieprocedure als de begeleiding van studenten met een hoger risicoprofiel tijdens de studie (Polder en Baten, 2010). Daarbij gaat het niet alleen om hoge rapportcijfers. Het profiel van de student is een referentiepunt van waaruit de opleiding en vorming begint van het individu en de groep als geheel. Pluriformiteit in persoonlijkheid, intelligentie en werkstijl bevorderen de interactie tussen studenten en daarmee het leerproces dat op Nyenrode is gericht op ondernemerschap en leiderschap, maar ook op samenwerking en inlevingsvermogen.

## CONCLUSIE

Psychologische kenmerken van studenten hebben invloed op hun studieprestaties. Het studieresultaat blijkt beter te voorspellen als deze meeweegt bij de selectie in plaats van uitsluitend met de GMAT-score. Op Nyenrode komt de voorspelling van het premaster-eindcijfer van studenten overeen met wat boven- en benedengemiddeld wordt verwacht. Dat het studieresultaat gemiddeld hoger ligt voor studenten die eerder wiskunde kregen, is wellicht een aanmoediging om het onderwijsbeleid van het kabinet-Rutte op dit punt te omarmen. Omdat elke vorm van selectie tijd en moeite vergt van de kandidaten en de staf die de resultaten moet beoordelen, is niet zomaar gezegd dat selectie aan de poort rendabel zal zijn voor elke opleiding in het hoger onderwijs.

## LITERATUUR

- Almlund, M., A.L. Duckworth, J.J. Heckman en T.D. Kautz (2011) Personality psychology and economics. *NBER Working Paper*, 16822.
- Bronneman-Helmers, H.M. (2009) De kwaliteit van het onderwijs. *Jaarboek Overheidsfinanciën*, 123–139.
- Carver, Ch.S. en M.F. Scheier (2011) *Perspectives on personality*. Upper Saddle River, NJ: Prentice Hall.
- Coppoolse, R., P. van Eijl en E. de Bruin (2010) Naar een instellingsbrede excellentiestrategie in het hbo: een onderzoek naar de ontwikkeling van talent. *Tijdschrift voor Hoger Onderwijs*, 28(4), 217–228.
- Crommentuyn, R. (2010) Decentrale selectie ter discussie. Tegenstanders van screenen aan de poort in het nauw. *Medisch Contact*, 65(9), 380–382.
- Drenth, P.J.D. (2004a) Selectie aan de poort werkt niet. *NRC Handelsblad*, 8 april.
- Drenth, P.J.D. (2004b) Selectie aan de poort van het hoger onderwijs. *Tijdschrift voor hoger onderwijs & management*, 11(4), 48–51.
- Drenth, P.J.D. (2010) *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn Stafleu Van Loghum.
- Grimm, L.G. en P.R. Yarnold (1995) *Reading and understanding multivariate statistics*. Washington: American Psychological Association.
- Groot, B. de (2008) Selectie aan de poort. Match student en studie is in ieders belang. *Het Financieel Dagblad*, 16 augustus.
- Kappe, R. en R.H. van der Flier (2010) Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic performance. *Journal of Research in Personality*, 44(1), 142–145.
- Lievens, F., D.S. Ones en S. Dilchert (2009) Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, 94(6), 1514–1535.
- Myors, B. (1994) A theoretical expectancy table calculator. *Behavior Research Methods*, 26(4), 467–469.
- Pindyck R.S. en D.L. Rubinfeld (2000) *Econometric Models and Economic Forecasts*. New York: McGraw-Hill.
- Polder K.-J. en J.W. Baten (2010) Gepercipieerde verklaringen voor lage en hoge rendementen in het hoger onderwijs. *Tijdschrift voor Hoger Onderwijs*, 28(3), 153–163.
- Resingh, W.C.M. en P.J.D. Drenth (2007) *Intelligentie: weten en meten*. Amsterdam: Nieuwezijds.
- Reumer, C. (2009) *Ervaringen met het premasterassessment op de Vrije Universiteit tussen 2003 en 2009*. Masterclass Digitaal Toetsen aan de SURF Academy. Amsterdam: Vrije Universiteit Amsterdam.
- Reumer, C. en M. van der Wende (2010) Excellence and diversity. The emergence of selective admission policies in Dutch higher education. *Center for Studies in Higher Education Research & Occasional Paper Series*, 15(10).
- Sackett, P.R., M.J. Borneman en B.S. Connelly (2008) High-stakes testing in higher education and employment. Appraising the evidence for validity and fairness. *American Psychologist*, 63(4), 215–227.
- Steege, M. van der, N. Vermeer en D. Lanser (2011) *Nederlandse onderwijsprestaties in perspectief*. Achtergronddocument bij CPB Policy Brief 2011(5). Den Haag: Centraal Planbureau.
- Urlings-Strop, L.C., T. Stijnen, A.P. Themmen en T.A. Splinter (2009) Selection of medical students: a controlled experiment. *Medical Education*, 43(2), 175–183.
- Muijzen, J.J. van (2005) Leiderschapsontwikkeling: het hanteren van paradoxen. In: Schramade, P.W.J.M. (red), *Handboek effectief opleiden*. Den Haag: Reed Business Information, 5.11-4.01–15.11-4.44.
- Veerman, C.P. (2010) *Differentiëren in drievoud*. Advies van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel. Den Haag: Commissie Toekomstbestendig Hoger Onderwijs Stelsel.