

Toetsen getoetst

Het begrijpen van het langzaam verlaagde niveau van
het vwo Centraal Examen (CE) Wiskunde en Natuurkunde.



Februari 2023

Auteurs

ir. Loek Zonnenberg, MBA
(docent wiskunde en Partner alumnus)
ir. Paul Rutten, MBA (Partner)

Colofon

Dit rapport is tot stand gekomen in samenwerking tussen 2 partijen: Loek Zonnenberg, docent wiskunde en voormalig Partner bij McKinsey & Company en Paul Rutten, Partner bij McKinsey & Company, op basis van hun gedeelde passie voor onderwijs. Ze hebben hun rapport uitgebracht onder de vlag van McKinsey.

Dit rapport is onafhankelijk, en reflecteert de inzichten van de auteurs, en is niet geschreven in opdracht van enig bedrijf, de overheid of ander instituut.

McKinsey & Company is een toonaangevend adviesbureau voor strategisch management, gespecialiseerd om instellingen in de publieke, private en sociale sector te helpen blijvende successen te behalen. McKinsey is al bijna een eeuw de leidende adviseur van overheden, organisaties en bedrijven wereldwijd en helpt hen om de meest dringende problemen op te lossen en om belangrijke, duurzame prestatieverbeteringen te behalen. Het besturingsmodel van McKinsey als een wereldwijd, privaat partnerschap waarborgt haar onafhankelijkheid en objectiviteit.

De auteurs bedanken de vele tientallen leraren, leerlingen, betrokkenen, experts, instanties en (oud-)collega's die mee hebben gedacht over de opzet van het onderzoek. In het bijzonder bedanken ze ook de leerlingen van het Maerlant-Lyceum te Den Haag die mee hebben gedaan aan het onderzoek door het maken van examenvragen. Daarnaast bedanken ze in het bijzonder Jeroen Gaudissaboïs, Iris van Beek, Janneke Woortmeijer en Sanne Basjes voor hun hulp.

Bij vragen kan contact opgenomen worden via Benelux-communications@mckinsey.com

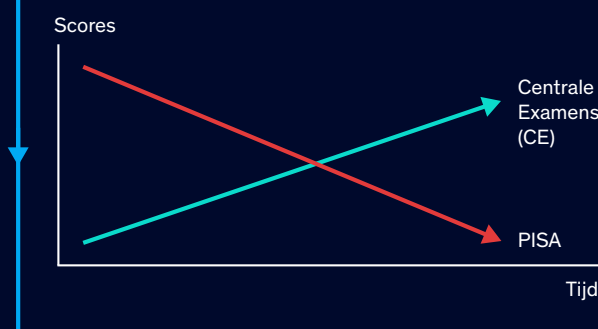
ISBN : 9789083265117

Het onderzoek in één oogopslag

Toetsen getoetst

Hoe kunnen we de paradox van de niveaudaling in het onderwijs verklaren?

Wat stellen we vast voor Wiskunde en Natuurkunde?



- 1 De vwo CE cijfers Wis- en Natuurkunde zijn de afgelopen 30 jaar gestegen
- 2 De PISA scores voor Wiskunde en Natuurwetenschappen dalen de afgelopen 20 jaar
- 3 Ook andere nationale en internationale onderzoeken duiden op achterblijvend niveau

Wat zou de oorzaak kunnen zijn?



PISA is geen adequate ontwikkelingsniveaumeting?



Er is geen verband te verwachten tussen PISA en het CE?



Het niveau van de vwo CE's is de afgelopen 30 jaar gedaald?



Nieuw onderzoek wijst naar drie grandoorzaken

1 CE's bevatten in de afgelopen jaren fors minder stof

De CE's toetsen nu 40-50% minder stof dan 30 jaar geleden

Ook de School Examens bevatten minder stof

2 De CE's van nu zijn makkelijker:

Minder denkstappen nodig om een vraag te beantwoorden

Meer makkelijke en minder moeilijke vragen

Bij Wiskunde is vooral de moeilijkere stof verdwenen

3 Het niveau van de vwo-populatie is gedaald

Minder goede landelijke scores nu op bijna identieke examenvragen dan 30 jaar geleden door:

- a Meer kandidaten: het aandeel vwo CE-kandidaten is 30 procent gestegen
- b Minder goed voorbereide vwo-examenpopulatie

Mogelijke paden voorwaarts

- 1 Erkenning dat het niveau is gedaald en een maatschappelijk debat en permanente aandacht voor kwaliteit
- 2 Meer onderzoek naar andere onderwijsvormen (havo, vmbo), andere vakken en naar maatregelen
- 3 Realisatie van een positieve kwaliteitstrendbreuk door de lat hoger te leggen en het onderwijssysteem te



Inhoudsopgave

Het onderzoek in één oogopslag	3
01 Introductie	6
02 Samenvatting	10
03 Context, doel & methode	14
04 Bevindingen	26
4.1 PISA is een adequate en zo absoluut mogelijke ontwikkelingsniveaumeting voor leesvaardigheid, rekenen en natuurwetenschappen	29
4.2 Er mag een verband verwacht worden tussen PISA en het CE Wiskunde en Natuurkunde	30
4.3 Het niveau van het vwo CE is gedaald voor Wiskunde en Natuurkunde	32
4.4 Mogelijke paden voorwaarts	37
05 Ondersteunende schema's	40
06 Begrippenlijst	80
Achtergrond auteurs	82

01

Introductie

Dit rapport bespreekt het niveau van Wiskunde en Natuurkunde op het vwo. Met niveau wordt in dit rapport zowel de beheersing van deze vakken door leerlingen die het vwo verlaten bedoeld, als de moeilijkheidsgraad waarop zij geëxamineerd zijn. In deze inleiding behandelen wij, de auteurs van dit rapport, de context voor ons onderzoek, de aanleiding en de specifieke relevantie van Wis- en Natuurkunde. Daarna bespreken we de veelvoudige doelen van het onderwijs en de spanning die daarin zit, daarvan afgeleid onze kernvragen en een deel van de beperkingen van onderzoek in het onderwijs.

Nederland staat hoog aangeschreven als kenniseconomie, met bloeiende hubs als Brainport Eindhoven. Op de toch al krappe arbeidsmarkt is er echter een groot tekort aan technisch, zorg-, en ander talent ontstaan. Dit heeft een remmende werking op de groei van de economie en beperkt de mogelijkheden van de overheid en het bedrijfsleven om op dossiers zoals klimaatverandering de stappen te zetten die we als maatschappij zouden willen. Talentontwikkeling is juist nu van extra groot belang, en specifiek ook een goede basis in exacte vakken. De ontwikkeling van deze kennis en vaardigheden kan het beste zo vroeg mogelijk worden ingezet. Stimulering van interesse in exacte vakken bij leerlingen, en het ervoor zorgen dat ze goed kunnen functioneren door ze in contact te brengen met prikkelende leerstof, is daarom een belangrijke taak van het onderwijs.

De aanleiding voor dit onderzoek is de discrepantie tussen enerzijds de sterk afgenomen resultaten van Nederlandse leerlingen op internationale testen als PISA, de afgenomen resultaten op taal en rekenen in het basisonderwijs zoals gesignaleerd door de Inspectie en de Onderwijsraad, onderzoek van diverse universiteiten en anderzijds de toegenomen cijfers op het Centraal Examen (CE). Deze trend is beter zichtbaar als je hem op een lange termijn beschouwt. Dit onderzoek kijkt naar de afgelopen 30 jaar omdat dit de langste periode is waarop data hierover redelijk beschikbaar en vergelijkbaar is. Het doel van dit onderzoek is een stap te zetten die kan leiden tot het begrijpen en verklaren van deze paradox. Wij hebben ons onderzoek beperkt tot het niveau van Wiskunde en Natuurkunde op het vwo-eindexamen om deze verklaringen in meer detail te kunnen onderzoeken.

Het vwo-diploma geeft toegang tot het volgen van een studie aan een universiteit of hogeschool, en wordt door werkgevers ook als maatgevend gezien. Bovendien heeft het vwo-diploma ook indirect een sterke invloed op de havo en andere schoolvormen omdat de havo moet aansluiten op het vwo. Een grondige kennis van Wis- en Natuurkunde is, net als een goede beheersing van de Nederlandse taal, noodzakelijk voor allerlei beroepen. Ook is die kennis van groot belang om studenten de opleidingen te laten voltooien die hen klaarstomen om een bijdrage te kunnen leveren aan de kenniseconomie. Deze vakken worden daarom basisvakken genoemd, analoog aan hoe taal en rekenen basisvaardigheden worden genoemd.

Leerlingen krijgen meer kansen en gaan beter presteren als ze vroeg de gelegenheid krijgen om binnen deze belangrijke basisvakken in aanraking te komen met meer en uitdagendere stof en vraagstukken (het pygmalion-effect). Daarnaast is het de bedoeling dat de toetsen die we gebruiken maatgevend zijn voor wat we als maatschappij verwachten van het onderwijs. Het is ook daarom belangrijk het niveau van het vwo-eindexamen goed te begrijpen.

Internationale verdragen en de wet benadrukken dat ieder kind recht heeft op goed onderwijs. Kansengelijkheid is daarbij een belangrijk uitgangspunt. De Commissie Steur benadrukte in 2019 in haar brief aan de minister voor Basis- en Voortgezet Onderwijs en Media dat het voor een leerling van groot belang is dat het niet uitmaakt in welk jaar hij of zij het examen maakt. Dit geldt dus ook voor het niveau van 30 jaar geleden versus nu.

Het beoordelen van wat het onderwijs precies levert is open voor debat. Volgens de breed geaccepteerde indeling van Biesta heeft onderwijs meerdere doelen: kwalificatie, socialisatie en persoonsvorming. Dit onderzoek richt zich vooral op kwalificatie, maar we weten dat deze andere factoren nauw verbonden zijn en kwalificatie geen waarde heeft zonder dat ook deze andere resultaten geleverd worden door het onderwijs. Ook richt dit onderzoek zich met name op de kwalificatie die getoetst wordt volgens de wettelijk vastgelegde programma's voor het schoolexamen en centraal eindexamen. We zijn we ons ervan bewust dat hiermee voornamelijk het cognitieve onderdeel van kwalificatie onderzocht wordt. Het onderwijs moet aan veel verwachtingen voldoen (meer dan scholen aankunnen, zie o.a. *Een verstevigd fundament*) en heeft na de Coronacrisis en door het lerarentekort met veel uitdagingen te maken. Werkdruk, leerachterstanden en de toegenomen ongelijkheid tussen scholen leiden allemaal tot extra vragen aan het onderwijs.

Het Nederlandse funderend onderwijs staat voor een aanzienlijke inspanning om de kwaliteit van het onderwijs te verhogen en een trendbreuk in kwaliteit te realiseren zodat Nederland internationaal weer concurrerend kan worden en blijven met de wereldtop (zie bijvoorbeeld *Staat van het Onderwijs 2022* van de Inspectie en *Een verstevigd fundament voor iedereen*).

Om alle hierboven genoemde redenen denken we dat het belangrijk is om het niveau van ons onderwijs goed te begrijpen. Zoals hierboven beschreven zien we echter tegengestelde trends c.q. beeld in onze CE-resultaten versus andere toetsen (d.w.z. vooral versus andere nationale en internationale vergelijkingen). In dit onderzoek willen we daarom systematisch de mogelijke verschillen en grondoorzaken identificeren. De mogelijke verklaringen die we onderzoeken zijn:

1. In hoeverre is PISA een adequate ontwikkelingsniveaumeting voor Wiskunde en Natuurwetenschappen?
2. Is er een verband te verwachten tussen PISA en het Centraal Examen (CE) voor dezelfde vakken?
3. Is het niveau van het vwo CE veranderd in de afgelopen 30 jaar voor Wiskunde en Natuurkunde?

Op basis van (inter-)nationaal gevalideerde datasets toont ons onderzoek aan dat het niveau

van Wis- en Natuurkunde op het Centraal Examen (CE) van het vwo is gedaald. We hebben aanwijzingen dat deze daling langzaam over een langere periode heeft plaatsgevonden. Deze langzame daling over deze lange periode telt op tot een aanzienlijke niveaudaling (zeer aanzienlijk voor Wiskunde en aanzienlijk voor Natuurkunde). Vervolgens probeert dit onderzoek feiten aan te dragen die zouden kunnen verhelderen welke oorzaken hieraan ten grondslag kunnen liggen. Aan het eind van dit onderzoeksrapport doen wij een aantal suggesties voor mogelijk te ondernemen acties die een positieve kwaliteitstrendbreuk mogelijk kunnen maken, maar we willen benadrukken dat dit onderzoek niet gericht was op het doen van aanbevelingen. We willen vooral helpen om de juiste vragen te stellen. Er zijn wat ons betreft dus ook goede alternatieve oplossingen voor de in dit rapport genoemde suggesties mogelijk.

Om een sluitend antwoord te vinden op de vragen die dit onderzoek oproept is het belangrijk dat er een maatschappelijke beweging ontstaat die onderkent welke feiten en conclusies reflecteren wat we als maatschappij van ons onderwijs verwachten. Verder onderzoek is nodig om te toetsen of dezelfde thema's spelen bij andere vakken, bij andere onderwijsvormen en hoe dit samenhangt met de algehele wensen rondom het curriculum. Pas dan kunnen we ook een positieve kwaliteitstrendbreuk realiseren door enerzijds de lat hoger neer te leggen en anderzijds het onderwijssysteem te versterken. Het is belangrijk dat zowel curriculum, toetsen, scholen, besturen, lerarenopleidingen, het ministerie en alle andere actoren consistent deze eisen goed kennen en implementeren. Daarbij is het belangrijk om ambities te stellen en de middelen te kiezen die passen bij de uitdagende context waarin scholen opereren anno 2023.

Onderzoek in het onderwijs is altijd inherent onzeker en iedere onderzoeksmethode heeft haar beperkingen. Wij hebben met dit onderzoek getracht om aan de hand van bovengenoemde datasets een aanzet te geven tot een nieuwe onderzoeksmethode om de discrepanties te kunnen verklaren. In het 'methode' hoofdstuk hebben we expliciet vermeld welke beperkingen onze methode heeft. Deze aanzet is expliciet bedoeld als een uitnodiging tot verder onderzoek en discussie.

02

Samenvatting

Nederland heeft een nog onopgeloste paradox: verschillende ontwikkelingsniveaumetingen (CE-uitslagen, PISA, universiteiten) voor het vwo op het gebied van Wiskunde en Natuurwetenschappen als Natuurkunde laten een discrepantie zien:

Echter andere nationale en internationale bronnen duiden op een niveaustijging, bijvoorbeeld:

-  De afgelopen 30 jaar zijn de vwo cijfers voor de Centrale Examens (CE's) Wiskunde en Natuurwetenschappen als Natuurkunde gestegen (zie bijv. examenverslagen).
-  Het aandeel vwo- (en havo-)leerlingen in het funderend onderwijs is gestegen volgens het CBS
-  Het aandeel hogeropgeleiden in de beroepsbevolking is toegenomen volgens het CBS.

Echter andere (nationale en internationale) bronnen duiden op een mogelijke niveaudaling, bijvoorbeeld:

-  PISA laat voor het vwo Wiskunde en Natuurwetenschappen een significante scoredaling zien over de afgelopen 15-20 jaar.
-  Instroomniveaumetingen van de TU Delft hebben het instapniveau zien dalen en de TU Delft ontmoedigt studenten met een cijfer lager dan een 7 voor Wiskunde B om te komen studeren (dit geeft onvoldoende studiesuccesgarantie).
-  Relatief steeds minder plekken bij de studie Luchtvaart- en Ruimtevaarttechniek kunnen worden toegewezen aan Nederlandse studenten door het dalen van hun instroomniveau.
-  De Universiteit Maastricht heeft aangetoond dat Nederlandse eerstejaars studenten slechter scoren op een instaptoets (3TU2005) dan hun internationale studiegenoten.

Dit rapport vindt een mogelijke verklaring voor bovengenoemde paradox in een afgenomen niveau van de centrale examinering die een niveaudaling van de leerlingen wellicht heeft gemaskeerd.

Allereerst ...

... toont analyse aan dat PISA een adequate en zo absoluut mogelijke ontwikkelingsniveaumeting is voor 15-jarigen voor Wiskunde en Natuurwetenschappen: het voldoet aan objectieve criteria van het SLO voor een goede toets. Wel is PISA net als elk vergelijkend onderzoek aan een vorm van statistische onzekerheid onderhevig. Echter aangezien enerzijds PISA de best beschikbare kalibratiemethoden gebruikt en er geen sprake is van een systematische onzekerheid in het nadeel Nederland en anderzijds de daling van de scores voor Wiskunde en Natuurwetenschappen dermate groot en consistent zijn over een langere periode (ook voor de 10% sterkste leerlingen) is het onwaarschijnlijk dat deze daling in Nederlandse (vwo) PISA scores door statistische onzekerheid veroorzaakt wordt.

Ten tweede ...

... is er wel degelijk een verband te verwachten tussen PISA en het CE voor Wiskunde en Natuurwetenschappen als Natuurkunde, ondanks dat sommigen van mening zijn dat PISA als test geen goede maatstaf zou zijn voor Nederlandse leerlingen en stellen dat de vergelijking tussen PISA en het CE mank gaat, bijvoorbeeld omdat er aan een "low stakes test" geen consequenties verbonden zijn. Bewijs dat er wel een verband te verwachten is, is onder andere dat historische data laten zien dat toetsen op verschillende leeftijden goed met elkaar gecorreleerd zijn ("cohortonderzoek").

Ten slotte ...

... laat een inhoudelijke vergelijking van de CE's (op basis van meer dan 1.500 geanalyseerde CE-vragen) en de CE-resultaten over een periode van 30 jaar zien dat een niveaudaling heeft plaatsgevonden bij Wiskunde (A+B) en Natuurkunde, bij zowel de Centrale Examen als de leerlingen:

- De vwo CE-populatie voor deze vakken is de afgelopen 30 jaar minder goed voorbereid. Vwo-leerlingen van nu scoren minder goed op (bijna) identieke CE-vragen dan vwo-leerlingen van 30 jaar geleden.
- Ondanks een zwakkere vwo-populatie zijn de CE-resultaten afgelopen 30 jaar aanzienlijk verbeterd (0,4 tot 1,2 punt hoger cijfer).
- De normhandhaving (nu middels de N-term destijds middels de cesuur) biedt hiervoor

Simpele doorrekening om de gestegen vwo CE cijfers te verklaren

	Wiskunde B	Wiskunde A	Natuurkunde
A: Stijging van de centrale examencijfers afgelopen 30 jaar (op schaal 1-10)	+1,2 =	+0,6 =	+0,4 =
B: Bijdrage van het eenvoudiger maken van het Centraal Examen	+2,35 +	+2,05 +	+0,70 +
C: Bijdrage door veranderde cesuur/normhandhaving	+0,1 +	-0,2 +	+0,0 +
D: Bijdrage door minder sterke en/of voorbereide examenkandidaten	-1,4 +	-1,25 +	-0,5 +
B-D: Onverklaarde stijging	+0,15	+0,0	+0,20

geen verklaring omdat deze nu ongeveer even hoog is als 30 jaar geleden. Dat wil zeggen het aantal tienden dat bij het cijfer wordt opgeteld is vergelijkbaar.

- Wel blijkt uit een analyse van meer dan 1.500 CE-vragen dat betreffende CE's de afgelopen 30 jaar aanzienlijk makkelijker geworden zijn. Er is bijvoorbeeld 40% tot 50% netto minder CE-stof en vooral de CE-stof met moeilijkere vragen is geschrapt (behalve bij Natuurkunde). Ook zijn de vragen eenvoudiger geworden: deze bevatten minder denkstappen en er zijn meer makkelijke en aanzienlijk minder moeilijke vragen. Daarnaast blijkt dat er iets coulanter wordt nagekeken. Ook de School Examens (SEs) voor Wiskunde (A+B) en Natuurkunde bevatten minder stof dan 30 jaar geleden.

Het stijgende CE-cijfer heeft deze veranderingen wellicht gemaskeerd. Onderzoek in het onderwijs is altijd inherent onzeker. We hebben in dit onderzoek ook veel correlaties onderzocht en correlatie betekent niet altijd een causaal verband. Daarnaast zijn de data vaak multi-interpretabel. Betere CE-cijfers zouden ook kunnen wijzen op bijvoorbeeld betere voorbereiding, maar we hebben geen data kunnen vinden die wijst op betere voorbereiding. Om de invloed hiervan te beperken, hebben we daarom met dit onderzoek getracht om een diversiteit aan databronnen te gebruiken, en aan de hand van (inter)nationaal gevalideerde datasets een aanzet te geven tot een nieuwe onderzoeksmethode om bovenstaande discrepanties te kunnen verklaren. Deze aanzet is expliciet bedoeld als een uitnodiging tot verder onderzoek en discussie.

Mogelijke paden voorwaarts

We noemen hieronder een drietal suggesties die onderwijsbeleidsmakers zouden kunnen overwegen. Deze zouden onderdeel kunnen uitmaken van een uitgebreider plan dat kan helpen de trend te keren:

1. Een breed maatschappelijk debat over het niveau en de verwachtingen van het vwo wis- en natuurkundeonderwijs inclusief het her- en erkennen dat het niveau van vwo Wiskunde en Natuurkunde is gedaald. Leraren en schoolleiders staan onder enorme druk, hebben een aanzienlijke administratieve last en missen geregeld de capaciteit en/of vaardigheden om deze uitdaging aan te kunnen. Om hen te ondersteunen is het noodzakelijk dat alle betrokken partijen zich hiervoor inzetten, van het ministerie van OC&W tot de Inspectie van het onderwijs, raden, schoolbesturen, schoolleiders en ondersteuners. Daarbij is het van belang dat deze partijen zich committeren om zich voor een lange termijn in te spannen om dit doel te behalen.
2. In het verlengde hiervan een onderzoek doen naar de investeringen die benodigd zijn om dit niveau te verhogen. Daarnaast zou dit onderzoek uitgebreid moeten worden naar andere vakken (in het bijzonder Nederlands), het gehele vwo-eindexamen (CE + SE) en andere onderwijsvormen met een CE (havo en vmbo),
3. Een realisatie van een positieve kwaliteitstrendbreuk door
 - De lat voor de vwo-eindtermen en voor het CE voor Wiskunde en Natuurkunde weer hoger te leggen, bijvoorbeeld meer en geavanceerdere stof die op hoger niveau wordt getoetst. Ook kan een oplossingsrichting overwogen worden om wiskunde D een verplicht (CE-)examenvak te maken voor het profiel Natuur & Techniek (i.p.v. een 2^e moderne vreemde taal) en een (CE-)keuze-examenvak voor het profiel Natuur & Gezondheid.
 - Het versterken van het onderwijssysteem om dit mogelijk te maken. Laat bijvoorbeeld de Inspectie op stelselniveau helpen bij de CE-normhandhaving (controle op het CvTE), borg formeel de continuering van dit onderzoek, rol een continue verbetercultuur uit op alle scholen, en volg verdere aanbevelingen uit het rapport *Een verstevigd fundament voor iedereen* op.

03

Context, doel & methode

Wiskunde en Natuurkunde zijn naast Nederlands de basisvakken voor allerlei beroepen, maar ook voor het vervolgonderwijs, zoals ook beschreven in de introductie. Op de arbeidsmarkt is er een tekort aan technici, leraren en in de zorg (naast allerlei andere tekorten), hetgeen een remmende werking kan hebben op de groei van de Nederlandse kenniseconomie maar ook bij belangrijke dossiers zoals klimaattransitie en modernisering van de zorg. Daarom is het belangrijk om het niveau van het Nederlandse onderwijs in deze basisvakken goed te evalueren.

We realiseren ons dat het onderwijs aan veel verwachtingen moet voldoen, maar onderzoek laat zien dat leerlingen ook de kans moeten krijgen om op deze belangrijke basisvakken vroeg in aanraking te komen met meer en geavanceerdere leerstof en vragen op uitdagender niveau. De lat haalbaar hoger leggen heeft een sterk positief effect op het ontwikkelingsniveau zoals bewezen is in diverse meta-onderzoeken. De lat op het vwo hoger neerleggen zou ook het niveau van de andere onderwijsvormen positief kunnen beïnvloeden (“trekken aan de bovenkant”), omdat het vwo het ijkpunt is voor doorstromende leerlingen vanuit bijvoorbeeld de havo.

In dit hoofdstuk beschrijven we in meer detail de aanleiding voor dit onderzoek, de maatschappelijke context en onze methode en haar beperkingen. Vervolgens bespreken we in het volgende hoofdstuk de conclusies, acties die overwogen kunnen worden en eerste ideeën voor volgende stappen en vervolgonderzoek.

Aanleiding voor dit onderzoek

De aanleiding voor dit onderzoek is de paradox tussen enerzijds de sterk afgenomen resultaten van Nederlandse leerlingen op internationale testen als PISA, de afgenomen resultaten op taal- en rekenen in het basisonderwijs zoals signaleerd door de Inspectie en de Onderwijsraad, en onderzoek van diverse universiteiten en anderzijds de gestegen cijfers op het CE en het toegenomen aandeel vwo leerlingen en het toegenomen aandeel hogeropgeleiden in de beroepsbevolking volgens het CBS.

Uit een vergelijkende analyse van de PIRLS/TIMSS-scores, PISA-scores en de PIAAC-scores op verschillende momenten in de tijd blijkt dat een afnemend resultaat bij jongere leerlingen en bij 15-jarigen typisch voorspellend zou moeten werken voor hun prestaties op 18-jarige leeftijd. Toch zien we in het CE geen daling van cijfers – we zien zelfs een forse stijging voor Wis- en Natuurkunde. Dat roept de vraag op waar dit verschil vandaan komt.

De Commissie Steur heeft hier (op verzoek van de Tweede Kamer) eerder onderzoek naar gedaan, maar kon deze verschillen tussen beide toetsen niet kwantitatief verklaren. De commissie concludeert “(...) dat de vaardigheden van leerlingen qua niveau, zoals getoetst door het CE, in de afgelopen tien jaar gelijk gebleven zijn of zijn toegenomen” en dat het normhandhavingproces van het Cito (N-termen) robuust is. En “Als overkoepelende verklaring voor de betere leerlingresultaten op het Centraal Examen (CE), draagt de commissie onder andere aan dat het belang van het CE is toegenomen”. Deze conclusie wordt vaak samengevat als “PISA is een low-stakes test”. De discrepantie tussen het CE en PISA wordt echter niet kwantitatief onderbouwd. De Commissie heeft niet vakinhoudelijk gekeken of en hoe het niveau van het CE over de jaren veranderd is.

Dit rapport is tot stand gekomen in samenwerking tussen 2 partijen: Loek Zonnenberg, docent wiskunde (zij-instromer) en voormalig Partner bij McKinsey & Company, en Paul Rutten, Partner bij McKinsey & Company, op basis van hun gedeelde passie voor onderwijs. Loek heeft initiële analyses op de vwo-eindexamens Wis- en Natuurkunde uitgevoerd en hier vanuit het onderwijsveld een positieve reactie op gekregen en Paul was hoofdauteur van het McKinsey rapport een verstevigd fundament (zie hieronder). Loek en Paul hebben hun rapport onder de vlag van McKinsey uitgebracht en dit rapport kan dus gezien worden als een McKinsey onderzoek.

McKinsey heeft eerder onderzoek gedaan naar het niveau van het funderend onderwijs op verzoek van het ministerie van OC&W, en in dat rapport *Een verstevigd fundament voor iedereen* heeft McKinsey het vermoeden uitgesproken dat de toename van het belang van het CE niet een afdoende verklaring lijkt voor de vastgestelde verschillen, en dat een afnemende score in PISA wel degelijk zou wijzen op een afnemend niveau van het Nederlands onderwijs.

PISA-scores van Nederlandse 15-jarigen namen sinds 2003 af voor zowel Wiskunde als Natuurwetenschappen, en in mindere mate voor leesvaardigheid. Deze daling speelt bij zowel sterke leerlingen als bij leerlingen die het moeilijker hebben (al blijkt wel dat er geen daling in leesvaardigheid heeft plaatsgevonden bij de 10% sterkste leerlingen). Daarnaast kennen de auteurs drie bronnen van data die een andere kant op wijzen dan verbeterde resultaten op het CE, namelijk ten eerste laten instroomniveaumetingen van de TU Delft in de periode na 1987 zien dat het niveau van Wiskunde bij instromende studenten aan het dalen was. Ten tweede, ziet de faculteit Luchtvaart- en Ruimtevaarttechniek dat Nederlandse studenten het bij hun toelatingstest om een van de 440 studieplekken te verkrijgen steeds vaker afleggen tegen studenten uit het buitenland, waardoor relatief steeds meer studenten uit het buitenland een van deze 440 studieplekken toegewezen krijgt. Ten slotte, laat een internationale vergelijking van wiskundige vaardigheden op de 3TU2005 toets door de Universiteit Maastricht zien dat Nederlandse studenten het minder goed doen dan hun internationale peers.

Het onderzoeken van de bovenstaande paradox paste niet in de reikwijdte van McKinsey's vorige onderzoek *Een Verstevigd Fundament*. Daarnaast waren er vragen van partijen in het veld of McKinsey deze paradox beter kon verklaren. Mede daarom hebben de auteurs van dit rapport nu voor eigen risico en rekening het initiatief genomen tot dit vervolgonderzoek, waarin systematisch is gekeken naar alle mogelijke oorzaken van deze verschillen over de afgelopen decennia. We hebben ten eerste een raamwerk opgesteld om zoveel mogelijk oorzaken te zoeken voor de discrepantie. Op elk van deze punten hebben we vervolgens zo goed mogelijk historische data verzameld die deze oorzaken onderbouwen of verwerpen.

Maatschappelijke context

Het onderwijs is van vitaal belang voor de maatschappij. Volgens de indeling van Biesta zijn voor de ontwikkeling van leerlingen kwalificatie, socialisatie en persoonsvorming alle drie van belang. De recente Coronacrisis heeft dit nog eens benadrukt. Zo heeft de Coronacrisis bewezen dat school meer is dan alleen kwalificatie: mentale weerbaarheid en de connectie met leeftijdsgenoten is minstens van even groot belang.

McKinsey heeft eerder in een *Een verstevigd fundament* geconstateerd dat de eisen die aan scholen gesteld worden met betrekking tot het onderwijs niet alleen veelzijdig zijn, maar ook stijgen na verloop van tijd. *Een verstevigd fundament* heeft verder laten zien dat kwalificaties ook goed gecorreleerd zijn met andere succesfactoren op school, zowel makkelijk meetbare (scores op toetsen) als minder makkelijk meetbare aspecten (bijvoorbeeld zelfstandigheid van leerlingen). Hieruit volgt dat een afname van kwalificaties extra zorgelijk is, ook al is het niet de gehele opdracht aan het onderwijs. Ook zijn kwalificaties van groot belang voor leerlingen in hun vervolg meer kansen te krijgen. Leerlingen moeten op een zo hoog mogelijk niveau de middelbare school verlaten, dus moeten ze op een zo hoog mogelijk niveau worden geëxamineerd (hoge verwachtingen leiden tot meer leeropbrengsten, zoals onder andere geïllustreerd in het meta-onderzoek van John Hattie). Het bovengenoemde rapport van 2020 liet ook zien dat de maatschappelijke verwachtingen aan het onderwijs hoog zijn, en nog altijd verder stijgen. Het is niet realistisch om te verwachten dat het onderwijs aan alle maatschappelijke verwachtingen kan voldoen, maar wel dat investeringen ten behoeve van continue innovatie en verbetering nodig zijn om aan deze toenemende verwachtingen tegemoet te komen.

Daarnaast wordt de werkdruk in het onderwijs als zeer hoog ervaren – en dit is door de Coronacrisis en het oplopende lerarentekort nog eens sterk toegenomen. Dit onderzoek is dan ook geen pleidooi om de werkdruk verder te verhogen, maar om oplossingen te zoeken die passend zijn binnen de context van het lerarentekort en de werkdruk, en daar ook de juiste middelen voor ter beschikking te stellen. Docenten kunnen bijvoorbeeld binnen hun normjaartaak meer tijd aan leerlingen besteden als hun administratieve last verminderd wordt (eliminieren, simplificeren, standaardiseren en automatiseren van taken). Daarnaast zou lesgeven op hoger niveau een extra motivatie kunnen bieden aan getalenteerde professionals om voor het docentenvak te kiezen. Dit onderzoek heeft echter niet in detail gekeken naar de rol van de leraar en hoe deze versterkt kan worden.

Het onderwijs in Nederland is in het algemeen van hoog niveau. Daarnaast is het Nederlandse onderwijs breed toegankelijk. Het Hoger Onderwijs is in de afgelopen decennia bijvoorbeeld voor meer Nederlanders weggelegd. De emanciperende functie van het onderwijs is een groot goed.

Ten slotte worden Wiskunde en Natuurkunde, naast Nederlands, als van zeer groot belang voor de maatschappij gezien. Ze zijn het fundament voor veel wetenschappen, niet alleen de traditionele exacte wetenschappen, maar ook voor bijvoorbeeld geneeskunde, klimaat, economische en sociale wetenschappen. Digitalisering en Big Data zijn belangrijke onderwerpen om op lange termijn concurrerend te zijn als kenniseconomie en Wiskunde is hiervoor een kritieke basis. We weten bovendien dat goede resultaten in Wiskunde een van de beste voorspellers is voor succes in het Wetenschappelijk Onderwijs. De toename van digitalisering, de stijgende vraag naar technici, het belang van klimaatdeskundigheid, om maar een paar voorbeelden te noemen, versterken de trend dat de Nederlandse maatschappij deze vaardigheden hard nodig heeft. Het is daarom goed om te onderzoeken hoe het funderend onderwijs in staat kan worden gesteld om beter gekwalificeerde leerlingen op te leiden, zonder dat dit afdoet aan de andere maatschappelijke doelen en uitdagingen die het onderwijs heeft.

Methode en haar beperkingen

In dit onderzoek is gezocht naar manieren om alle langjarige ontwikkelingen in de resultaten van het eindexamen Wiskunde en Natuurkunde te objectiveren en zo veel als mogelijk te kwantificeren. Om deze reden is begonnen met het opstellen van een raamwerk van mogelijke oorzaken (redenen) van de discrepantie tussen de ontwikkeling van het CE versus andere bronnen. Dit raamwerk neemt alle factoren expliciet mee die een verklaring kunnen geven voor de verandering van de resultaten van het CE (bijvoorbeeld gemiddeld cijfer en percentage voldoende). Voor elk van deze factoren is vervolgens gezocht naar bronnen uit de periode 1990 tot en met 2021, zowel van de resultaten zelf als de onderliggende drijvers van die resultaten. Deze periode omvat ook de laatste grote curriculumwijzigingen na de introductie van Wiskunde A en Wiskunde B als vak.

Resultaten in het onderwijs zijn van jaar tot jaar moeilijk te vergelijken. Er is variatie in de leerlingpopulatie, de examenvragen, de normering en statistische onzekerheid. Deze onzekerheid ("ruis") kunnen we deels vermijden door te kijken naar een zo lang mogelijke trend, en door meerdere jaren te middelen. Maar ook dan zijn er nog allerlei beperkingen.

Er zijn slechts beperkt makkelijk toegankelijke longitudinale datasets die deze vergelijking mogelijk maken. Curriculumwijzigingen maken dit extra uitdagend. Voor de jaren 1983 en 1984 bestonden Wiskunde A en B nog niet en werden de vakken Wiskunde 1 en 2 gegeven met deels onvergelykbare stof. Wiskunde 2 bestond toen voornamelijk uit Lineaire Algebra dat sindsdien niet meer in die uitgebreide vorm op het vwo gegeven is. Gegeven deze curriculumwijziging lijkt het logisch om een datum te kiezen enige jaren na 1983/84, zodat de invoeringseffecten van deze wijziging niet in onze analyse terug zijn te vinden. In de eerste

paar jaar na 1983/1984 werden Wiskunde A en B als pilot afgenomen naast Wiskunde 1 en 2.

In veel gevallen zijn historische bronnen vergeleken met recente data, bijvoorbeeld boeken met oude examenvragen, Cito TIAs en de vakbladen Euclides en NVOX om oude officiële landelijke p-waarden zoals gepubliceerd door Cito te vinden en te vergelijken met huidige online bronnen (p-waarden zijn het deel van de punten die landelijk voor een vraag werden behaald).

In al deze bronnen is gezocht naar een zo lang mogelijke periode om de effecten van verschillen in prestaties van individuele jaren of van individuele CE's uit te middelen. Daarnaast dient te worden opgemerkt dat er aanwijzingen zijn dat het niveau al vanaf of voor het jaar 2000 aan het dalen was zoals te zien is in de instroomniveaumetingen van het TU Delft-onderzoek. Een bijkomende uitdaging, zoals hierboven al aangegeven, was dat Wiskunde en Natuurkunde in de periode 2002-2010 bestond uit deelvakken 1 en 2, wat een andere indeling was dan de vakken Wiskunde A en B, en dus een andere kandidatenpopulatie had. Daarom hebben we gekozen om de examenperiode 1990-1995 (6 jaren aan Centrale Eindexamens) te vergelijken met de examenperiode 2015-2021 (6 jaren aan Centrale Eindexamens) met alleen voor de analyse met betrekking tot de stofverandering een focus op de periode 2018-2021 voor Wiskunde en 2016-2021 voor Natuurkunde, gezien de laatste curriculumwijziging die vlak hiervoor plaatsvond.

In 2020 was er geen Centraal Examen en ook het eindexamen 2021 is natuurlijk beïnvloed door Corona. Daarom hebben we gecontroleerd of de resultaten significant zouden veranderen als de examens van het Coronajaar 2021 wel of niet worden meegenomen. Dit is niet het geval.

Voor de analyses met p-waarden hebben we alleen de resultaten voor het eerste tijdvak meegenomen. Een aantal van de geraadpleegde docentexperts benadrukken dat volgens hen de sterkste leerlingenpopulatie deelnam aan het eerste tijdvak in 2021.

Daarnaast is gekozen om voor wat het vak Wiskunde betreft de analyses te beperken tot Wiskunde A en Wiskunde B omdat deze ook in het nieuwe vakkenpakket verreweg de twee grootste vakken zijn. Wiskunde D is geen examenvak en zowel Wiskunde D als Wiskunde C wordt slechts door een relatief klein deel (elk zo'n 5%) van de leerlingen gevolgd. Wiskunde A en B zijn dus het meest representatief voor het wiskundeonderwijs op het vwo omdat ze het grootste deel van de leerlingen omvatten.

Er is een samenvattend raamwerk ontworpen, waarin de drie belangrijkste vragen worden samengevat over het verschil tussen PISA en het CE.

Ten eerste ...

... is de vraag gesteld in hoeverre PISA een adequate ontwikkelingsniveaumeting is voor Wiskunde en Natuurwetenschappen (en taal/leesvaardigheid). Hiervoor is de methode van PISA in detail bekeken en afgezet tegen de criteria voor een goede toets zoals gegeven door het SLO. Hierna zijn de data uit de methodologische verantwoordingen van PISA bestudeerd en hebben interviews met PISA plaatsgevonden over de validiteit en methodiek van PISA – in het bijzonder ook met dhr. Andreas Schleicher (PISA Director for Education and Skills Special Advisor on Education Policy to the Secretary-General). Ook zijn individuele PISA-vragen en analyses van de OECD bestudeerd en zijn historische data geanalyseerd die meer inzicht geven over de validiteit van hun methode, zoals onder meer de relatie tussen PISA-scores en voorbereiding en motivatie voor PISA. Ook is de recente publicatie van het CPB *Een blik op de Nederlandse positie in internationale onderwijsrankings* meegenomen.

Ten tweede ...

... is de vraag gesteld of het redelijk is om een verband te verwachten tussen PISA en het Centraal Examen (CE) voor dezelfde vakken. Hiervoor is gekeken naar het verband tussen PISA en het CE en andere toetsen op andere leeftijden, en is de inhoud van PISA vergeleken met de inhoud van het reguliere curriculum. Daarbij is gebruik gemaakt van de documentatie en data van de OECD, waarbij niet alleen naar PISA is gekeken maar ook

naar de relaties tussen PISA, TIMSS (internationale toets voor 10-jarigen voor rekenen en natuurwetenschappen), PIRLS (internationale toets voor 10-jarigen voor taal) en PIAAC (internationale toets voor volwassenen voor taal en rekenen), stuk voor stuk internationaal geaccepteerde testen voor het niveau van leerlingen en volwassenen op verschillende leeftijden.

Ten slotte ...

... is gekeken naar het niveau van de Centrale Eindexamens en School Examens voor Wiskunde en Natuurkunde op het vwo in de afgelopen 30 jaar. Deze periode komt goed overeen met dezelfde periode als waarin de PISA-resultaten fors zijn afgenomen. De CE-resultaten (gemiddelde cijfer en percentage voldoende) zijn geanalyseerd, waarbij is afgevraagd waar deze CE-resultaten van afhangen. Daarin is voor deze vakken onderscheiden:

- Wat is de ontwikkeling van het niveau van het CE, oftewel hoe eenvoudig (of moeilijk) is het CE zelf?
- Hoe soepel (of streng) is de normhandhaving (ijking/kalibratie) en gebeurt dit op meer absolute of relatieve wijze?
- Wat is de ontwikkeling van het niveau van de vwo CE-kandidaten, oftewel hoe “sterk” (of “zwak”) zijn de vwo CE-kandidaten?

In de analyses naar het niveau van de School Examens is alleen de stofverandering over afgelopen 30 jaar onderzocht.

Bij ieder van deze vragen is gekeken naar de onderliggende drijvers. Voor elk van de onderliggende drijvers is bewijs verzameld, bij voorkeur een kwantitatieve analyse. Bij de moeilijkheidsgraad van het CE is gelet op de breedte (“hoeveelheid”) en diepte (“moeilijkheidsgraad”) van de getoetste CE-(leer)stof, het vereiste tempo en uithoudingsvermogen waarin inhoudelijk wis- en natuurkundig werk moet worden verricht, de moeilijkheidsgraad van de gestelde vragen en toegestane hulpmiddelen, hoe streng het correctievoorschrift is en hoe streng er wordt nagekeken.

Hiervoor kijken we onder andere naar meer dan 1.500 CE-opgaves (toetsitems) die zijn opgenomen in de eindexamens in de afgelopen 6 jaar en 25-30 jaar geleden, maar ook naar de normering van de vragen, publicaties in vakbladen over cijfers en slagingspercentages, “makkelijke” en “moeilijke” vragen op basis van het percentage leerlingen die (bijna) alle of maar heel weinig punten voor een vraag behalen, of het aantal denkstappen (complexiteit) dat nodig is om een vraag te kunnen beantwoorden. Daarnaast is geverifieerd of recente examens meer woorden bevatten dan vroeger, omdat er vaak beweerd wordt dat Centrale Examens “taliger” geworden zijn, waarbij meer leeswerk sommige leerlingen zou benadelen. Hiervoor zijn de woorden geteld van recente en oude examens en is uitgerekend hoeveel meer leestijd er gemiddeld is bijgekomen gegeven deze toename. Voor deze analyses gebruiken we examenbundels, examenwebsites en gegevens uit vakbladen.

Bij de normhandhaving hebben we gekeken naar de hoogte van de definitieve N-term of cesuurverschuiving destijds (de aanpassing die het Cito toepast om de cijfers te bepalen van leerlingen), de bepaling van de N-term, en specifiek of deze voornamelijk een absolute of relatieve normhandhaving (ijking/kalibratie) is en hoe frequent de technische N-term wordt aangepast, om welke reden en welke kant op.

Bij het niveau van de vwo CE-kandidaten voor Wiskunde en Natuurkunde, zijn veranderingen in de leerlingenpopulatie geanalyseerd. Dit omvat onder andere het niveau van de gehele CE-populatie (alle onderwijsvormen), de relatieve grootte van de vwo-populatie voor Wiskunde of Natuurkunde ten opzichte van de gehele CE-populatie, maar ook hoe goed de leerlingen zijn voorbereid. De grootte van deze populatie is te vinden in CBS-data en examenverslagen, net als de grootte van de beroepsbevolking met een hbo/wetenschappelijke achtergrond. Over voorbereiding van leerlingen bestaat helaas weinig

data. Er is ook gekeken naar de uitgaven aan bijlesbedrijven en verdere indicaties van gebruik van bijles. Een andere invalshoek die genomen is het vergelijken hoe leerlingen van nu en van 30 jaar geleden landelijk scoren op bijna identieke vragen. Om deze vraagparen van bijna identieke vragen te vinden, heeft een systematische matching van alle CE-vragen van tijdvak 1 tussen beide perioden plaatsgevonden. Elke individuele vraag is eerst toegekend aan het desbetreffende examendomein. Vervolgens zijn de gevonden vragen per domein beoordeeld op vergelijkbaarheid. Dit heeft voor Wiskunde B drie matches, voor Wiskunde A twee matches en voor Natuurkunde negen matches opgeleverd.

Voor onze analyses zijn ons geen bestaande longitudinale datasets bekend, dus zijn deze door onszelf samengesteld uit verschillende bronnen. Daarbij is onder andere gebruik gemaakt van vwo-examenprogramma's, Centrale Examens en Correctievoorschriften op OCW- en vakwebsites, examenbundels, vwo samengevat, en examenverslagen, Cito TIAs, interviews met toets- en vakexperts, vakbladen Euclides en NVOX (voor het ophalen van oude p-waarden). Daarbij zijn alle examenvragen individueel gerubriceerd op onderwerp, gecombineerd met p-waardes en is vergelijkbaarheid van alle vragen ingeschat. Daarnaast is gebruik gemaakt van publicaties van het Cito met betrekking tot normhandhaving.

Bij onze analyses hebben we een test gedaan op betrouwbaarheid, o.a. Pearson's R2 en op basis van controle of betrouwbaarheidsintervallen van twee verschillende datapunten niet aanzienlijk overlappen (hierbij is de betrouwbaarheid vastgesteld en dat was over het algemeen hoog, tussen de 85% en 99%). Omwille van de leesbaarheid hebben we deze data niet overal in de tekst toegevoegd. Voor het vak Natuurkunde hebben we, omdat het gevonden scoreverschil op (bijna) identieke vragen bij Natuurkunde kleiner is dan bij Wiskunde A en B, extra getest of deze scoreverschil toe te schrijven is aan een niveauverandering van leerlingen en niet bijvoorbeeld veroorzaakt is doordat de vragen van nu (2015-2021) uit de vraagparen moeilijker zijn dan die van toen (1990-1995), of omgekeerd.

Hiertoe is een aantal van de vraagparen getest bij een aantal leerlingen in een statistische tekentoets opzet waarbij dezelfde leerling (bereidwillige leerlingen van het Maerlant-Lyceum te Den Haag) zowel de vraag van toen als nu maakt. Het voordeel van een dergelijke opzet is dat niet een groot aantal leerlingen nodig is om statistisch enigszins betrouwbare resultaten te verkrijgen.

Hierbij wordt aangenomen dat de kans dat een leerling per toeval een vraag van nu beter maakt 50% en een vraag van toen ook 50%. De vragen die even goed worden gemaakt worden niet meegenomen, ondanks dat deze zouden bevestigen dat de vraag minimaal even moeilijk of makkelijk was.

De week voor de toets hebben de leerlingen een korte herhaalles gekregen over de stof om ervoor te zorgen dat alle leerlingen een gelijk minimaal benodigd kennisniveau hebben. Een serieuze deelname werd beloond met een gratis lunch.

De leerlingen hebben aangegeven een serieuze poging tot beantwoording van de vragen te hebben gedaan en dat de vragen niet bij hen bekend waren.

Bovenstaande methode stelt ons naar onze mening in staat om een goede analyse te maken van het niveau voor Wiskunde en Natuurkunde op het (Centrale) Eindexamen vwo. Op elk van de ondersteunende schema's bij dit rapport staat vermeld welke bronnen zijn gebruikt en waar relevant is een extra methodologische toelichting gegeven bij de uitgevoerde analyse.

We hebben diverse toets- en inhoudelijk vakexperts uitgenodigd om onze methodiek en analyses te valideren en ons te wijzen op blinde vlekken.

Elk van de betrokken experts heeft gevalideerd met betrekking tot het eigen gebied van expertise. Dit zijn onder andere:

em. prof. dr. Jan Karel Lenstra

oud-decaan faculteit Wiskunde & Informatica en
hoogleraar optimalisering TU Eindhoven,
oud-directeur Centrum Wiskunde & Informatica

em. prof. dr. Gerrit Timmer

hoogleraar (Bedrijfs)Econometrie Vrije Universiteit
Amsterdam, mede-oprichter Ortec

prof. dr. Vinod Subramaniam

voorzitter College van Bestuur Universiteit
Twente, hoogleraar biofysica

em. prof. dr. ir. Wim van Saarloos

emeritus hoogleraar theoretische natuurkunde
Lorentz Instituut Universiteit Leiden, oud-
voorzitter Koninklijke Nederlandse Akademie van
Wetenschappen

prof. dr. ir. Ton van der Steen

hoofd biomedische engineering Erasmus MC

dr. ir. Christa Hooijer

voorzitter Nederlandse Natuurkundige
Vereniging (NNV), Technisch Natuurkundig
ingenieur Universiteit Twente, gepromoveerd UvA

drs. Noortje de Graaf

directeur Nederlandse Natuurkundige Vereniging
(NNV), doctoraal natuurkunde UvA

prof. dr. Peter van der Heijden

hoogleraar statistiek ten behoeve van de sociale
wetenschappen, psychometrist, Universiteit
Utrecht

**em. prof. dr. Paul A. Kirschner, dr.
h.c.**

emeritus hoogleraar Onderwijspsychologie

prof. dr. Jasper Knoester

decaan van de Faculteit der Wiskunde en
Natuurwetenschappen Universiteit Leiden

prof. dr. Gerard T. Barkema

bèta vice-decaan Universiteit Utrecht, voorzitter
landelijke bèta vice-decanen overleg

prof. dr. Sijbrand de Jong

decaan van de Faculteit Natuurwetenschappen,
Wiskunde en Informatica van de Radboud
Universiteit en hoogleraar experimentele
natuurkunde, Radboud Universiteit Nijmegen

prof. dr. Peter van der Straten

voormalig opleidingsdirecteur BSc Natuur- en
Sterrenkunde en hoogleraar experimentele
fotonfysica Universiteit Utrecht

prof. dr. Nicolo de Groot

vice-decaan bèta-wetenschappen en hoogleraar
deeltjesfysica, Radboud Universiteit Nijmegen

prof. dr. Rob Timmermans

vice-decaan Faculty of Science & Engineering
en hoogleraar theoretische deeltjesfysica
Rijksuniversiteit Groningen

em. prof. dr. Frans Keune

emeritus hoogleraar algebra Radboud Universiteit

prof. dr. Jos Thijssen

opleidingsdirecteur MSc Applied Physics bij
de TU Delft, hoogleraar Technische Natuurkunde

prof. dr. ir. Mark Voskuil

hoogleraar Wapen- en Luchtvaartsystemen
bij de Faculteit Militaire Wetenschappen van
de Nederlands Defensie Academie

ir. Joris Melkert

directeur onderwijs en UHD faculteit
Luchtvaart- en Ruimtevaarttechniek TU Delft

ir. Margot Kok

Directeur Onderwijsbeleid faculteit
betawetenschappen Universiteit Utrecht,
technisch natuurkundig ingenieur TU Twente

dr. Hanno van Keulen

Opleidingsdirecteur MSc Science Education &
Communication, TU Delft

drs. Jan van Riswick

vakdidacticus en doctoraal Natuurkunde
Radboud Universiteit, eerstegraads docent
natuurkunde (1986-heden)

dr. Ed van den Berg

gepensioneerd vakdidacticus Natuurkunde
Universiteit Twente

ir. Enno van der Laan

vakdidacticus Natuurkunde Rijksuniversiteit
Groningen

dhr. Bernard La Rivière

vakdidacticus Natuurkunde en lector Hogeschool
van Amsterdam

dr. Wim Caspers

bestuur Nederlandse Vakvereniging van
Wiskundeleraren NVvW, eerstegraads
vwo-docent (Wiskunde) in residence TU Delft EWI

Robert Zibret

NVON sectiehoofd Natuurkunde

drs. Heleen van der Ree

econometrist EUR, beleidsmedewerker
Nederlandse Vereniging van Wiskundeleraren

ir. Heleen Neggers

eerstegraadsdocent Wiskunde, technisch
wiskundig ingenieur TU Eindhoven

Jim van Bekhoven, MEd

eerstegraadsdocent Wiskunde

Ingrid Hoovers

eerstegraadsdocent Wiskunde

dr. Jelmer Wagenaar

docent Universiteit Leiden en eerstegraadsdocent
Natuurkunde bij het Stedelijk Gymnasium Leiden

dr. Arjan Keurentjes

eerstegraadsdocent Natuurkunde,
gepromoveerd natuurkundige Universiteit Leiden

ir. Jan Willem Bosch

eerstegraadsdocent Natuurkunde, technisch
natuurkundig ingenieur aan de TU Delft

Christiaan Spapé, MEd

eerstegraadsdocent Natuurkunde

dr. Sezgin Cihangir

directeur Nederlands Mathematisch Instituut

dr. ir. Menel Rahrah

gepromoveerd technisch wiskundige aan
de TU Delft

dr. ir. Ivo Vink

gepromoveerd technisch natuurkundige
aan de TU Delft

dr. ir. Meindert van der Meulen

technisch natuurkundig ingenieur aan de TU Delft
en gepromoveerd theoretisch natuurkundige aan
de UvA

Jenny Lu, BSc

Honours Master Student Technische Wiskunde
en bijlesleraar met een 10 voor Wiskunde B en
Natuurkunde op het vwo CE

Uit gesprekken met experts zijn de volgende opmerkingen en kanttekeningen bij onze analyses naar voren gekomen:

- Dit onderzoek omvat alleen kwalificatie, en biedt dus geen overzicht van de ontwikkeling van persoonsvorming en socialisatie. Voor deze laatste twee is de meetbaarheid en beschikbaarheid van data veel minder. Dit betekent dat in dit onderzoek niet is getoetst of leerlingen op andere manieren vaardiger zijn geworden dan alleen gemeten als kwalificatie, en er zijn ook geen data over “onmeetbare” (maar wel belangrijke) outputs van het onderwijs. Sommige experts gaven aan dat vooral kennis/cognitieve vaardigheden ondergesneeuwd zijn geraakt en dit geldt volgens hen ook voor het universitaire vervolgonderwijs: “aversie tegen excelleren: er wordt te veel naar het gemiddelde gekeken en niet naar onze toppers”, “echte wetenschappers leiden we niet meer op”, “universiteiten mogen geen eisen stellen”, “perverse financiële prikkel op universiteiten om studenten erdoorheen te trekken”, “de stofverandering bij Natuur- en Wiskunde verklaart precies waarom Nederlandse studenten beduidend minder goed geworden zijn in L&R-kernvakken als Dynamica”. Dit onderzoek doet op andere vaardigheden echter geen uitspraken.
- Wiskunde is een taal die je jong al moet leren, dus niet alleen vlak voor het eindexamen of op de universiteit.
- Dit onderzoek kan de inputs van het onderwijs niet scheiden van het ingaande niveau van leerlingen vanwege gebrek aan data. Er is dus niet apart gekeken naar de kwaliteit en mogelijk dalende status/waardering van leraren, de kwaliteit van hun lessen, de organisatie van scholen of “remedial teaching”, ofwel alle vormen van bijles dan wel binnen of buiten school.
- Dit onderzoek heeft niet in detail kunnen kijken naar examenvoorbereiding, in het bijzonder naar schaduwonderwijs. Bijles door ouders of andere leerlingen heeft altijd een rol gespeeld, maar lijkt de laatste jaren door groei bij commerciële aanbieders van bijles prominenter aanwezig. Er zijn door ons geen data gevonden die het effect van bijles op onderwijsresultaten in Nederland in de afgelopen jaren kwantificeert. Wel is het duidelijk dat schaduwonderwijs slechts een fractie van de personeelskosten van het totale onderwijs omvat (0,2 versus 16 miljard euro), wat erop lijkt te wijzen dat het schaduw-onderwijs niet meer dan 1-2% van het resultaat van het onderwijs kan verklaren als men aanneemt dat schaduwonderwijs niet meer dan een factor 2 meer effectief is dan regulier onderwijs in termen van leeropbrengsten. Ook merkten een aantal experts op dat bijles vaak ook een averechts effect kan hebben, omdat een leerling door de bijles minder doet tijdens de les (“ik kan het nog met bijles inhalen”). Bijles is soms sommen maken met een student of docent die souffleert waarbij de leerling dan denkt in staat te zijn de sommen te kunnen maken terwijl dat zelfstandig nog niet het geval is. Deze leerlingen merken dan bij de toets pas dat dat niet zo is (“tijdens bijles lukte het altijd”).
- Dit onderzoek omvat alleen Wiskunde en Natuurkunde. Het is waarschijnlijk dat eisen

voor andere vakken en vaardigheden, samen met de breedte van het vakkenpakket, scholen op dit moment dwingt om ook keuzes te maken binnen de basisvakken Wiskunde en Natuurkunde – dit valt buiten de scope van dit onderzoek. Dit onderzoek heeft niet gekeken naar Wiskunde C en D. Wiskunde D bestond twintig jaar geleden nog niet. Het merendeel van de geschrapte, meer geavanceerde onderwerpen is nog wel in Wiskunde D terug te vinden. Wiskunde D is echter geen goede maatstaf voor het niveau van Wiskunde in het middelbaar onderwijs omdat slechts 2.000-3.000 vwo-leerlingen jaarlijks dit vak afronden (van de ongeveer 42,000 vwo CE-kandidaten), wat maar ongeveer 5-7% van het totaal aantal vwo-leerlingen in het examenjaar is dat veel kleiner is dan het aandeel Wiskunde B. Wiskunde D is geen onderdeel van het CE.

- Bij een aantal experts bestaat de indruk dat er minder contacturen voor Wiskunde A en B en Natuurkunde in de bovenbouw vwo zijn. Anderen merkten op dat docenten minder opgaven klassikaal hoeven uit te werken omdat deze uitwerkingenboeken nu standaard voor leerlingen verkrijgbaar zijn.
- Over examenstof die niet meer in het Centraal Examen (CE), maar nog wel in het School Examen (SE) zit is opgemerkt dat dit niet wil zeggen dat die goed en op het juiste niveau wordt getoetst. De verschillen in cijfers tussen scholen zijn groot, leraren focussen op het CE waardoor de SE stof die niet in het CE zit te weinig aandacht krijgt en de SE stof is vaak te summier omschreven. Een deel van de experts ziet ook geen reden om stof die in het CE getoetst kan worden, bijvoorbeeld statistiek en kansrekening bij Wiskunde, alleen in het SE te toetsen. Dit wordt wel als logisch gezien

voor onderdelen als een practicum waar het format van het Centraal Examen zich niet goed voor leent.

- De grootte van onze steekproef voor de analyse met vergelijkbare vragen is relatief klein. Dit komt omdat er helaas niet meer vragen beschikbaar zijn die twee keer gebruikt zijn op een examen. De resultaten van onze analyses met vergelijkbare vragen zijn wel statistisch significant voor de getrokken conclusies in dit onderzoek, maar blijven gebaseerd op een relatief kleine set.
- Met betrekking tot niveau en IQ van leerlingen op het vwo, hebben we niet onderzocht wat het effect is van grotere deelname van populaties die in het verleden niet op het vwo zaten.
- Alle data in het onderwijs zijn voor meerdere uitleg vatbaar (bijv. hogere scores op het CE kunnen zowel het gevolg zijn van een makkelijkere toets als van betere voorbereiding). Er is systematisch geprobeerd hierover feiten te verzamelen, maar deze kunnen niet altijd de alternatieve uitleg uitsluiten. In het bijzonder hebben we geen data kunnen vinden die duidelijk laten zien of leerlingen meer gemotiveerd zijn voor het CE door de jaren heen door het toegenomen belang van het CE voor lotingsstudies etc., wat soms wel als verklaring genoemd wordt.
- Een van de trends die geobserveerd zijn zoals curriculumontwikkeling en aantal woorden, zijn het gevolg van bewuste beleidskeuzes. Er is in dit onderzoek niet onderzocht of die keuzes het gewenste effect behaald hebben, er is enkel gekeken wat de data zeggen over de resultaten van leerlingen op de examens in de loop der tijd.

04

Bevindingen

Nederland heeft een nog onopgeloste paradox: verschillende ontwikkelingsniveaumetingen (CE-uitslagen, PISA, universiteiten) voor het vwo op het gebied van Wiskunde en Natuurkunde laten een discrepantie zien.

In dit hoofdstuk beschrijven we onze bevindingen. Ten eerste beschrijven we data die de aanleiding zijn voor dit onderzoek, namelijk de trends die al bekend waren voorafgaand aan dit onderzoek. We bekijken daarin de resultaten van PISA, het Centraal Examen (CE) en diverse andere onderzoeken. Daarna analyseren we eerst de geschiktheid van PISA als ontwikkelingsniveaumeting, het te verwachten verband tussen PISA en het CE en ten slotte het niveau van het vwo CE voor Wis- en Natuurkunde.

De nationale Centrale Examens (CE's) en het vervolgonderwijs laten over afgelopen 30 jaar een positief verloop zien

De afgelopen 30 jaar zijn de vwo CE-cijfers voor Wiskunde en Natuurkunde met 0,4-1,2 punt gestegen (Schema 1). Het gemiddeld behaalde vwo eindexamencijfer (SE en CE) is een goede voorspeller voor het succes in vervolgstudies (Schema 2). Het eindexamencijfer voor Wiskunde B als vak is een goede voorspeller voor het succes aan een technische universiteit (Schema 10).

Ook is het aandeel vwo- (en havo-)leerlingen in het funderend onderwijs sinds 1990 met ongeveer 50% toegenomen (Schema 3). Dit is in lijn met het streven van de Lissabon-doelstellingen die Nederland onderschrijft, en de vertaling daarvan in de HOOP, waarin gearticuleerd is dat meer dan 50% van de leerlingen hogeropgeleiden zou moeten worden. Mede hierdoor is het aandeel hogeropgeleiden in de beroepsbevolking sinds 2005 met 43% toegenomen (Schema 4). Dit betekent betere kansen op de arbeidsmarkt voor een groot deel van de bevolking. Er is niet onderzocht of het ontwikkelingsniveau van de hogeropgeleiden minimaal gelijk is gebleven (valt buiten de scope van dit onderzoek).

Andere nationale en internationale onderzoeken laten echter een minder positief beeld zien

De Inspectie heeft in verschillende publicaties van de Staat van Onderwijs afgelopen jaren aangegeven dat de onderwijsresultaten dalen (Schema 5). In een recent advies *Taal en Rekenen in het vizier* uit ook de Onderwijsraad zorgen over de beheersing van taal en rekenen (Schema 6). Daarnaast dalen al 15-20 jaar de PISA scores van Nederlandse kinderen onafhankelijk van het schooltype waarop ze zitten (Schema 7). Deze daling heeft ervoor gezorgd dat we

20-25% minder leerlingen hebben die als “sterk” worden gekwalificeerd, en inmiddels twee keer zoveel leerlingen die als “zwak” worden gekwalificeerd, wat overigens ook betekent dat de ongelijkheid in Nederland is toegenomen. De daling bij de sterkste 10% (vaak vwo-leerlingen) is het sterkst voor Wiskunde en Natuurwetenschappen, bij Leesvaardigheid is er geen langjarige trend (Schema 8). De daling van ongeveer 20 PISA punten voor Wiskunde en Natuurwetenschappen komt in PISA termen ruwweg overeen met een half tot een heel schooljaar aan voortgang.

Daarnaast hebben de auteurs drie bronnen van data gevonden die een andere kant op lijken te wijzen dan verbeterde resultaten op het CE. Allereerst lijken ankervraagstukken uit de instroomniveaumetingen van de TU Delft erop te duiden dat het instapniveau van eerstejaarsstudenten na 1987 en met name in de periode voor 2005 sterk gedaald is (Schema 9). In de periode 2005-2010 lijkt het niveau licht te zijn toegenomen, maar niet tot het niveau van voor 1987. Daarnaast is rond 2014 geconstateerd dat een vwo-diploma met een cijfer voor Wiskunde B lager dan een 7 onvoldoende garantie geeft op studiesucces bij een technische studie aan de TU Delft (Schema 10). Een voldoende op de middelbare school lijkt daardoor inmiddels niet dezelfde waarde te hebben: de TU Delft ontmoedigt via hun website inmiddels dan ook studenten met een cijfer lager dan een 7 voor Wiskunde B om te komen studeren (Schema 11). Een bijkomende oorzaak was dat scholen met lagere CE-resultaten deze compenseerden (tussen 2005-2008) met te makkelijke School Examens (Schema 12), wat ook kansengelijkheid in de hand werkt. Het grote verschil in kansen voor kinderen tussen scholen is in *Een verstevigd fundament* al verder toegelicht. Helaas maakt het uit naar welke school een kind gaat.

Een tweede bron is de toelatingstoetsen van de Faculteit Luchtvaart- en Ruimtevaarttechniek (L&R) van de TU Delft die het niveau van aspirant-studenten toetst om vanwege een numerus fixus plekken te kunnen toekennen voor hun bachelorprogramma (Schema 13). Op basis van deze toelatingstesten kunnen, gecorrigeerd voor instroom, de laatste 6-7 jaar steeds minder plekken worden toegekend aan Nederlandse studenten door het dalen van hun instroomniveau (Schema 14). Daar waar Nederlandse studenten ten opzichte van hun instroom eerst een oververtegenwoordiging hadden in het aandeel toegekende plekken, is dat inmiddels een ondervertegenwoordiging. Er wordt vooral terrein verloren aan andere EU-landen. Als je kijkt naar de vertegenwoordiging van Nederlandse studenten in de top-100 van de toelatingstoets dan is deze trend nog sterker zichtbaar (Schema 15). Deze resultaten duiden op een afgenomen niveau in recente jaren (2017-2022) van wis- en natuurkundevaardigheden op het vwo ten opzichte van studenten buiten Nederland, met name versus de EU.

Ten derde hebben we een internationale vergelijking door de Universiteit Maastricht op wiskundige vaardigheden op de 3TU2005 toets bestudeerd. Deze laat zien dat Nederlandse studenten 20 procent lager scoren dan Duitstalige of Internationale Baccalaureaat studenten (Schema 16).

Bovenstaande geconstateerde discrepantie roept een drietal vragen op (Schema 17):

1. In hoeverre is PISA een adequate ontwikkelingsniveaumeting voor Wiskunde en Natuurwetenschappen (en Taal)?
2. Is er een verband te verwachten tussen PISA en het Centraal Examen (CE) voor dezelfde vakken?
3. Is het niveau van het vwo CE veranderd in de afgelopen 30 jaar voor Wiskunde en Natuurkunde?

Hieronder worden de feiten behandeld die zijn verzameld bij ieder van deze mogelijke verklaringen.

4.1

PISA is een adequate en zo absoluut mogelijke ontwikkelingsniveaumeting voor leesvaardigheid, rekenen en natuurwetenschappen

The “Program for International Student Assessment” (PISA) van de OECD is een internationale studie die voor het eerst gehouden werd in 1997, en in Nederland voor het eerst volwaardig in 2003. Nederland participeerde al in het jaar 2000, maar de “response rate” van scholen was toen te laag voor een statistisch betrouwbare PISA score.

PISA toetst elke 3 jaar leesvaardigheid, wiskundige geletterdheid en natuurwetenschappelijke geletterdheid. Het laatste PISA onderzoek vond plaats in 2018.

PISA is erop gericht om vaardigheden te testen die in de maatschappij van belang zijn. Het is dus expliciet niet een toets gericht op het hebben van kennis, maar wel op het toepassen ervan. In de praktijk is er echter een grote overeenkomst tussen de vragen voor Wiskunde en Natuurwetenschappen en het standaardcurriculum, zoals bijvoorbeeld een vraag waarbij de stelling van Pythagoras nodig is om de schuine zijde van een driehoek te kunnen berekenen of het kunnen doen van berekeningen met een wiskundige formule. PISA geeft aan dat experts uit Nederland aangeven dat voor leesvaardigheid 85% van de vragen van PISA direct voorkomen in het curriculum Nederlands voor de middelbare school.

Daarnaast is PISA doelbewust een “low-stakes” test omdat PISA wil meten of je opgedane vaardigheden op een willekeurig moment in je leven kunt toepassen. Dit vindt de OECD relevanter dan het specifiek kunnen reproduceren en toepassen van geleerde kennis op één moment. Deelnemers van bijna alle deelnemende landen gaven aan dat ze zich beter op PISA zouden hebben voorbereid als het een “high-stakes” test was geweest.

PISA voldoet aan de objectieve criteria van het SLO voor goede toetsing. Het is daarmee dus een adequate ontwikkelingsniveaumeting voor het funderend onderwijs, en dus voor 15-jarige vwo-leerlingen voor taal/leesvaardigheid, Wiskunde en Natuurwetenschappen (Schema 18). PISA is ook internationaal geaccepteerd als de best beschikbare standaard, ondanks het feit dat op elke gestandaardiseerde test natuurlijk ook kritiek is, dus ook op PISA. Een kritiekpunt is (naast de overlap met het curriculum) bijvoorbeeld dat PISA meer gericht is op kwalificatie, en minder of kinderen ook een volwaardig burger zijn. PISA brengt daar zelf tegenin dat leesvaardigheid, Wiskunde en begrip van Natuurwetenschappen van groot belang zijn voor de maatschappij. Dit komt overeen met de bevinding in *Een verstevigd fundament* dat resultaten op socialisatie en persoonsvorming gecorreleerd zijn met goede resultaten op kwalificatie, ofwel goed onderwijs op kwalificatie is voorspellend voor goed onderwijs op allerlei andere aspecten. Daarnaast doet PISA al jaren aanvullend onderzoek naar deelthema's om haar data een beter beeld te laten geven, zoals bijvoorbeeld op motivatie of “growth mindset”.

Een ander kritiekpunt is dat PISA niet specifiek de stof van een eindexamen toetst. Uit analyses voor Taal van PISA blijkt dat de overlap tussen PISA en de stof die in het (standaard) nationaal curriculum zit geen groot effect heeft op de PISA score (Schema 24). Daarnaast heeft de OECD voor het vak Wiskunde aan landen gevraagd om hun eigen prioriteiten aan te geven, en heeft daarmee de PISA-scores opnieuw berekend. Uit deze herberekening bleek dat landen over vrijwel de gehele linie dezelfde internationale scores en ranking zouden krijgen als de nationale prioriteiten aangehouden worden (Schema 25).

Daarbij is het ook van belang om op te merken dat PISA een toets is met (vrijwel) absolute

normering. Dat wil zeggen dat (vrijwel) identieke opgaves gebruikt worden door de jaren heen om te zorgen dat het niveau van PISA constant gehouden wordt, ook als de opgaven worden vernieuwd. Dit maakt longitudinale trends in PISA goed vergelijkbaar, zeker in één enkel land. Dit soort absolute normering is alleen mogelijk bij een “low-stakes” test (anders zou testvoorbereiding de testresultaten te veel verstoren), en dat is waarom PISA er expliciet naar streeft om een infrequente, “low-stakes” test te blijven in plaats van een jaarlijkse barometer.

Tenslotte is een kritiekpunt dat PISA net als elk vergelijkend onderzoek aan een vorm van statistische onzekerheid onderhevig is, zie bijvoorbeeld de recente CPB-publicatie over de positie van Nederland in internationale onderwijsrankings. Echter aangezien enerzijds PISA de best beschikbare kalibratiemethoden gebruikt en er geen sprake is van een systematische onzekerheid in het nadeel Nederland en anderzijds de daling van de scores voor Wiskunde en Natuurwetenschappen dermate groot en consistent zijn over een langere periode (ook voor de 10% sterkste leerlingen) is het onwaarschijnlijk dat deze daling in Nederlandse (vwo) PISA scores door statistische onzekerheid veroorzaakt wordt (Schema 19).

4.2

Er mag een verband verwacht worden tussen PISA en het CE Wiskunde en Natuurkunde

PISA voldoet aan de objectieve criteria van het SLO voor goede toetsing. PISA-resultaten zijn dus een goede weergave van het ontwikkelingsniveau van 15-jarigen. Gegeven de overeenkomst in vaardigheden die in PISA getest worden en in het CE, verwachten we ook een correlatie tussen de uitkomsten bij 15-jarigen en bij 16-19-jarigen bij hun volgende examen. Die (uitkomsten) tonen longitudinale analyses tussen TIMMS / PIRLS (metingen bij 10-jarigen), PISA (15-jarigen) en PIAAC (metingen op latere leeftijd waaronder 16-19-jarigen) aan: resultaten op 10-jarige leeftijd zijn goed voorspellend voor resultaten bij 15-jarigen, en deze zijn goed voorspellend voor resultaten op 16- tot 19-jarige leeftijd (Schema 26).

De commissie Steur ondersteund door het Cito stelt dat de vergelijking tussen PISA en het CE mank gaat (Schema 20, 21 en 22). Ze geven hiervoor met name zes argumenten: vergelijkbaarheid, het feit dat het CE later plaatsvindt (16-19-jarigen) dan PISA (15-jarigen), het feit dat PISA geen consequenties kent (“low-stakes” test), de lagere motivatie van leerlingen vergeleken met andere landen, de toegenomen motivatie van Nederlandse leerlingen voor het CE en ten slotte het feit dat PISA in 2015 de methodiek digitaal heeft gemaakt. Op veel van deze items zien we echter geen data die de grote verschillen tussen 20 jaar geleden en nu verklaren (Schema 23 tot en met 29). Zo was PISA 20 jaar geleden ook een “low-stakes” test, is er geen bewijs dat leerlingen op den duur minder gemotiveerd zijn geworden, en ook niet dat lagere motivatie leidt tot significant lagere PISA-scores. Ook is het CE altijd belangrijk geweest voor leerlingen omdat ze dit nodig hadden voor zo’n 50% voor hun diploma en daarmee voor hun vervolgopleiding. Wel zijn er tegenwoordig meer studies met loting. Daarnaast geeft de digitalisering van PISA in 2015 ook geen verklaring voor de daling van de resultaten in de 15 jaar daarvoor.

Het argument van de “low-stakes test” wordt als een van de belangrijkste verklaringen gezien, en daarom zijn we hier dieper op ingegaan. Schema 27 laat zien wat het verband is tussen de motivatie en voorbereiding van leerlingen in verschillende landen en hun resultaat op PISA. Hieruit blijkt een lichte afname van de PISA-score als motivatie en voorbereiding voor PISA omhoog gaan – precies andersom dan verwacht zou worden als een “high-stakes” test tot hogere resultaten zou leiden (maar de statistische betrouwbaarheid is laag).

Het is echter overtuigend niet zo dat voorbereiding en motivatie voorspellende factoren zijn voor resultaten op PISA. Motivatie als voorspeller voor resultaten wordt ook betwijfeld door onderzoek van de Nederlandse Inspectie. PISA laat daarentegen zien dat een “growth mindset” wel een positief effect heeft op de PISA resultaten. Deze “growth mindset” zoals onder andere gepropageerd door Carol Dweck is het geloof in de incrementele theorie van intelligentie, ofwel dat iemand gelooft dat zijn/haar vaardigheden en intelligentie zich in de loop van de tijd kunnen ontwikkelen en toenemen door inzet (Schema 29).

Naast voorbereiding en motivatie is het voor een toets minstens zo belangrijk of leerlingen ook aandacht besteden aan de antwoorden en of ze de toets afmaken. PISA houdt hierover data bij, en die laten zien dat Nederlandse kinderen internationaal ruim bovengemiddeld aandacht besteden aan de opdrachten (niet doorklikken) en de toets ook veelal afmaken. Dit geeft aan dat Nederlandse kinderen ondanks de “low-stakes” context wel degelijk de PISA toets serieus maken. Dit in tegenstelling tot wat verwacht zou worden als Nederlandse leerlingen minder gemotiveerd zouden zijn dan landen waarin aan PISA meer gewicht wordt gegeven op scholen. Wel is de “response time” (de totale tijd besteed aan de PISA vragen) in Nederland lager dan bij andere landen, maar dit kan bijvoorbeeld ook worden verklaard door het altijd nog relatief hoge PISA-vaardigheidsniveau in Nederland dat onze leerlingen in staat stelt sneller het antwoord op vragen te vinden. Uit de data blijkt dat er geen verband is tussen een lagere “response time” en een lagere PISA score, dus dit lijkt ook geen verklaring te zijn (Schema 28).

Tenslotte is PISA expliciet ontworpen als een “low-stakes” test en maakt het gebruik van diverse strategieën om het een “low-stakes” test te laten blijven (zoals onder andere het afwisselen van scholen en het voorkomen dat PISA landelijk afgenomen wordt). PISA is zodoende opgezet en gevalideerd met het “low-stakes” karakter als uitgangspunt.

Uit de data-analyse blijkt dus dat PISA een goede zo absoluut mogelijke methode is, die wel degelijk het verschil laat zien tussen 2000 en 2018 op een vergelijkbare wijze en die ook grotendeels voorspellend zou moeten zijn voor leerlingen op 18-jarige leeftijd. De PISA scores en trends zijn daarom een goede indicatie voor het ontwikkelingsniveau in het funderend onderwijs, zeker wanneer de daling zo groot is, en ook specifiek als we Nederland met Nederland vergelijken (en niet in vergelijkende rankings).

De Commissie Steur wees er ook op dat het onderwijs in de bovenbouw afwijkt van onderwijs in de onderbouw en dat dit een mogelijke verklaring is waarom leerlingen anders kunnen scoren op het CE dan in PISA. Dit vormt echter nog niet een verklaring voor de toenemende discrepantie. In Nederland is geconstateerd dat de CE-cijfers voor Wiskunde en Natuurkunde in de afgelopen 20 jaar zijn gestegen. Als het CE inderdaad geen niveaudaling en zelfs een stijging van de cijfers laat zien, zou dit moeten betekenen dat leerlingen in toenemende mate een steeds grotere versnelde inhaalslag van één tot vier jaar in stof na PISA kunnen maken naar het CE toe. Ofwel, in de laatste 20 jaar zouden leerlingen in de laatste jaren van hun schoolcarrière steeds sneller en effectiever zijn gaan studeren. Dit wordt niet door data ondersteund, er zijn in de afgelopen 15-20 jaar geen structurele interventies in het Nederlandse onderwijs geweest die een zodanig grote inhaalslag zouden kunnen verklaren.

4.3

Het niveau van het vwo CE is gedaald voor Wiskunde en Natuurkunde

De derde mogelijke verklaring voor de schijnbare afname van het niveau van Wiskunde en Natuurkunde is dat het niveau van het vwo CE voor deze vakken de afgelopen 30 jaar veranderd is. Deze verklaring wordt in dit hoofdstuk nader onderzocht. Daarnaast hebben we geanalyseerd of het SE de afgelopen 30 jaar veranderd is. Ten slotte stippen we ook kort aan hoe momenteel Wis- en Natuurkunde in de vwo-profielen zijn gepositioneerd.

Om het vwo CE Wis- en Natuurkunde over de afgelopen 30 jaar te onderzoeken, vergelijken we de periode 1990-1995 (6 examenjaren) met 2015-2021 (6 examenjaren) met alleen voor de analyse die de stofverandering in kaart brengt een focus op de periode 2018-2021 voor Wiskunde en 2016-2021 voor Natuurkunde gezien de laatste curriculumwijziging die vlak hiervoor plaatsvond (Schema 30).

Allereerst kijken we (A) hoe de CE-resultaten veranderd zijn. Daarna analyseren we welke factoren hierbij een rol hebben gespeeld: (B) Wat is de ontwikkeling in de moeilijkheid van het CE?, (C) Hoe streng is de normhandhaving (ijking/kalibratie) en gebeurt dit op meer absolute of relatieve wijze? en (D) Wat is de ontwikkeling van het niveau van de vwo CE-kandidaten? Op Schema 31 hebben we de logische probleemboom verder uitgewerkt. We behandelen hieronder de resultaten van de analyses bij elk van de vier onderdelen van de probleemboom.

A. Significant betere CE-resultaten nu dan 30 jaar geleden

Als we de periode 1990-1995 met 2015-2021 vergelijken, dan zien we dat de Centrale Eindexamencijfers zijn toegenomen bij Wiskunde A, B en Natuurkunde. Daaraan gecorreleerd is het aantal voldoende aanzienlijk toegenomen. Het gemiddelde cijfer voor Wiskunde B is gestegen van 5,8 (58% voldoende) naar 7,0 (84% voldoende), voor Wiskunde A van 6,3 (70% voldoende) naar 6,9 (79% voldoende) en voor Natuurkunde van 6,2 (69% voldoende) naar 6,6 (74% voldoende) (Schema 32).

B. Het CE Wiskunde en Natuurkunde is makkelijker geworden

Uit onze analyses blijkt dat de moeilijkheidsgraad van het CE Wiskunde A en B zeer aanzienlijk is afgenomen en van het CE Natuurkunde aanzienlijk is afgenomen.

De breedte ('hoeveelheid') en diepte ('moeilijkheidsgraad') zijn bekeken. De breedte van de CE-leerstof is (zeer) fors afgenomen voor Wiskunde A en B en Natuurkunde. Tussen de 40% en 50% van de stof is netto van het CE verdwenen (Schema 33). Er is in detail geëvalueerd welke stof uit het curriculum is verdwenen en welke is toegevoegd aan de hand van examenprogramma's en evaluatie van vragen op het CE (Schema's 34 t/m 45). Van de verdwenen stof zijn voorbeelden toegevoegd. Verder blijkt uit de analyse van meer dan 1.500 CE toetsitems (vragen/opgaven) dat er nauwelijks stof is toegevoegd. Wat wel is toegevoegd, staat niet in verhouding tot de grote hoeveelheid thema's die verwijderd zijn (Schema 46). De diepte blijkt bij Wiskunde A en B bovendien afgenomen omdat de elementen die zijn weggehaald significant moeilijker zijn dan wat is overgebleven (Schema 47). Dit geldt niet voor Natuurkunde, waar een klein deel moeilijker stof is geïntroduceerd (met name quantum-tunneling).

Om zeker te weten of het valide is om de puntenwaardes van opgaves te kunnen vergelijken over de jaren heen, is bij Wiskunde B en Natuurkunde gekeken of scorepunten voor vergelijkbare vaak voorkomende wiskundige en natuurkundige handelingen ongeveer

hetzelfde zijn. Dit is het geval (Schema 48). Voor Wiskunde A was deze vergelijking niet mogelijk omdat de stof te ingrijpend veranderd is voor een statistisch betrouwbaar resultaat.

Wel is geconstateerd dat de “hoeveelheid” inhoudelijk werk van de vragen op het CE afgenomen is. Dit is te vergelijken door de maximaal te behalen scorepunten in 3 uur examentijd naast elkaar te zetten (Schema 49). Hieruit blijkt dat de te behalen punten met 10-17% afgenomen zijn. Er wordt ook veel gesproken of het CE Wiskunde en Natuurkunde niet taliger geworden is. Dit blijkt het geval te zijn. Om dit te toetsen, is het aantal woorden in een CE geteld en is gekeken naar het verschil tussen 1990-1995 en 2015-2021 (Schema 49). Hieruit blijkt dat het aantal woorden met niet meer dan 300-700 woorden toegenomen is. Vragen zijn context-rijker geworden. Dit representeert ongeveer 5 à 10 minuten van de examentijd en heeft daarmee geen grote invloed op het vereiste tempo. Kijkend naar het CE is het tempo dat van leerlingen geëist wordt nu lager (Schema 49).

Als de stof verminderd is, is het nog steeds mogelijk dat leerlingen meer uitgedaagd worden door moeilijkere vragen (althoewel er minder stofcombinatiemogelijkheden voor een vraag zijn). Het tegenovergestelde is echter het geval. Kijkend naar de Centrale Examen worden er nu minder complexe of samengestelde vragen gesteld (minder denkstappen per vraag) dan 30 jaar geleden (Schema 50). Er is geconstateerd dat er een verband is tussen het aantal denkstappen per vraag en de p-waarde: hoe meer denkstappen per vraag des te lager de p-waarde, dus hoe meer denkstappen per vraag hoe moeilijker de vraag (Schema 51). Ook geven experts aan dat het lastiger is moeilijkere vragen te stellen over eenvoudigere stof dan over meer geavanceerde stof.

We hebben gekeken welke vragen relatief makkelijk zijn voor de deelnemers aan het CE en welke vragen relatief moeilijk zijn. Dit is gedaan door te kijken naar de zogenaamde p-waarde van deze toetsitems (welk deel van de punten is door de leerlingen behaald). Makkelijke vragen definiëren we als een p-waarde van 85 procent of hoger en moeilijke vragen als een p-waarde van 35 procent of minder. Als we de examens van 30 jaar geleden vergelijken met die van nu is het duidelijk dat er nu meer makkelijke en aanzienlijk minder moeilijke vragen op het CE zijn dan 30 jaar geleden (Schema 52).

Een mogelijke verklaring van dit effect zou kunnen zijn dat de CE-kandidaten over de gehele linie van de CE-stof nu veel sterker of beter voorbereid zijn dan 30 jaar geleden waardoor zij de moeilijke vragen beter maken en dus voor een kleiner deel van het CE de p-waarde 35% of lager is, en een groter deel van de overige vragen beter beantwoorden waardoor een groter deel van het examen een p-waarde van 85% of hoger heeft. Echter, vooruitlopend op onderdeel D (verder) in dit rapport zijn er data die duiden op een niveaudaling van de vwo CE-kandidaten, en dus in ieder geval niet op een significante niveaustijging.

Daarnaast zijn er nu ook meer hulpmiddelen toegestaan in het CE, waaronder de grafische rekenmachine, staat een deel van de wiskundeformules op papier en worden er hints gegeven, zoals aangegeven in vaktijdschriften en in het correctievoorschrift (bijvoorbeeld hulpgrafieken en het opknippen van vragen in deelstukken waarbij in de eerste vraag het antwoord gegeven wordt, “toon aan dat...”) (Schema 53). Ook is opgevallen dat bij Natuurkunde minder vragen gesteld worden waarbij een beroep op wiskundige vaardigheden wordt gedaan en leerlingen bij grotere vraagstukken meer “aan de hand worden meegenomen” door de benodigde deelstappen om het probleem op te lossen als recept erbij te geven in plaats van de leerling deze zelf te laten bedenken.

Ten slotte is ook het correctievoorschrift iets versoepeld, door bijvoorbeeld minder precieze notaties te vereisen bij Wiskunde of meer en eerdere deelscores bij Natuurkunde (Schema 54). Er wordt verwacht dat dit leidt tot hogere scores voor soortgelijke Centrale Examen. De inschatting is dat dit gemiddeld één à twee extra scorepunten per kandidaat oplevert, dus een relatief klein effect op het totale CE cijfer (ruwweg 0,1 tot 0,2 punt op het hele CE).

C. De normhandhaving van het vwo CE is niet gebaseerd op een nauwkeurige absolute ijking

Ten behoeve van de normhandhaving gebruikt het Cito nu N-termen om de cijfers voor examens aan te passen aan de moeilijkheid van dat specifieke examen en 30 jaar geleden een cesuur(verschuiving). Voor de bepaling van de N-termen verzamelt het Cito de data en stelt het College voor Toetsen en Examens (CvTE) de N-termen vast. In dit onderzoek gebruiken we de N-termen (minus 1) en cesuur(verschuivingen) om het aantal tienden dat het cijfer wordt aangepast in beide perioden te berekenen (Schema 55). Het is daarom een logische vraag of het aantal tienden dat het cijfer wordt aangepast, is veranderd mettertijd. De huidige N-termen zijn vergelijkbaar met de cesuur van 30 jaar geleden. Dus is het aantal tienden dat het CE-cijfer wordt aangepast nagenoeg gelijk en biedt het geen verklaring voor de CE-cijferstijging tussen beide onderzochte perioden. Tussentijds zijn er wel schommelingen geweest (Schema 56). Voor een uitgebreidere analyse hiervan verwijzen we naar het Cito-rapport *Vaardigheidsontwikkelingen volgens PISA en examens*, dat onderliggend is aan de analyse van de Commissie Steur.

Het Cito heeft een brede set methodes die het inzet om het niveau van de Centrale Examens te bewaken (Schema 57). Uit deze methodes blijkt dat het vwo Centraal Examen niet de meest absoluut vergelijkbare toets is. Het Cito geeft aan dat de methodes toegepast op het vwo 3 maal (pre- & posttest) tot 9 maal (standaardbepaling) onnauwkeuriger zijn dan het meest nauwkeurige ijkingsinstrument, namelijk het opnemen van identieke of sterk vergelijkbare opgaves (ankervraagstukken) zoals in PISA wordt gedaan. In een “low-stakes” test als PISA is dit uiteraard makkelijker te realiseren dan in een “high-stakes” test als het CE, omdat dit mogelijk ertoe leidt dat leerlingen zich voorbereiden op deze specifieke opgave. Het Cito gebruikt deze methode dus ook niet voor het vwo CE. Daarnaast heeft het Cito ook te maken met budgettaire beperkingen omdat de meest absolute vergelijkingen ook zeer bewerkelijk zijn.

Het Cito concludeert op basis van gelijkblijvende N-termen en analyse van de vaardigheidsontwikkeling dat de normering voor de eindexamens gelijk is gebleven. Dit veronderstelt echter dat de populatie voor de examens over langere termijn hetzelfde is gebleven, wat niet het geval is. Deze populatie is sterk gegroeid, dat wil zeggen, het aandeel vwo-kandidaten is sterk gegroeid. De slagingspercentages zijn echter niet afgenomen – zoals we eerder zagen is het aandeel voldoende toegenomen. Daarnaast normeerde het Cito eind vorige eeuw voornamelijk op basis van de deelnemende populatie, terwijl toen het aandeel vwo-leerlingen sterk steeg. Dit bemoeilijkt uiteraard het absoluut handhaven van het niveau.

D. Het niveau van de vwo-examenpopulatie is lager dan 30 jaar geleden

Het gemiddelde IQ van de hele populatie is afgelopen 40 jaar nagenoeg constant gebleven (of zeer licht afgenomen) volgens onderzoek.

Aan het CE Wiskunde A en B en Natuurkunde nemen nu relatief meer kandidaten deel dan 30 jaar geleden, dus is deze CE-populatie relatief zwakker geworden omdat een grotere groep uit de hele populatie (iedereen in Nederland) nu deelneemt aan het vwo-CE voor Wiskunde A, Wiskunde B en Natuurkunde (Schema 58):

- Het aandeel vwo-CE kandidaten is ongeveer 30 procent gestegen.
- Binnen het vwo CE is het aandeel Wiskunde B 4pp gestegen en Natuurkunde 9pp. Wel is het aandeel Wiskunde A 13pp gedaald, maar de toename van het aantal vwo-CE kandidaten gecombineerd met deze afname leidt netto toch tot een groter aandeel Wiskunde A van de totale populatie CE-kandidaten.

Ook is het niet aannemelijk dat de huidige CE populatie beter is voorbereid om drie redenen:

1. De voorbereidingstijd: de huidige generatie heeft meer afleidingen tot zijn/haar beschikking
2. De kwaliteit en hoeveelheid van het oefenmateriaal: 30 jaar geleden was ook 15-20 jaar aan oefenexamens beschikbaar
3. De hoeveelheid en de kwaliteit van de begeleiding: het aandeel academische vwo bèta-docenten is afgenomen en bijlessen waren 30 jaar geleden ook beschikbaar (weliswaar minder formeel en geïnstitutionaliseerd dan nu). Het totale effect (niet voor een individuele leerling) van de bijlesindustrie wordt beperkt geschat, omdat de omzet van de bijlesindustrie t.o.v. regulier vo-lerarensalarissen minder dan 2% is. Harde data over het gebruik van bijles bij informele aanbieders ontbreken over de afgelopen 30 jaar.

In de afgelopen decennia zijn de zak-slaagregelingen eerst afgezwakt, en in de afgelopen 10 jaar weer aangescherpt. In 2013 is de kernvakkenregeling ingevoerd, waarbij voor vwo voor Nederlands, Engels en Wiskunde maximaal één vijf mocht worden gehaald. Dit is vergelijkbaar met hoe streng het examen was in de periode rond 1990-1995, toen er maximaal één vier mocht zijn of twee vijven, maar waarbij het gemiddelde 6,0 of hoger moest zijn, en waarbij het CE voor 50% meetelde.

Ondanks het feit dat het Cito geen identieke vragen stelt (zoals hierboven besproken), zijn er wel vragen te vinden in de CE's van 30 jaar geleden en nu die zeer sterk op elkaar lijken. We hebben deze opgaves opgezocht en zien dat leerlingen nu minder goede scores (p-waarden) hebben op vrijwel identieke vragen: 16pp minder goed bij Wiskunde B, 14pp bij Wiskunde A en 6pp bij Natuurkunde (Schema 59 tot en met 65).

We kunnen met deze analyse geen onderscheid maken tussen de extra instroom van vwo-leerlingen of de voorbereiding van de CE-kandidaten (hun daadwerkelijke niveau). Gegeven de grootte van de daling is het waarschijnlijk dat beide een rol spelen. Een aantal experts merkt hierbij op dat vaak de betere havo-leerlingen vwo gaan doen en dat die niet onder doen voor een gemiddelde vwo-leerling.

Omdat de scoredaling op (bijna) identieke vragen bij Natuurkunde kleiner is dan bij Wiskunde A en B, is een klein deelonderzoek uitgevoerd om te testen of deze scoredaling toe te schrijven is aan een niveaudaling van leerlingen of aan de oorzaak dat de vragen van nu (2015-2021) uit de vraagparen moeilijker zijn dan die van 30 jaar geleden (1990-1995).

Hiertoe is een aantal van de vraagparen van (bijna) identieke vragen getest bij leerlingen in een statistische tekentoets opzet waarbij dezelfde leerling zowel de vraag van 30 jaar geleden als nu maakt. Het voordeel van een dergelijke opzet is dat niet een groot aantal leerlingen nodig is om enigszins statistisch betrouwbare resultaten te verkrijgen. Hierbij wordt aangenomen dat de kans dat een leerling per toeval een vraag van nu beter maakt 50% is en een vraag van 30 jaar geleden ook 50%. In een tekentoets worden de vragen die even goed worden gemaakt niet meegenomen, ondanks dat ze eigenlijk ook bevestigen dat de vragen van nu niet moeilijker zijn dan van 30 jaar geleden.

Uit deze test zijn de volgende resultaten naar voren gekomen (Schema 66):

- De deelnemende leerlingen zijn qua niveau goed gekalibreerd met het huidig landelijk niveau (van alle vwo leerlingen): op exact dezelfde CE-vragen van nu scoren zij 61% goed versus 56% landelijk door de leerlingen van nu.
- Echter lijkt hun niveau beduidend onder dat van de leerlingen van 30 jaar geleden te liggen: op exact dezelfde CE vragen van toen scoren zij 37% goed versus 64% landelijk door de leerlingen van 30 jaar geleden.

- De vragen van 30 jaar geleden uit deze drie vraagparen zijn even moeilijk of zelfs moeilijker dan de vragen van nu:
 - 1. Dezelfde leerlingen van nu scoren 61% op de vragen van nu, terwijl zij 37% goed scoren op de vragen van 30 jaar geleden.
 - 2. Voor drie geteste vraagparen maken meer leerlingen de vraag van nu beter dan de vraag van 30 jaar geleden:
 - » Vraagpaar 1: de vraag van nu wordt beter gemaakt door dezelfde leerlingen. De kans dat dit per toeval was gebeurd is ongeveer 6%
 - » Vraagpaar 7: de vraag van nu wordt beter gemaakt door dezelfde leerlingen. De kans dat dit per toeval was gebeurd is ongeveer 23%
 - » Vraagpaar 8: de vraag van nu wordt beter gemaakt door dezelfde leerlingen. De kans dat dit per toeval was gebeurd is ongeveer 19%.

Hieruit blijkt dat de landelijke scoredaling van 8pp landelijk scoredaling (64% versus 54%) op deze drie vraagparen goed overeenkomt met de 6pp scoredaling voor alle 9 vraagparen (67% versus 61%). Hierdoor is er geen reden om de gevonden 6pp scoredaling bij Natuurkunde uit Schema 59 aan te passen.

We zijn ons ervan bewust dit extra deelonderzoek voor Natuurkunde een klein aantal leerlingen betreft, maar het lijkt zeer onwaarschijnlijk dat de gevonden niveaudaling van de kandidaten veroorzaakt wordt doordat de vragen van nu versus 30 jaar geleden uit de vraagparen over de hele linie moeilijker zijn.

A-D Het niveau van het CE is gedaald

We hebben de resultaten van de bovenstaande 3 paragrafen nog eens samengevat (Schema 67). Als we over de hele linie kijken, concluderen we dat het niveau van vwo CE Wiskunde A of B zeer aanzienlijk en van Natuurkunde aanzienlijk is gedaald, waardoor meer leerlingen nu een voldoende halen voor deze vakken, terwijl het niveau van vwo CE-kandidaten lijkt te zijn afgenomen.

Specifiek hebben we ook geprobeerd elk van bovenstaande factoren te kwantificeren met een simpele doorrekening. Deze vertaling is nooit helemaal precies, maar het is wel mogelijk om een orde van grootte te geven (Schema 69 tot en met 71). Uit onze analyse blijkt dat de stijging van de CE-cijfers grotendeels verklaard zou kunnen worden uit de hierboven genoemde factoren. We denken dat deze analyse binnen een ordegrootte van zeg 0,2 punt op het cijfer een verklaring kan geven, waarbij we aantekenen dat de onzekerheid in onze schattingen waarschijnlijk ook in deze ordegrootte ligt.

Naast het CE, is er ook nog een School Examen (SE). Een deel van de stof die verdwenen is van het CE wordt nog wel getoetst op het SE. Het kan zo zijn dat stof uit het CE is verdwenen, maar dat het SE de afgelopen 30 jaar is verzwakt. Om dit te verifiëren hebben we ook de stofverandering van het SE in kaart gebracht. Het blijkt dat ook de stof in het SE de afgelopen 30 jaar netto aanzienlijk is afgenomen: met ongeveer 45% bij Wiskunde B, met ongeveer 25% bij Wiskunde A en met 20-25% bij Natuurkunde (Schema 72). De details van deze schattingen zijn te vinden in Schema 73 tot en met 75. Een ander probleem van het schrappen van wiskunde- en natuurkundestof uit het CE is dat een herhaling mist waardoor de stof minder blijft hangen. Wiskundige en natuurkundige vaardigheden zijn veelal hiërarchisch, ofwel ze bouwen op, en hiermee wordt een kans op nog een keer leren en internaliseren ontnomen. Verder beperkt de stofvermindering het aantal stof-vraagcombinaties waarmee bij leerlingen overkoepelend inzicht getest kan worden zoals hierboven al vermeld.

Ten slotte ...

... hebben we nog geanalyseerd op welke wijze Wiskunde en Natuurkunde zijn ingepast in het totale curriculum. Hierbij vielen ons volgende zaken op:

Profiel Natuur & Techniek: Wiskunde D is niet verplicht en een 2^e moderne vreemde taal wel, terwijl Wiskunde D relevanter is voor een groter aantal vervolgstudies die aansluiten op dit profiel. Daarnaast kan een 2^e moderne vreemde taal een blokkade vormen voor een deel van de getalenteerde bèta's die minder talig zijn om een vwo diploma te halen.

Profiel Natuur & Gezondheid: Slechts 1 Wiskunde A of B (niet beide) kan gekozen worden en Wiskunde D is geen keuzevak. Een 2^e moderne vreemde taal is wel verplicht. Wiskunde is relevanter voor een veel groter aantal vervolgstudies. Daarnaast kan een 2^e moderne vreemde taal een blokkade vormen voor een deel van de getalenteerde bèta's die minder talig zijn om een vwo-diploma te halen.

Profiel Economie & Maatschappij: Er kan alleen gekozen worden tussen Wiskunde A en B (niet beide) en Wiskunde D is geen keuzevak, terwijl een uitgebreid Wiskundepakket relevant is voor een aantal vervolgstudies in dit profiel (bijvoorbeeld econometrie).

Profiel Cultuur & Maatschappij: Wiskunde C is een verplicht CE-vak. Wiskunde C bevat het "minimum" wat nodig is voor leerlingen om te functioneren in de maatschappij.

We begrijpen dat de keuzes voor het curriculum altijd gebaseerd zijn op afwegingen tussen verschillende belangen, verschillende vakken en ook afhankelijk is van de technische mogelijkheden om een profiel in het rooster te passen. Dit onderzoek strekt zich dan ook niet uit tot een integrale afweging van het curriculum.

4.4

Mogelijke paden voorwaarts

Dit onderzoek is in eerste instantie gericht op het verzamelen van de feitenbasis om te bepalen of het niveau van het Nederlandse onderwijs in Wiskunde en Natuurkunde daadwerkelijk is afgenomen. We hebben daarom geen integraal plan met aanbevelingen ontwikkeld. In plaats daarvan willen we daarom de onderstaande drie suggesties aanbieden als een uitnodiging aan partijen in het veld, waaronder diegenen die verantwoordelijk zijn voor het onderwijsbeleid, om een alomvattend en passend plan te ontwikkelen om positieve kwaliteitstrendbreuk te realiseren.

Ten eerste ...

... is het nog onvoldoende bekend en erkend dat de resultaten van het vwo teruglopen. Borging van deze basisvaardigheden voor vwo Wiskunde en Natuurkunde zijn fundamenteel voor onze mondiale positie als kenniseconomie en naast welvaart ook voor ons welzijn. Een maatschappelijk debat met de belangrijkste betrokken partijen lijkt een noodzakelijke eerste stap om een beter inzicht te krijgen in het dalende niveau van ons onderwijs en de middelen die kunnen worden ingezet om deze trend te keren. In de coronatijd heeft dit debat begrijpelijkerwijs meer op de achtergrond gestaan. Uit ons onderzoek blijkt dat wat betreft het vwo Wiskunde zeer aanzienlijk en Natuurkunde het niveau aanzienlijk is afgenomen van zowel het CE zelf als van de CE-kandidaten, ook al voorafgaand aan de coronatijd. Lockdowns en schoolsluitingen hebben dit de afgelopen jaren waarschijnlijk verergerd. Hieruit kan worden geconcludeerd dat er een flinke interventie nodig is, maar ook dat een langjarige inspanning nodig is. Deze interventie om de trend te keren kan alleen plaatsvinden als daar een aanzienlijk draagvlak voor is.

Ten tweede ...

... moet in het verlengde van de verhoogde ambitie een onderzoek gestart worden naar de investeringen die benodigd zijn om dit niveau te verhogen. Daarvoor moet dit onderzoek

Ten derde ...

uitgebreid worden naar andere CE-vakken (in het bijzonder Nederlands), het gehele vwo-eindexamen (CE + SE) en andere onderwijsvormen met een CE (havo en vmbo).

... zou een positieve kwaliteitstrendbreuk gerealiseerd moeten worden door enerzijds de lat hoger te leggen, en anderzijds het onderwijssysteem te versterken.

Het hoger leggen van de lat is goed voor de kwaliteit van het onderwijs op het vwo en hopelijk werkt de niveauverhoging ook uiteindelijk door bij de andere onderwijsvormen (havo en vmbo). Wiskunde en Natuurkunde zijn kernvakken en leggen een beter fundament voor de basisvaardigheden, ze helpen om beter te functioneren in de maatschappij, en goede resultaten op deze vakken bevorderen de instroom naar het hoger onderwijs en helpen het hoger onderwijs een hoger niveau gediplomeerde bachelor- en masterstudenten op te leiden. Deze kernvakken moeten dus op niveau zijn om studenten beter te kunnen laten presteren in het vervolgonderwijs. De OECD verwacht op basis van hun onderzoek dat veeleisender onderwijsprogramma's leerlingen beter kunnen laten presteren. Hoge verwachtingen zijn ook volgens meta-onderzoek een belangrijk middel om leerlingen te helpen goed te presteren. Een betere basis in deze vakken stellen de Nederlandse vwo-leerlingen ook in staat om internationaal te concurreren.

Specifiek zouden beleidsmakers de volgende acties kunnen overwegen om de lat hoger te leggen:

- Maak de stof die alleen getoetst wordt in het SE ook weer CE-stof (dit is een goed herhaalmoment en leert het combineren van de stof). Uitzondering hierop zijn onderdelen waarvoor het format van het CE zich minder leent, bijvoorbeeld een natuurkundig practicum.
- Breng (deels) "oude" CE-stof weer terug in het CE-programma.
- Breng de CE-lengte weer terug tot 90 scorepunten (10-15 scorepunten meer aan inhoudelijk wis- en natuurkundig werk).
- Verhoog de complexiteit (aantal denkstappen) en moeilijkheidsgraad van de CE-vragen.
- Verminder het aantal makkelijke vragen en breng minimaal 2 à 3 echt moeilijke vragen extra terug per CE (dit biedt meer uitdaging voor de meest getalenteerde vwo-leerlingen).
- Herijk via de N-term de CE's op een gemiddeld cijfer tussen de 6,0 en 6,5
- Breng een deel van uit het CE Wiskunde A en B verdwenen stof weer terug in het CE door Wiskunde D een CE-vak te maken (of een soortgelijk alternatief), en een kernvak voor het profiel Natuur en Techniek (i.p.v. de 2^e moderne vreemde taal): Wiskunde D is voor vervolgstudies voor dit profiel een veel relevanter vak dan een 2^e moderne vreemde taal (naast Nederlands en Engels). Het zou nuttig kunnen zijn om Wiskunde D als CE-vak te hebben omdat dit dezelfde onafhankelijke niveaumeting is voor iedereen, meer gewicht en status aan het vak geeft, en meer testen van toepassen bij verschillende CE-onderdelen tegelijkertijd betekent. Nu doen slechts 2000 à 3000 leerlingen per jaar examen in dit vak. Zelfs als we deze leerlingen Wiskunde D meetellen bevat het vwo-curriculum nu fors minder Wiskunde. Wiskunde D zou een uitdagend CE zijn. Voorheen (30 jaar geleden) koos 15-20% van de vwo leerlingen Wiskunde A en B. Dit zou de doelgroep kunnen zijn. De vraag is of leerlingen in zijn algemeenheid niet meer keuzevrijheid zouden moeten krijgen om hun vakkenpakket samen te stellen.

Als tweede onderdeel van het verhogen van het kwaliteitsniveau moet het onderwijssysteem structureel versterkt worden.

De hierboven genoemde suggesties kunnen uiteraard niet van het ene op het andere jaar doorgevoerd worden zonder na te denken over de consequenties. Abrupte veranderingen van de norm leiden tot ongewenste effecten zoals kinderen die onterecht niet terecht kunnen in

het vervolgonderwijs. Dit soort veranderingen vergen daarom zorgvuldige voorbereiding van leraren, leerlingen, schoolbesturen, schoolleiders en andere actoren in het onderwijssysteem.

In het rapport *Een verstevigd fundament* is al opgeroepen tot een substantiële inspanning om het Nederlands onderwijs te verbeteren, ingezet door de sector maar nauw begeleid door het ministerie. Dit zal inspanningen vereisen binnen de hele onderwijsketen: door de minister, het ministerie, de Inspectie, de raden, schoolbesturen, schoolleiders, leraren en ondersteunende organisaties. Dit zou onder andere het volgende kunnen omvatten:

- Laat – in het kader van het belang van een onafhankelijke toezichthouder die de kwaliteit van het (examen)systeem als geheel in de gaten houdt – de Inspectie op stelselniveau actief helpen bij de normhandhaving en het controleren van het CvTE.
- De verankering en continuering van dit onderzoek volgens deze nieuwe onderzoeksmethode in het onderwijsstelsel te verankeren.
- De uitrol van de verbetercultuur naar alle scholen te versnellen, met extra aandacht voor de kwaliteit van leraren en “evidence-informed” werken, zodat alle scholen ook in staat worden gesteld het beste uit zowel de leraren als de leerlingen te halen.
- Extra middelen inzetten om het lerarentekort op te lossen.

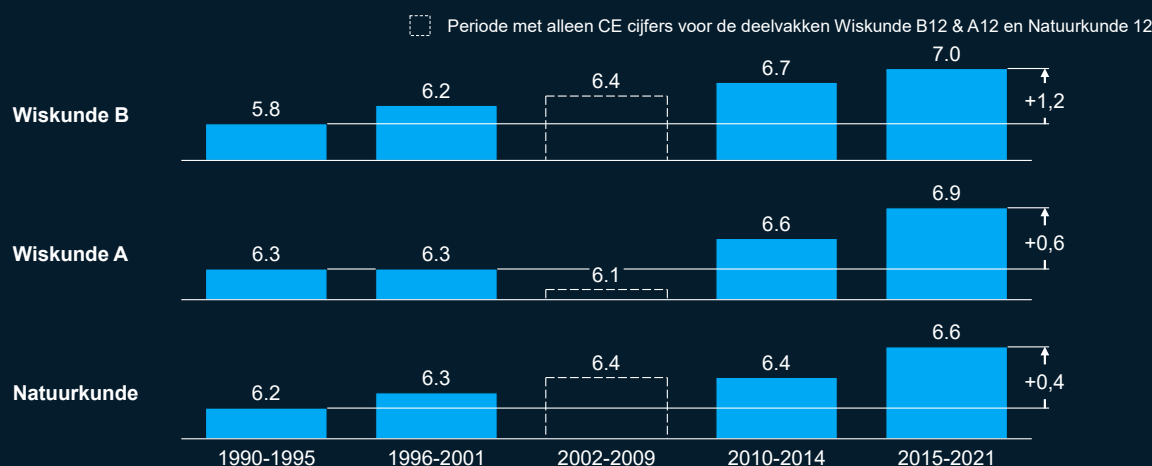
We verwijzen daarbij naar *Een verstevigd fundament voor iedereen* voor een langere lijst mogelijke actiepunten, eerste ideeën over de benodigde financiële middelen en over de volgorde van maatregelen.

05

Ondersteunende schema's

Schema 1

De afgelopen 30 jaar zijn de vwo Centraal Examen (CE) cijfers voor Wiskunde en Natuurkunde met 0,4-1,2 punt gestegen



Noot 1: Waargenomen stijgingen van 1990-1995 tot en met 2015-2021 zijn statistisch significant. Voor de jaren 2002 t/m 2009 zijn de CE cijfers voor de deelvakken Wiskunde B12 en A12 en Natuurkunde 12 gebruikt. Deze CE cijfers zijn minder representatief, doordat slechts een deel van de populatie die Wiskunde B (B1 of B12), Wiskunde A (A1 en A12) en Natuurkunde (1 en 12) Centraal Examen in dit deelvak deed. Dit waren met name leerlingen van het profiel N&T voor Wiskunde B12 en Natuurkunde 12 en van het profiel E&M voor Wiskunde A12. In 2020 was er geen CE en 2021 was het "Corona" examenjaar.

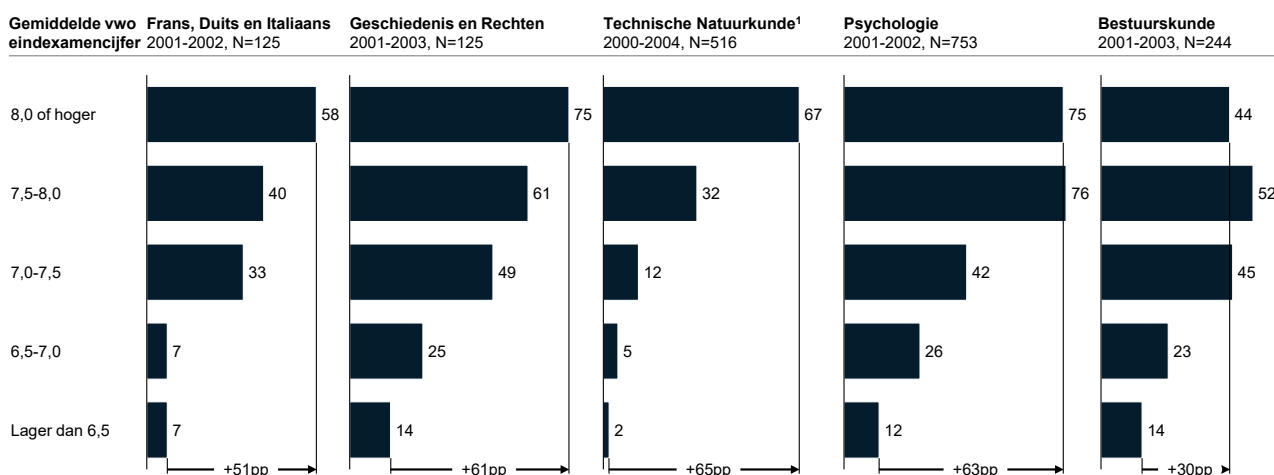
Noot 2: Naast Natuurkunde zijn ook de CE cijfers voor de andere Natuurwetenschappenvakken afgelopen 20-30 jaar gestegen. Op basis van gevonden publiek beschikbare data: Scheikunde van 6,4 in 1990-1994 naar 6,6 in 2015-2021, Biologie van 6,1 in 1990-1994 naar 6,3 in 2015-2021 en Aardrijkskunde van 6,2 in 2000-2005 naar 6,3 in 2015-2021 (voor Aardrijkskunde geen publieke data gevonden voor de periode 1990-1995).

Bron: Examenverslagen voor de periode 2000-2021; Vakbladen Euclides (Wiskunde) en NVOX (Natuurkunde) voor de periode 1990-1999; Analyse door auteurs

Schema 2

Het vwo eindexamencijfer is een goede voorspeller voor het succes in vervolgstudies

Percentage studenten dat propedeuse in één studiejaar haalt ("P in 1"); periode 2000-2004



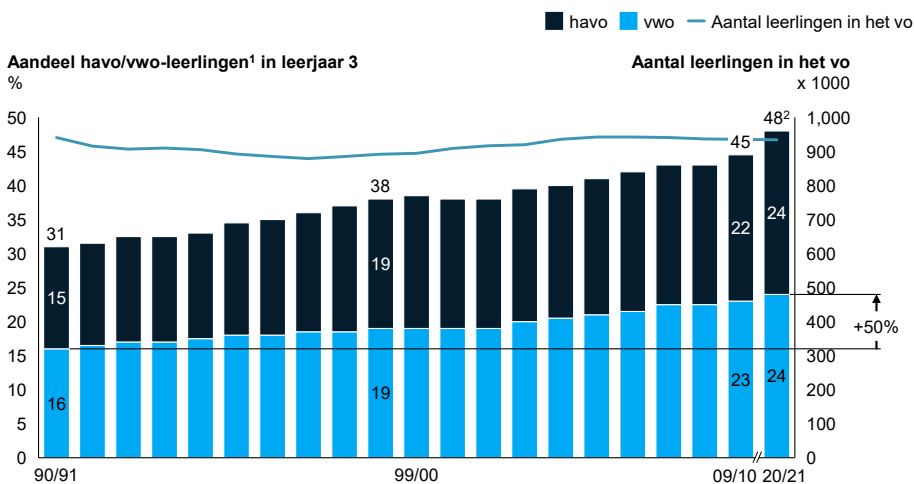
Noot: De data zijn van de periode 2000-2004. Een soortgelijk verband is gevonden door Van der Heijden, Hessen en Wubbels (2012, Studietoetsen of -falen van eerstejaarsstudenten voorspellen: een nieuwe aanpak).

1. Gemiddelde cijfer voor Wiskunde B en Natuurkunde

Bron: TU Delft onderzoek door dr. ir. Eric Logtenberg; Rapport 148: Presenteren in het vwo en het HO (ICLON, Universiteit Leiden en Cito, maart 2005); Analyse door auteurs

Schema 3

Het aandeel vwo (en havo) leerlingen in leerjaar 3 van het voortgezet onderwijs (vo) is sinds 1990 met ongeveer 50% toegenomen



Noot 1: De data voor de periode tussen 09/10 en 20/21 is beschikbaar, maar niet weergegeven. De toename heeft voornamelijk in de periode tussen 90/91 en 09/10 plaatsgevonden

Noot 2: De twee meest rechter staafjes in de grafiek zijn een overgang van 10 jaar

1. Exclusief voortgezet speciaal onderwijs (vso) en speciaal onderwijs (so)

2. Van de 193.527 zijn niet meegeteld: 5.981 in "algemene leerjaren 1-2-3" en 5.691 op het praktijkonderwijs. Van de meegetelde 181.855 leerlingen in leerjaar 3 zaten er 44.177 op de havo, 43.359 op het vwo. Deze methodiek is hetzelfde voor de andere jaren

Bron: CBS rapport Onderwijs de Cijfers (20-1-2011) van Chantal Melser en CBS Statline data 2020-2021; Analyse door auteurs

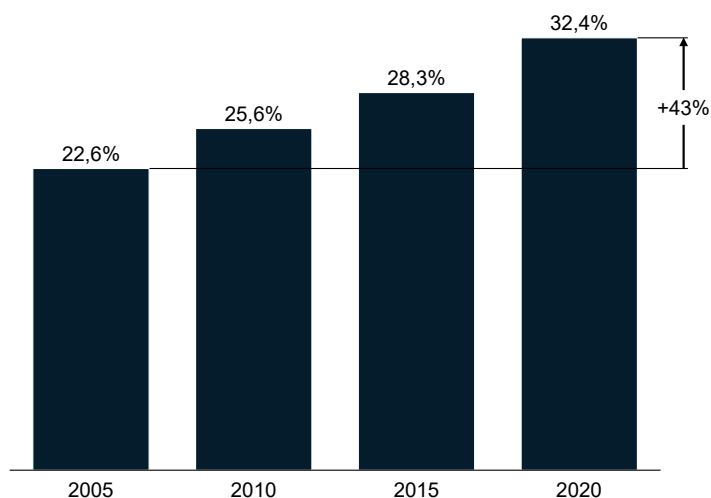
In 1990 zat 31% van de leerlingen in leerjaar 3 op het havo of vwo. Dit is toegenomen tot 48% in 2020/2021. Dit is een relatieve stijging van ongeveer 55%.

In 1990 zat 16% van de leerlingen in leerjaar 3 op het vwo. Dit is toegenomen tot 24% in 2020/2021. Dit is een relatieve stijging van ongeveer 50%.

Schema 4

Het aandeel hogeropgeleiden in de beroepsbevolking is sinds 2005 met 43% toegenomen

Percentage van de beroepsbevolking dat minimaal een WO of HBO Bachelor diploma heeft



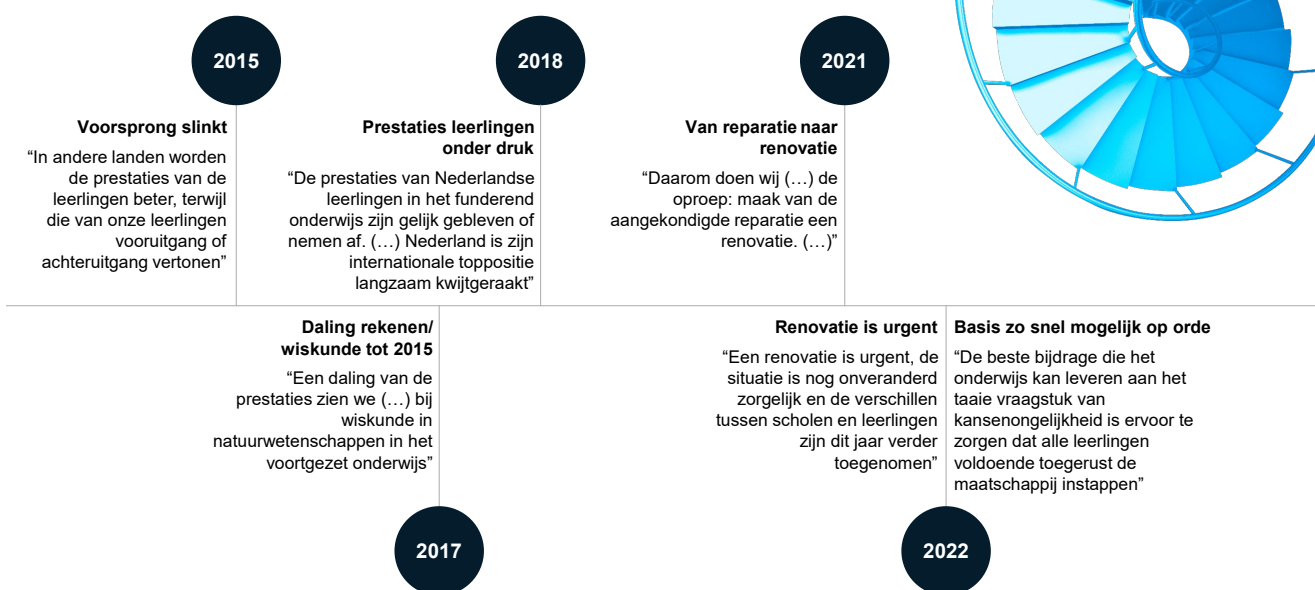
Bron: CBS data "Bevolking; hoogst behaald onderwijsniveau en onderwijsrichting"; Analyse door auteurs

Het hogere aandeel havo en vwo leerlingen in het funderend onderwijs (en de hogere CE cijfers voor Wis- en Natuurkunde) hebben meer hogeropgeleiden opgeleverd. Het aandeel hogeropgeleiden (minimaal een HBO of WO bachelor) is sinds 2005 gestegen van 22,6% naar 32,4%. **Dit is een relatieve stijging van 43%.**

Er is niet onderzocht of het niveau van de hoger opgeleiden minimaal gelijk is gebleven (valt buiten de scope van dit onderzoek).

Schema 5

De Inspectie heeft in verschillende publicaties aangegeven dat de onderwijsresultaten dalen



Bron: Verschillende publicaties van De Staat van het Onderwijs door de Inspectie van het Onderwijs

Schema 6

In een recent advies uit de Onderwijsraad zorgen over de beheersing van rekenen (en taal)

Selectie uit de inhoudsopgave van een recent Advies van de Onderwijsraad



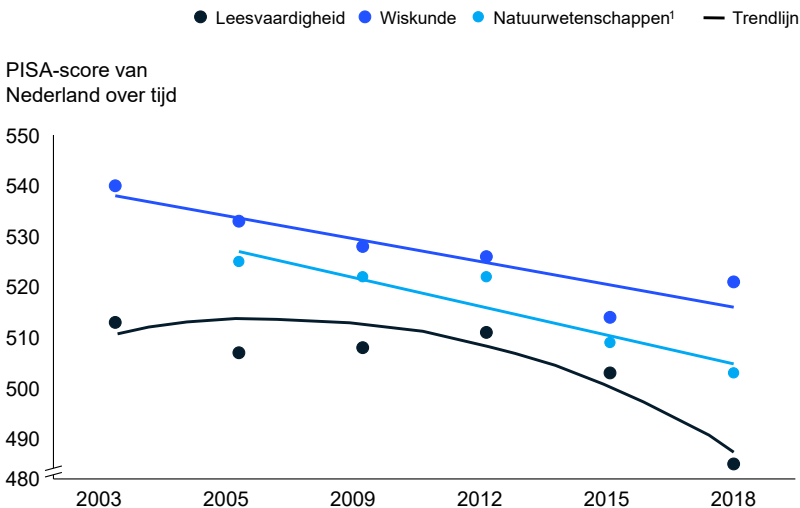
1	Aanleiding	
	Zorgen over de beheersing van taal en rekenen	11
1.1	Goede beheersing van taal en rekenen is onontbeerlijk	12
1.2	Adequaat zicht op de beheersing van taal en rekenen ontbreekt	13
1.3	Beheersing taal en rekenen geeft reden tot zorg	16
1.4	Beheersing taal en rekenen ook beïnvloed door factoren buiten het onderwijs	19
1.5	Adviesvraag: Hoe kan het onderwijs de beheersing van taal en rekenen duurzaam verbeteren?	20
2	Advies	
	Focus aanhoudend op taal en rekenen	23
2.1	Curriculum en toetsing van taal en rekenen zijn nu onevenwichtig	26
2.2	Het taal- en rekenonderwijs in de klas kent tekortkomingen	29
2.3	Lerarenopleidingen leiden niet goed genoeg op voor onderwijs in taal en rekenen	33
2.4	Professionalisering in taal en rekenen is ontoereikend	37
2.5	Kennis uit de wetenschap komt onvoldoende terecht in het onderwijs	39

Bron: Advies van de Onderwijsraad "Taal en rekenen in het vizier" (2022)

Schema 7

PISA laat over afgelopen 15-20 jaar een significante niveaudaling van 15-jarigen zien over de gehele linie van het funderend onderwijs

Verandering in scores tussen 2003 en 2018, per PISA-onderdeel



1. Combinatie van Natuurkunde, Scheikunde, Biologie en Aardrijkskunde ("Science")

Bron: PISA; McKinsey onderzoek naar de doelmatigheid en toereikendheid van het funderend onderwijs (primair en voortgezet): "Een verstevigd fundament voor iedereen, April 2020" (Paul Rutten et al); Aanvullende analyse door de auteurs

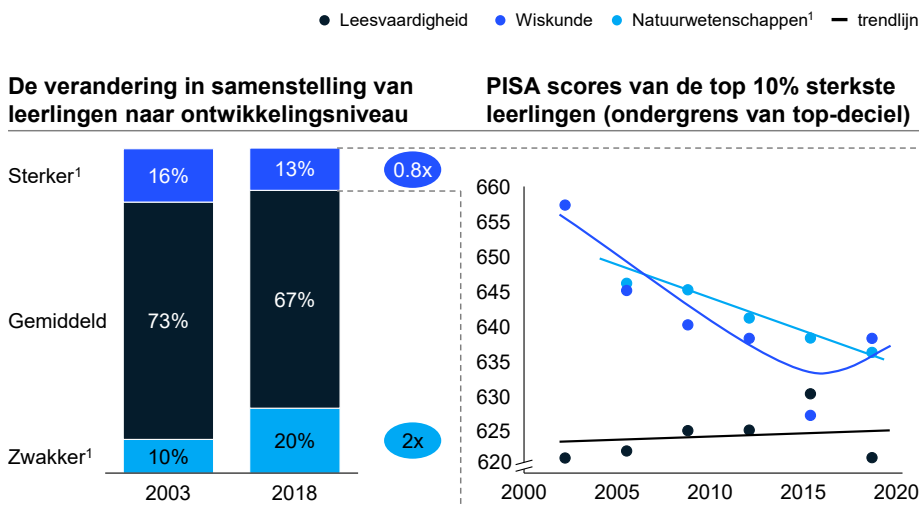
Zowel Wiskunde als Natuurwetenschappen zijn sinds 2003 met ongeveer 20 PISA punten gedaald. Dit komt in PISA termen ongeveer **overeen met een half jaar tot een jaar aan voortgang**.

De daling per onderwijsvorm sinds 2003 is voor het vwo 13 punten, de havo 22 punten en het vmbo 33 punten.

Ook relatief is Nederland gezakt: voor Wiskunde van plek 4 in 2003 naar plek 9 in 2018 en voor Natuurwetenschappen van plek 8 in 2003 naar plek 16 in 2018. Het aantal deelnemende landen is wel toegenomen.

Schema 8

Ook het aandeel "sterke" vo-leerlingen is met 20-25% gedaald en hun PISA scores zijn met ~15 punten afgenomen voor de exacte vakken



1. Combinatie van Natuurkunde, Scheikunde, Biologie en Aardrijkskunde ("Science")

2. "Sterker" is hoogste twee mogelijke PISA score categorieën (niveau 5 of 6) en "zwakker" is laagste twee mogelijke PISA score categorieën (niveau 1 of 2)

Bron: PISA; OESO; IEA; McKinsey onderzoek naar de doelmatigheid en toereikendheid van het funderend onderwijs (primair en voortgezet): "Een verstevigd fundament voor iedereen, April 2020" (Paul Rutten et al); Aanvullende analyse door de auteurs

Het aandeel sterke leerlingen in het VO gedaald van **16,4%** naar **12,7%**. Dit is een afname van ongeveer **23%**.

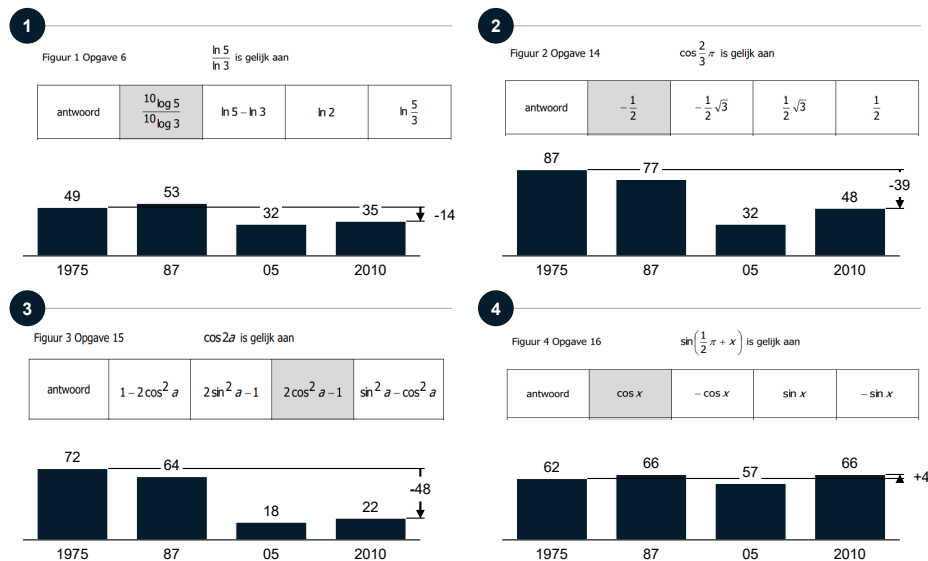
Ook is hun PISA score voor **Wiskunde** en **Natuurwetenschappen** met ~15 PISA punten gedaald.

Bij **leesvaardigheid (taal)** geen duidelijke trend zichtbaar

Schema 9

Instapniveaumetingen aan de TU Delft lijken te duiden op een niveaudaling voornamelijk na 1987

Percentage goed op 4 ankervraagstukken¹ bij instapniveautoets aan TU Delft



1. Het rapport verstrekt door de TU Delft bevatte alleen de scores op deze 4 ankervraagstukken in 4 jaren (1975, 1987, 2005, 2010)

Bron: TU Delft onderzoek door dr. Wim Caspers; Analyse door auteurs

De TU Delft heeft instapniveaumetingen gedaan (jaarlijks afgenomen in de periode 2004-2010). Deze instaptoetsen worden zonder rekenmachine afgenomen.

In de figuur links vind je de resultaten van 4 ankeropgaven door de jaren heen.

De grootste scoredaling heeft plaats-gevonden na 1987 en voor 2005. In 2005-2010 is er een licht herstel van de scores.

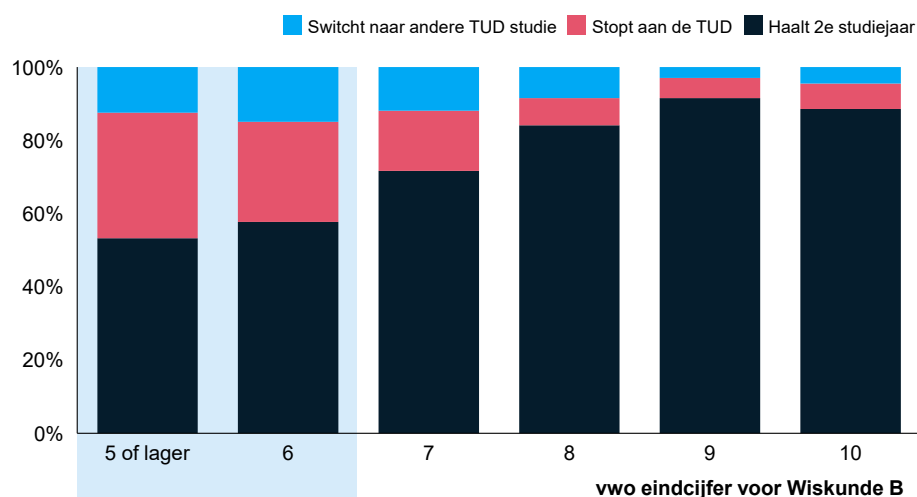
In 1975 en 1987 waren er 8 keuzeopties en in 2005 en 2010 4 keuzeopties.

In de periode 2005-2010 is het niveau licht toegenomen, maar ruim niet tot de scores van voor 1987.

Schema 10

Een vwo eindexamencijfer voor Wiskunde B lager dan een 7 geeft onvoldoende garantie voor het succes bij technische vervolgstudies

TU Delft (TUD) cohort (alle technische studies) aangevangen in 2014¹



1. Het onderzoek was onder studenten van alle TU Delft studies die in 2014 waren gestart. Dit zijn er >1.000, dus statistisch significant

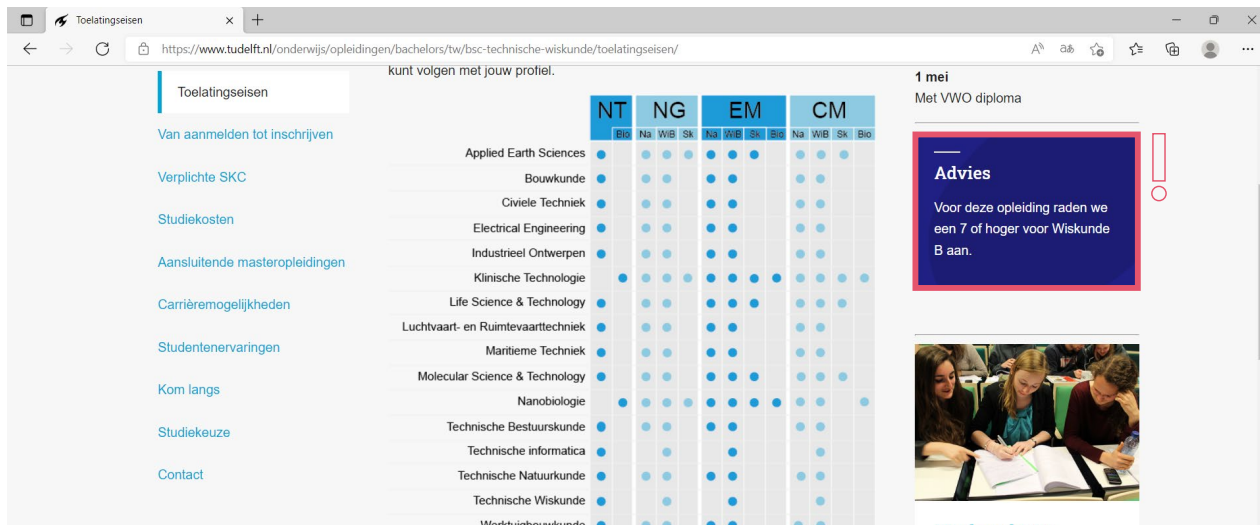
Bron: TU Delft onderzoek door dr. ir. Eric Logtenberg. Gepubliceerd in Delta (2016): <https://www.delta.tudelft.nl/article/tu-meer-studiesucces-met-hoger-Wiskundecijfer>; analyse door auteurs

Met een 6 (of lager) voor Wiskunde B heb je ongeveer 50-55% kans om het eerste jaar van een TU Delft studie te halen, bij een 7 is dat 70% en bij een 9 of 10 is dat 90%. Hetzelfde beeld zie je voor de kans om je bachelor in drie jaar te halen

Bij studies als Technische Natuurkunde of Technische Wiskunde is het verband met Wiskunde B sterker dan bij studies als Bouwkunde, Industrieel Ontwerpen of Technische Bestuurskunde

Schema 11

De TU Delft ontmoedigt studenten met een examencijfer lager dan een 7 voor Wiskunde B om te komen studeren



Bron: Website van de TU Delft <https://www.tudelft.nl/onderwijs/opleidingen/bachelors/tw/bsc-technische-wiskunde/toelatingseisen/>; Analyse door auteurs

Schema 12

Dat een vwo diploma onvoldoende garantie geeft voor studiesucces kan deels worden verklaard doordat scholen 10-15 jaar geleden mindere CE prestaties compenseerden met te makkelijke School Examens (SE): in de grafiek rechts zie hoe lager het CE-cijfer (meer links op de horizontale as), des te groter het verschil tussen SE en CE-cijfer (meer bovenaan op de verticale as)

Kansengelijkheid: afhankelijk van de schoolkeuze was het in 2005-2008 het eenvoudiger je vwo diploma te halen

In 2005-2008 lag het SE cijfer 0,4 punt hoger dan het CE. In 2019 waren de cijfers gelijk.

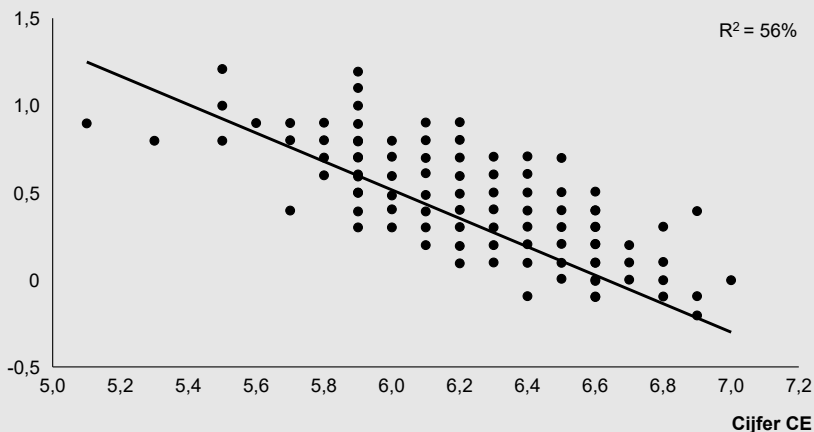
Het verschil tussen CE en SE cijfer is in de periode 2005-2020 kleiner geworden. Dit kan ook gekomen zijn door een toegenomen CE-cijfer i.p.v. een afgenomen SE cijfer. Voor vwo Wiskunde A en B lag het CE cijfer in 2018 hoger dan het SE cijfer (bron: Wiskunde E-brief nr 832, 13 januari 2019).

De sturing door Inspectie dat het verschil tussen SE en CE niet groter dan een half punt mag zijn is recent losgelaten.

Bron: Elsevier 2008; Analyse door auteurs

Verskil SE en CE cijfer (positief is hoger SE cijfer)

Alle vwo afdelingen van scholen in Nederland; gemiddelde 2005-2008; N = 498



Elke vwo afdeling is weergegeven met een stip. Er kunnen meer dan één vwo-afdelingen op een stip weergegeven zijn

Schema 13

De Faculteit Luchtvaart- en Ruimtevaarttechniek (L&R) van de TU Delft toetst het niveau van aspirant-studenten om plekken te kunnen toekennen voor hun bachelorprogramma

De bacheloropleiding L&R heeft een "numerus fixus" (beperking op de instroom) waarmee het elk jaar een opleidingsplek kan bieden aan 440 nieuwe studenten. Vanaf 2017 hanteert de studie een 100% decentrale selectie.

De toekenning van deze 440 plekken vindt plaats middels een selectieprocedure die uit meerdere onderdelen bestaat, waaronder een cognitieve toets. Deze toets, die 75% van de eindscore bepaalt, bestaat uit 3 deeltolsten:

1) Wiskunde, 2) Natuurkunde, en 3) Introductie tot Aerospace Engineering (aan de hand van eerstejaars studiemateriaal). Na de toets worden de studenten gerangschikt op basis van hun prestaties. De 440 studenten met de hoogst behaalde totaalscores in de selectie krijgen een opleidingsplek aangeboden.

Om toegelaten te kunnen worden tot de studie checkt de universiteit of aspirant-studenten met een verzoek tot inschrijving aan de vooropleidingseisen voldoen (zoals Wiskunde B en Natuurkunde als examenvak). Het aantal aspirant-studenten met een vooraanmelding voor de opleiding is over de afgelopen jaren toegenomen van 1174 (2017) tot 2283 (2022). Hiervan neemt per jaar ongeveer twee derde deel aan de cognitieve toets.

De Wiskunde en Natuurkunde onderdelen van de toets zijn volledig ontworpen op de stof die bekend is voor Nederlandse vwo-leerlingen (in lijn met de vwo-eindtermen van het CE). Dit voordeel geldt niet voor kandidaten met een niet-Nederlandse vooropleiding.

De toets wordt afgenomen vlak voor de vwo CE's (maart van het jaar) zodat de stof bij Nederlandse studenten bekend verondersteld is. Het gemiddelde vwo-eindexamencijfer van Nederlandse instromende studenten was afgelopen 5-10 jaar nagenoeg constant. De Faculteit Luchtvaart- en Ruimtevaarttechniek heeft de wens elk jaar minimaal 50% Nederlandse studenten in te laten stromen, maar kan hier niet op sturen.

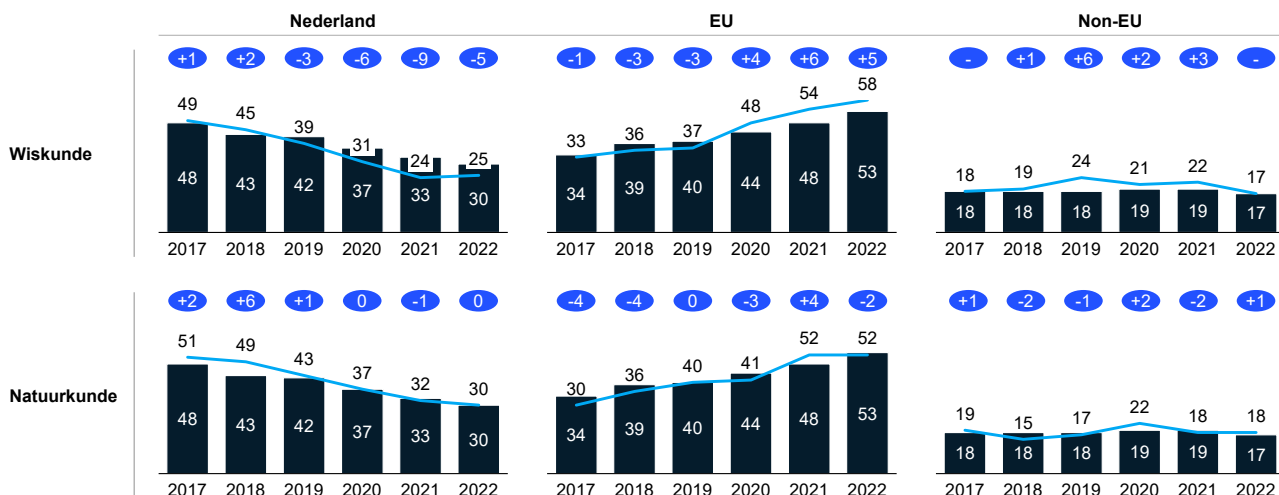
Het aantal vooraanmeldingen en daarmee deelnemers aan de toets is over de afgelopen jaren flink toegenomen. Echter het aandeel van kandidaten uit de EU/EFTA neemt relatief veel harder toe. Daarbij nemen ook de relatieve prestaties in de selectieprocedure toe van kandidaten uit de EU/EFTA, wat ten koste gaat van opleidingsplekken voor Nederlandse studenten.

Bron: Analyse door Luchtvaart- en Ruimtevaarttechniek (ir. Joris Melkert en Floris Keehnen, MA)

Schema 14

L&R: zowel het aandeel deelnemende Nederlandse studenten als hun relatieve prestaties op de toelatingstoets nemen af t.o.v. EU-peers

Over (+) en onder (-) vertegenwoordiging in de top-500 toets scores t.o.v. aandeel toets deelname — % van de top-100 toets scores ■ % van de toets deelnames

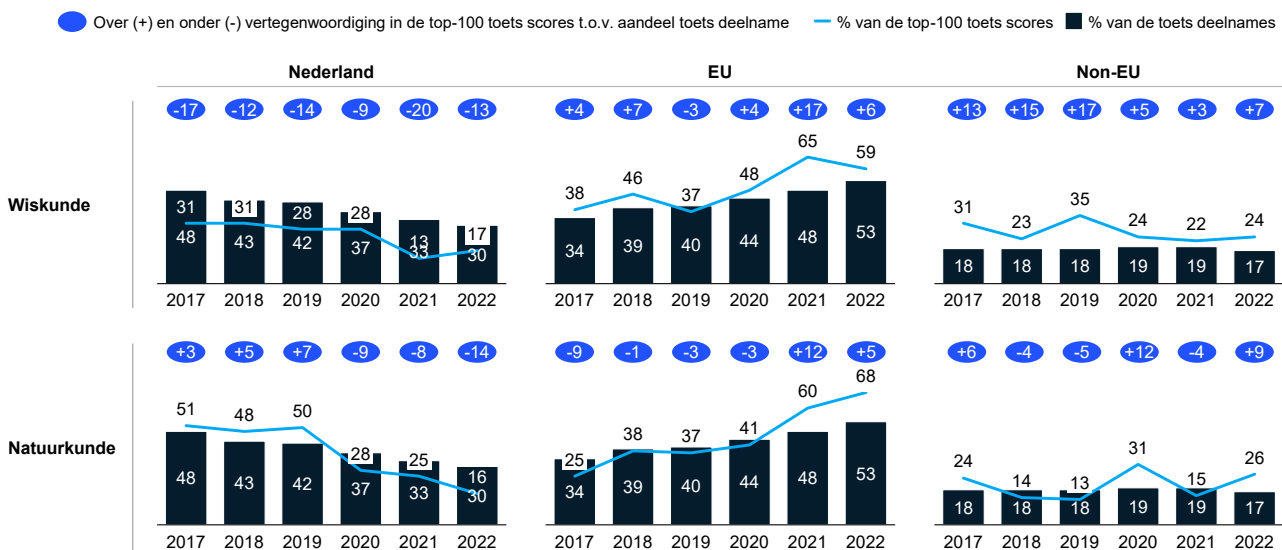


Noot: Het aantal studenten dat heeft deelgenomen aan de toets is toegenomen van 719 (2017) naar 1500 (2022) voor wiskunde en van 714 (2017) naar 1465 (2022) voor natuurkunde. Het aandeel Nederlandse studenten is gedaald met name door een toegenomen toestroom vanuit de EU

Bron: Analyse door Luchtvaart- en Ruimtevaarttechniek (ir. Joris Melkert en Floris Keehnen, MA)

Schema 15

L&R: Nederlandse studenten t.o.v. aandeel in toetsdeelnames ondervertegenwoordigd in de top-100 toetsscores versus EU-peers



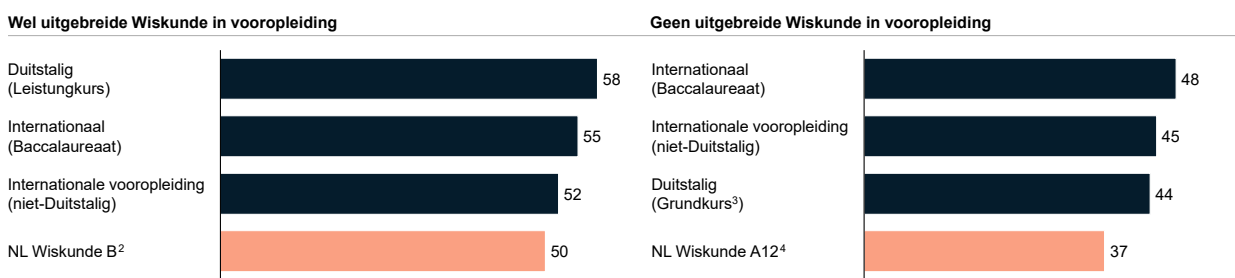
Noot: Het aantal studenten dat heeft deelgenomen aan de toets is toegenomen van 719 (2017) naar 1500 (2022) voor wiskunde en van 714 (2017) naar 1465 (2022) voor natuurkunde. Het aandeel Nederlandse studenten is gedaald met name door een toegenomen toestroom vanuit de EU

Bron: Analyse door Luchtvaart- en Ruimtevaarttechniek (ir. Joris Melkert en Floris Keehnen, MA)

Schema 16

Een internationale vergelijking van wiskundige vaardigheden onder eerstejaars studenten lijkt erop wijzen dat Nederland achter blijft

Score (0-100) op 3TU2005¹ toets onder eerstejaars studenten economie en internationale bedrijfskunde (cohort 2008) aan de Universiteit Maastricht; N = 679



- Citaat uit de publicatie: "Zorg om een niveaudaling van het Nederlandse Wiskunde- en rekenonderwijs is een onderwerp dat volop in de belangstelling staat, en dat instellingen van hoger onderwijs heeft genoopt tot het inrichten van reparatieprogramma's. Bijspijkeronderwijs Wiskunde en het afnemen van instaptoetsen zijn de bekendste onderdelen"
- Een eventueel selectie-effect (de "betere" buitenlandse studenten gaan in Maastricht economie en internationale bedrijfskunde studeren) kan niet worden uitgesloten
- Quote van de opleidingsdirecteur van een Technische Universiteit: "Met de Wiskunde vakken uit onze Master worstelen onze Nederlandse studenten in vergelijking met hun internationale peers die deze vakken met twee vingers in de neus doen"
- Quote van een hoogleraar bij Defensie: "Belgische en Zuid-Europese studenten zijn veel sterker bij Luchtvaart- en Ruimtevaart bij de TU Delft dan hun Nederlandse peers"

1. Wiskundetoets ontwikkeld door drie Technische Universiteiten (Delft, Eindhoven, Enschede)

2. Gewogen gemiddelde van 17 Wiskunde B12 studenten (score: 52) en 33 Wiskunde B1 studenten (score: 48)

3. Gewogen gemiddelde van 115 met Wiskunde in eindexamen (score: 41) en 184 Wiskunde B1 studenten (score: 45)

4. Gebaseerd op Wiskunde A12, dus exclusief Wiskunde A1 studenten

Bron: De rol van de instaptoets door dr. Wim Caspers (TU Delft) en dr. Dirk Tempelaar (Universiteit Maastricht); Interviews; Analyse door auteurs

Dit roept een drietal vragen op ...

- 1** In hoeverre is PISA een adequate ontwikkelingsniveaumeting voor Wiskunde en Natuurwetenschappen (en Taal)?
- 2** Is er een verband te verwachten tussen PISA en het Centraal Examen (CE) voor dezelfde vakken?
- 3** Is het niveau van het vwo CE veranderd in de afgelopen 30 jaar voor Wiskunde en Natuurkunde?

Schema 18

1: PISA is een adequate ontwikkelingsniveaumeting voor het funderend onderwijs voor taal, wiskunde en natuurwetenschappen

Criteria goede toets	Voldoet?	Rationaal
1 Validiteit	✓	- Inhoudsvaliditeit: PISA is niet volledig, maar meet de basisvaardigheden en de belangrijkste aspecten daarvan. - Taal: informatie opzoeken, begrijpen, evalueren en reflecteren - Wiskunde: verandering en relatie, ruimte en vormen, hoeveelheid, onzekerheid en data - Natuurwetenschappen: fenomenen wetenschappelijk kunnen uitleggen, evalueren en ontwerpen van wetenschappelijk onderzoek, wetenschappelijk interpreteren van data en bewijs - Begripsvaliditeit: PISA meet direct bovenstaande vaardigheden in de vragen - Op basis van feedback rondes met nationale experts wordt de vraagselectie continu verfijnd en bijna alle vragen worden zeer relevant genoemd
2 Betrouwbaarheid	✓	- Beide type vragen hebben hoge mate van betrouwbaarheid - Meerkeuzevragen (60%) worden per definitie op dezelfde manier beoordeeld - Antwoord open vragen (40%): a) simpele door AI en b) complexer door getrainde correctoren - Een nagenoeg en een zo goed mogelijke absolute niveaumeting: 1) een groot deel van de vragen is identiek en 2) blinde ijkingsvragen (Anchor in Package) worden gebruikt wanneer een domein wordt vernieuwd (dit gebeurt elke 9 jaar)
3 Specificiteit	✓	De vragen kunnen slechts worden beantwoord als de leerling de onderliggende parate kennis en vaardigheden heeft voor taal (bijv. woordenschat), Wiskunde (bijv. formules als de stelling van Pythagoras, grafieken interpreteren) en natuurwetenschappen (bijv. biologische celstructuur, wetten van Newton, chemische reacties) en deze kan toepassen in "dagelijkse" situaties
4 Moeilijkheid	✓	- Vragen van verschillende moeilijkheidsniveau (1 t/m 6) worden gebruikt - Bij taal werkt PISA nu met vragen die zich aanpassen aan het niveau van de kandidaat
5 Discriminerend vermogen	✓	De gemiddelde PISA score is 500 met 100 als standaardafwijking. PISA scores in Nederland variëren van ongeveer 200 tot 700 wat een spreiding van 5 standaardafwijkingen is. Hiermee kun je goed onderscheiden. Dit is een goede verdeling voor ongeveer 99% van de populatie
6 Transparantie	✓	De kennis en vaardigheden die in het funderend onderwijs worden onderwezen zijn afdoende om de PISA toets goed te kunnen maken. Voorbeeld PISA toetsen zijn beschikbaar en worden in Nederland ook gebruikt
7 Beschikbare tijd	✓	Er is twee uur beschikbaar voor de test. Dit is voldoende voor de 3 domeinen en vragen. Nederlandse leerlingen beantwoorden 99% van de vragen (voldoende tijd) en over 97% van de vragen wordt meer dan 5 seconde nagedacht (geen doorklikken)
8 Technische kwaliteit	✓	- Taal: de gebruikte taal is simpel en duidelijk en de situaties herkenbaar (dagelijkse situaties) - Toets: goede gangbare toets lay-outs worden gebruikt (conform gerenommeerde andere toetsen). Wel is recent overgestapt van schriftelijk naar digitaal toetsen

Bron: SLO, PISA rapportages en data rondom methodologische verantwoording en een interview met Andreas Schleicher (PISA Director for Education and Skills Special Advisor on Education Policy to the Secretary-General); Analyse door auteurs

1: PISA is – zoals gebruikelijk voor een vergelijkend onderzoek – aan een vorm van statistische onzekerheid onderhevig

Duiding van recente CPB-publicatie over de positie van Nederland in internationale onderwijsrankings

Kernpunten uit recent CPB rapport

Internationaal peilings-onderzoek duidt op een verslechtering van de basisvaardigheden onder Nederlandse jongeren.

In deze publicatie onderzoeken we in hoeverre deze onderzoeken een goede graadmeter zijn voor onderwijsprestaties.

We constateren dat de positie van Nederland gevoelig is voor instroom van nieuwe deelnemende landen en methodologische problemen.

Het is onbekend of de optelsom van deze problemen de positie van Nederland gunstig of ongunstig beïnvloedt.

Duiding van het CPB rapport

Het recente CPB rapport “Een blik op de Nederlandse positie in internationale onderwijsrankings” stelt dat PISA aan een vorm van statistische onzekerheid onderhevig is. Dit is juist, net als elk ander vergelijkend onderzoek dat is.

Over de totale richting en de grootte van deze statistische onzekerheid doet het CPB rapport geen uitspraak.

- **Totale richting:** De door PISA waargenomen niveaudaling kan dus zowel meer als minder zijn. Sommige door het CPB rapport omschreven equivalentieproblemen pakken gunstig voor Nederland uit (bijvoorbeeld relatief weinig niet-serieuze leerlingen bij natuurwetenschappen en een hoge non-respons) en andere ongunstig (bijvoorbeeld verandering in het toetsingskader bij taal in 2018).
- **Totale grootte:** Het CPB rapport geeft aan dat PISA al het mogelijke doet om deze statistische onzekerheid te minimaliseren en transparant te maken om over de tijd te kunnen vergelijken middels “link errors”. Voorbeelden van het minimaliseren van de statistische onzekerheid is het gebruik van ankervraagstukken wat volgens het Cito de meest nauwkeurige kalibratie hiervoor is en is het gebruik van een zorgvuldige vertaalprocedure. Ondanks deze inspanningen kan PISA niet de niveaumeting geheel absoluut maken en blijft er een mate van statistisch onzekerheid bestaan.

Aangezien enerzijds PISA de best beschikbare kalibratiemethoden gebruikt en er geen sprake is van een systematische onzekerheid in het nadeel Nederland en anderzijds de daling van de scores voor wiskunde en natuurwetenschappen dermate groot en consistent zijn over een langere periode (ook voor de 10% sterkste leerlingen) is het onwaarschijnlijk dat daling in Nederlandse (vwo) PISA scores door statistische onzekerheid veroorzaakt wordt

Bron: CPB publicatie december 2022 – “Een blik op de Nederlandse positie in internationale onderwijsrankings”; Vaardigheidsontwikkelingen volgens PISA en examens (2019) door Cito auteurs Paul van der Molen, Sanneke Schouwstra, Remco Feskens en Marieke van Onna; Analyse door auteurs

2: De Commissie Steur samengesteld door het NRO concludeerde dat de vergelijking tussen PISA en het CE mank gaat en er geen sprake is van teruglopende onderwijsresultaten (1/3)

1. Zijn de prestaties van PISA voortdurend dalend op alle niveaus en alle domeinen?

De commissie constateert dat de prestaties van Nederlandse leerlingen voor PISA niet voortdurend dalen op alle niveaus en alle domeinen, maar wel voor bepaalde niveaus op een aantal domeinen, ook vergeleken met andere OESO-landen. Het algemene beeld van de ontwikkeling van prestaties van Nederlandse leerlingen voor PISA is deels stabiel en deels dalend.

2. Is het gemiddelde cijfer op het centraal examen (CE) gelijkblijvend tot stijgend (in samenhang met de gewijzigde kernvakkenregeling)?

Wij hebben als commissie geconstateerd dat in de periode van 2008 tot 2017 het over alle vakken gemiddelde cijfer op het CE voor het vmbo-bb en vwo aanzienlijk is gestegen (resp. 0,38 en 0,22 punt), voor de overige schooltypen en leerwegen is er een lichtere stijging.

3. Blijven de cijfers op het schoolexamen (SE) gelijk?

De commissie constateert dat in de periode 2008-2017 het gemiddelde cijfer voor alle vakken van het schoolexamen tezamen stabiel is. Tussen 2008 en 2012 daalde het gemiddelde cijfer licht, en daarna trad een lichte stijging op. Hierbij merken we wel op dat de schoolexamens per school anders ingericht kunnen zijn, vastgelegd in het programma van toetsing en afsluiting (PTA). Zowel de ordening in het PTA als de validiteit van de toetsen en evaluaties loopt uiteen.

4. Welke rol speelt de N-term en hoe robuust is deze procedure?

De commissie constateert dat de N-termprocedure een belangrijke rol speelt bij het vaststellen van de examencijfers op het CE over de jaren heen, met beperkte effecten op de hoogte van die cijfers. De commissie is ervan overtuigd dat de N-termprocedure nodig is om op een verantwoorde wijze te examineren en om examencijfers over de jaren heen met elkaar te kunnen vergelijken. Voor een leerling is het van groot belang dat het niet uitmaakt in welk jaar hij of zij het examen maakt. De N-termprocedure is een essentieel correctiemechanisme dat corrigeert voor de moeilijkheid van het examen waardoor de cijfers van jaar tot jaar vergelijkbaar zijn. De daarbij gekozen methodiek (per vak, per niveau) lijkt volgens de commissie binnen de budgettaire ruimte adequaat.⁴

Daarnaast constateert de commissie dat de N-termprocedure over het algemeen robuust is. Dat betekent dat het goed lukt om de examens op een vergelijkbaar moeilijkheidsniveau te houden zodat CE-cijfers over jaren heen goed te vergelijken zijn en dat examenprestaties niet kunstmatig opgehoogd zijn door middel van de N-term. Uit het Cito-onderzoek blijkt dat het effect van de bijstelling van de technische N-term op de uitslag van het examen verwaarloosbaar klein is.

5. Hoe verhouden de metingen van schoolexamen, centraal schriftelijk en PISA zich tot elkaar en wat betekent dit voor het antwoord op vraag 6?

Wij constateren als commissie dat de examens (zowel SE als CE) en PISA niet goed te vergelijken zijn vanwege verschillen tussen gemeten vaardigheden, de leerlingenpopulatie die (deels) afwijkt, de belangen die op het spel staan voor leerlingen en hun daarmee samenhangende motivatie, en daarbij aansluitend de toegenomen nadruk op het CE. Vanwege die toenemende nadruk op het CE is het SE meer op het CE gaan lijken, zoals de commissie Ten Dam heeft geconcludeerd, waardoor de focus binnen het curriculum meer is komen te liggen op het ontwikkelen en toetsen van vakkennis en vakspecifieke vaardigheden binnen de examendomeinen.⁶ Eerder merkten we al op dat de resultaten op het CE een vergelijking over de jaren goed mogelijk maken; door de verschillende invulling die scholen aan de schoolexamens kunnen geven is een vergelijking over de jaren heen voor het SE minder betrouwbaar.

6. Wat is de reële ontwikkeling van de vaardigheden van scholieren in de afgelopen 10 jaar (minimaal voor Engels en wiskunde)?

De commissie baseert zich voor de beantwoording van deze vraag primair op de resultaten van het CE en PISA. Wij constateren dat de vaardigheden van Nederlandse leerlingen zoals getoetst in het CE de afgelopen jaren over het geheel stabiel zijn of stijgen. Specifiek voor Engels betreft dit overwegend een stijging voor alle leerlingen terwijl de vaardigheidsontwikkeling voor wiskunde een divers beeld laat zien voor de verschillende schooltypen en leerwegen. Tegelijkertijd blijven de PISA-resultaten deels stabiel en ten dele dalen ze. Dit laatste wijst mogelijk op een verminderde vaardigheid om algemene kennis op dagelijkse problemen toe te passen.

Bron: Brief getiteld “Aanbieding bevindingen ontwikkeling examenresultaten VO n.a.v. motie Van Meenen c.s.” d.d. 28 maart 2019 aan de toenmalig Minister van OC&W Slob; Analyse door auteurs

2: De Commissie Steur samengesteld door het NRO concludeerde dat de vergelijking tussen PISA en het CE mank gaat en er geen sprake is van teruglopende onderwijsresultaten (2/3)

7. In hoeverre bevorderen/belemmeren de N-term en andere factoren het zicht op het beeld als bij 6 bedoeld?

De commissie concludeert dat de N-term het zicht op de reële ontwikkeling van vaardigheden niet belemmert maar juist bevordert. De N-term zorgt ervoor dat cijfers over de jaren heen een goede weerspiegeling zijn van de vaardigheidsontwikkeling van de leerlingen. Hierbij dient opgemerkt te worden dat de nauwkeurigheid van de normering afhangt van de gebruikte normhandhavingmethode. Grofweg gezegd, is 'Anchor in Package' drie keer zo nauwkeurig als 'pretest en posttest' en is 'pretest en posttest' weer drie keer zo nauwkeurig als 'standaardbepaling'.

De in de afgelopen 10 jaar genomen beleidsmaatregelen (o.a. verscherping van exameneisen) belemmeren enigszins het zicht op de reële ontwikkeling van vaardigheden, omdat niet duidelijk is hoe de vaardigheden zich zouden hebben ontwikkeld zonder deze interventies. In 2012 werd de zak-/slaagregeling aangepast en moesten leerlingen gemiddeld minimaal een 5,5 voor het CE halen. Deze eis werd in 2013 verder aangescherpt (kernvakkenregeling); er mocht voor havo/vwo voor de vakken Nederlands, Engels en wiskunde nog maximaal één 5 worden gehaald (voor SE en CE gezamenlijk). Vanaf 2014 moest in het vmbo minimaal een 5 voor Nederlands behaald worden om te kunnen slagen (SE plus CE). Deze wijzigingen hebben tot meer aandacht van scholen geleid voor het onderwijzen van de specifieke vaardigheden die nodig zijn voor een leerling om aan de exameneisen te kunnen voldoen (waaronder 'teaching to the test'). Daarbij spelen examentrainingen die binnen en buiten de scholen steeds vaker worden aangeboden eventueel ook een rol. Scholen hebben zich meer op het CE gericht, waarbij het SE meer op het CE is gaan lijken.⁷ Het belang voor scholen en leerlingen van goede CE-uitslagen is duidelijk toegenomen, terwijl het belang voor en de motivatie van leerlingen voor de PISA-toets laag wordt ingeschat.

De vraag is of het beeld ook beïnvloed wordt door wijzigingen in de diverse leerlingenpopulaties die aan het examen hebben deelgenomen. Deze vraag kan niet eenduidig worden beantwoord vanuit de cijfers van het Cito-rapport omdat hier diverse tegengestelde effecten spelen. Wel valt het op dat in de jaren 2010-2012 scholen lijken voor te sorteren op de verscherpte exameneisen (met meer zittenblijvers en afstromers); daarna vallen deze effecten weg, mogelijk doordat men beter weet waar men aan toe is in de aangepaste situatie. Ook valt het op dat sinds 2015 het aantal leerlingen dat niet is aangemeld voor het centraal examen licht stijgt (met uitzondering van vmbo-bb waar het tot 2016 duidelijk daalt en daarna in 2017 weer toeneemt).

Voor uitgebreide informatie: paragraaf 2.2.1, hoofdstuk 3 tot paragraaf 3.5, hoofdstuk 5 en paragraaf 6.2 van het Cito-rapport.

De commissie komt in deze conclusie tot beantwoording van de hoofdvragen.

- I. Kunnen de kwaliteit van de examens en de ontwikkeling van de vaardigheden van scholieren optimaal worden vastgesteld en gevolgd?

De commissie is ervan overtuigd dat voor wat betreft de normering het CvTE en het Cito de juiste middelen in handen hebben om de kwaliteit van het CE over de jaren heen te waarborgen en te monitoren, in ieder geval voor de vakken waarbij de meer geavanceerde normhandhavingsmethoden mogelijk zijn. Deze meer geavanceerde methoden worden toegepast bij de meeste examens voor de grootste groepen leerlingen. Indien de wens leeft om de examens van meer vakken op het hoogst mogelijke niveau te normeren met de meer geavanceerde methoden, dan is daarvoor meer budget nodig. Dit vraagt om een politieke beoordeling en afweging.

Bij bovenstaande dient aangetekend te worden dat de complexiteit van de normeringsprocedure maakt dat deze snel vragen kan oproepen. Hoewel dit niet helemaal te vermijden is, is de commissie wel van mening dat een grotere transparantie over de beslissingen in de N-termprocedure (o.a. de vaststelling, de mogelijke bijstellingen en de impact daarvan) bij kan dragen aan het vertrouwen daarin. In het bijgaande Cito-rapport is daar in samenwerking met het CvTE een begin mee gemaakt.

Wat betreft het systematisch monitoren van de ontwikkeling van de vaardigheden van VO-leerlingen is het Nederlandse CE een belangrijk instrument. Het is een toets van goede kwaliteit. Hoewel PISA ook een toets van goede kwaliteit is, heeft de commissie twijfel over de validiteit van PISA met betrekking tot het meten van vaardigheden. Het is een 'low stakes' toets en niet uitgesloten moet worden dat PISA hierdoor geen goed beeld geeft van vaardigheden van leerlingen. Ook moet worden benadrukt dat het CE en PISA verschillende vaardigheden van leerlingen in kaart brengen. Het feit dat de ontwikkeling van de resultaten van deze twee soorten toetsen van elkaar kan afwijken, zegt daarom niets over de vaardigheidsontwikkeling van de leerlingen in het algemeen. Uiteraard kan een daling op een van beide toetsen op zich gezien reden zijn om nader onderzoek te doen naar de kwaliteit van het aangeboden onderwijs dat de betreffende vaardigheden moet aanleren. Daarbij is het wel van belang rekening te houden met de motivatie van de leerlingen voor de toets en de leeftijd waarop de toets wordt afgenomen.

Bron: Brief getiteld "Aanbieding bevindingen ontwikkeling examenresultaten VO n.a.v. motie Van Meenen c.s." d.d. 28 maart 2019 aan de toenmalig Minister van OC&W Slob; Analyse door auteurs

2: De Commissie Steur samengesteld door het NRO concludeerde dat de vergelijking tussen PISA en het CE mank gaat en er geen sprake is van teruglopende onderwijsresultaten (3/3)

- II. Wat valt er – met de nuance vanuit het antwoord op hoofdvraag I – te zeggen over de ontwikkeling van de vaardigheden in de afgelopen 10 jaar?

De ontwikkelingen rondom de examens in het voortgezet onderwijs en de afnames van PISA in de afgelopen tien jaar laten zien dat de resultaten op het CE stabiel zijn gebleven en op een aantal onderdelen zijn verbeterd. Het SE laat een stabiel beeld zien met een paar afwijkingen, meestal naar boven. Tegelijkertijd zijn de PISA-scores gelijk of (duidelijk) afgenomen, ook in vergelijking met de overige OESO-landen.

Onze conclusie is dan ook dat de vaardigheden van leerlingen qua niveau, zoals getoetst door het CE, in de afgelopen tien jaar gelijk gebleven zijn of zijn toegenomen. Dit wijst overigens niet op een toename van de vaardigheden over de volle breedte, want dan zouden ook de resultaten op het SE zijn toegenomen (alhoewel de validiteit van het SE niet van eenzelfde niveau is als het CE). Een vergelijking met de (deels gelijkblijvende en deels dalende) PISA-scores is niet mogelijk omdat het hier om een 'low stakes' toets gaat, deels afgenomen bij een andere leeftijdsgroep en in een andere fase van hun opleiding, en die een andere vraagstelling als basis heeft.

Wij als commissie komen tot de volgende verklaring voor de betere leerlingresultaten op het CE. Wij baseren ons daarbij op het Cito-onderzoek en de gevoerde gesprekken en panels. Nader onderzoek kan dit beeld aanvullen of nuanceren. Overkoepelende verklaring is het toegenomen belang van het CE.

- De lat voor het slagen is door het beleid van de overheid (aanpassing zak-/slaagregeling) hoger komen te liggen (met sterk effect op de resultaten van het CE).
- De experts die wij spraken, gaven aan dat binnen de vakverenigingen van leraren de neiging bestaat om het belang van een vak af te meten aan de status als examenvak met een daarbij passende examensyllabus die – bij herziening – vaker zwaarder dan lichter wordt gemaakt.
- Het SE wordt te weinig ingezet om de eigenheid van het onderwijs van de school te honoreren maar functioneert steeds meer als een voorbereiding op het CE.⁸ Ook het bewaken van een maximale marge tussen het gemiddelde SE-cijfer en CE-cijfer per opleiding heeft hier vermoedelijk aan bijgedragen.⁹
- De focus op het examen en de examencijfers in het dagelijkse onderwijs in de bovenbouw van het VO is toegenomen, onder andere blijvend uit de toename van examentrainingen in schoolverband of op particulier initiatief ('teaching to the test').
- De examenresultaten zijn in de afgelopen jaren door de openbaarheid van schoolprestaties én het toezicht van de Inspectie van het Onderwijs (risicogericht vanaf 2007 met nadruk op de leeropbrengsten) op schoolniveau zwaar gaan wegen.
- Ouders en leerlingen hechten oververminderd belang aan het behalen van een diploma ten behoeve van de vervolgoopleidingen van de eindexamenkandidaten. Dit wordt versterkt door het belang van goede cijfers bij loting en intake voor een deel van de opleidingen in het hoger onderwijs.

Tenslotte

Al met al zijn de resultaten behaald op het CE in de afgelopen tien jaar stabiel gebleven of verbeterd. De verbetering wijst op een toegenomen vaardigheidsniveau van de leerlingen; de leerlingen bezitten op een aantal gebieden meer kennis en vaardigheden binnen de gedefinieerde examendomeinen.

Tijdens dit onderzoek en de gesprekken met betrokken partijen heeft de commissie de stellige indruk gekregen dat de sterke focus op het examen, en de start van het SE in de voorexamenklassen, naast betere resultaten tegelijk een versmalling van het onderwijs heeft opgeleverd met een te grote nadruk op het behalen van goede toetsresultaten ('teaching to the test'). Ook de Onderwijsraad heeft recent op dit risico gewezen in haar advies "Toets wijzer" waarin ze spreekt van een grote nadruk op beslissende, centraal ingerichte en vooral kwantitatieve toetsing.

Het beeld waarbij het CE het onderwijs in de bovenbouw stevig in zijn greep heeft met het risico van verschrining werd door veel gesprekspartners bevestigd. Het belang van het CE als objectief ijkpunt voor het behaalde opleidingsniveau door leerlingen is niettemin evident. Onze vraag is echter of het hier een 'natuurlijke' afsluiting van een opleiding betreft of dat het examen tegenwoordig zijn schaduw zozeer vooruit werpt, dat het examen het bovenbouwonderwijs in het VO in hoge mate is gaan bepalen. Leraren, vakgroepen en schoolleiders lijken weinig ruimte te zien om het onderwijs in te richten naar hun eigen visie om zo bijvoorbeeld meer aandacht te besteden aan bredere vorming of vakoverstijgend werken. Hiervoor zou een andere invulling van het SE goede mogelijkheden bieden in lijn met hoe het SE bedoeld is. Ook zou er voor de scholen ruimte moeten zijn om meer aandacht te besteden aan de vaardigheidsontwikkeling van leerlingen op het terrein van meer alledaagse vraagstukken, zoals door PISA getoetst.

Wij willen u van harte aanbevelen om door middel van nader onderzoek in het VO na te gaan hoe groot het gewicht en de uitstraling is van het CE voor het dagelijkse onderwijs (inclusief de invulling van het SE) en - als dit gewicht erg groot is, zoals wij veronderstellen – te overwegen of dit consequenties zou moeten hebben voor de aard, de invulling en de omvang van het SE en het CE in het voortgezet onderwijs.

Bron: Brief getiteld "Aanbieding bevindingen ontwikkeling examenresultaten VO n.a.v. motie Van Meenen c.s." d.d. 28 maart 2019 aan de toenmalig Minister van OC&W Slob; Analyse door auteurs

2: Een verband tussen de PISA en de Centrale Examens (CE) scores zou verwacht mogen worden voor vergelijkbare vakken

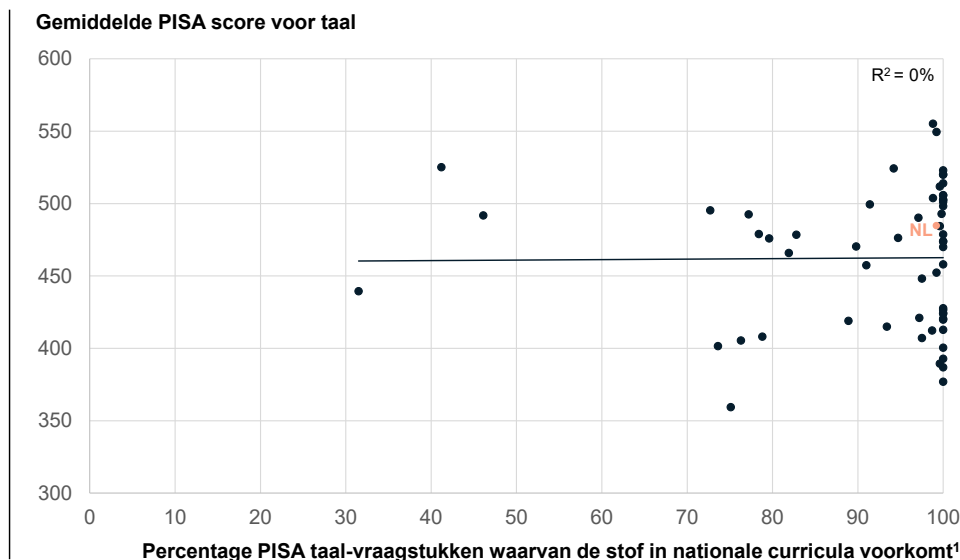
Tegenargumenten	Valide?	Rationaal
a PISA en de Centrale Examens zijn niet geheel vergelijkbaar	✗	- Gesteld wordt dat PISA de toepassing van kennis en vaardigheden in de dagelijkse praktijk test, terwijl de CE's duidelijke vaktechnische onderwerpen testen. Dat zou geen grote rol moeten spelen, aangezien PISA en de CE's vergelijkbare vakken bevatten: Nederlands leesvaardigheid, Wiskunde en Natuurwetenschappen (o.a. Natuurkunde, Scheikunde, Biologie en Aardrijkskunde). Als iemand goed is in bijvoorbeeld Wiskunde, dan verwacht je dat deze leerling zowel op PISA als op het CE daar goed in is. Het type vragen is minder van belang. De CE's voor Wiskunde en Natuurkunde bevatten steeds meer toepassingen in de praktijk. Als dit wel zo is, zijn de CE's dan wel een goede toets? - De vorm van de door PISA gestelde vragen (bijvoorbeeld contextrijk) sluit goed aan bij het onderwijs in Nederland - Taal: het maakt niet uit of de stof in het nationale curriculum zit. NL experts vinden 85% van de vragen zeer relevant (Schema 24) - Wiskunde: er is geen fundamenteel verschil wanneer PISA vragen meer vergelijkbaar zijn met het standaard nationale curriculum (Schema 25)
b CE's vinden later plaats dan PISA	✗	- De CE's (16-19 jaar) vinden later plaats dan PISA (15 jaar). Dit verklaart echter niet de PISA daling. - Als PISA daalt en de CE scores blijven gelijk of stijgen zelfs, dan zouden leerlingen een steeds grotere versnelde inhaalslag van 1 tot 4 jaar na PISA maken. Dit is onwaarschijnlijk. Er zijn geen aanwijzingen dat we in de bovenbouw steeds effectiever onderwijs zijn gaan geven - Internationale testen laten ook zien dat de resultaten van 15-jarigen zeer goed voorspellend zijn voor het resultaat van 16-19-jarigen (Schema 26)
c PISA heeft geen consequenties voor de leerlingen "low stakes"	✗	- CE's hebben inderdaad meer consequenties, maar dit is een bewuste keuze omdat PISA wil meten wat je daadwerkelijk in het leven kan - Dit gegeven is een constante factor over de jaren heen die dus niet eventuele dalingen of stijgingen verklaren - PISA test voornamelijk algemene vaardigheden i.p.v. kennis, dus speelt toets voorbereiding en motivatie voor PISA een veel minder grote rol in de manier van toetsen (Schema 27). Een groei-mindset heeft wel een positief effect op de behaalde PISA score (Schema 29)
d NL leerlingen minder gemotiveerd voor PISA dan leerlingen van andere landen	✗	- Aangezien we niet weten of er een significante verandering hierin is geweest in de afgelopen 15-20 jaar is dit geen argument. Als de motivatie inderdaad gedaald zou zijn, dan is het zelfs aannemelijk dat dit zich niet beperkt heeft tot alleen PISA, maar meer voor onderwijs in zijn algemeenheid. Dit zou dan eerder een bevestiging zijn van een algehele niveaudaling. NL scoort 7,4/10 op motivatie versus 7,6 gem in OECD - In NL wordt slechts 1% van de toets items niet gemaakt (daarmee staan we in de top 10 beste landen) en daarnaast wordt over 97% van de PISA-items meer dan 5 sec nagedacht (top 20 beste landen). Beide datapunten wijzen erop dat NL leerlingen PISA serieus nemen (Schema 27) - De "response time" in Nederland is relatief laag, maar "response time" is geen goede voorspeller van de PISA score (Schema 28)
e De motivatie van NL leerlingen is toegenomen voor het CE en dus betere CE kandidaten	✗	- Als reden wordt gegeven dat de exameneisen zijn verzaagd en het CE een groter gewicht heeft. Hiervoor wordt geen feitenbasis aangereikt voor oorzaak en gevolg, de causaliteit lijkt eerder een aanname. Het CE telde in 1990-95 voor 50% mee en moest je tekortpunten compenseren - De moeilijkheidsgraad van het CE zelf wordt niet als mogelijke factor meegenomen. Als de lat voor een examen (haalbaar) hoger ligt, dan gaan leerlingen zich beter voorbereiden (Hattie) en het is onbekend of dit daadwerkelijk is gebeurd. - Er wordt genoemd dat er meer afstroom is waardoor er nu betere examenkandidaten zijn. Dit kan ook door niveaudaling van de leerlingen gekomen zijn. Bij gelijke lat (moeilijkheid toetsen) in de voorexamenklassen, houd je nog steeds even goede examenkandidaten over
f PISA wordt sinds 2015 digitaal afgenomen	✗	- Dit verklaart niet de daling volgens PISA voorafgaand aan 2015 - Dit verklaart ook niet waarom andere reeds goed presterende landen als Singapore de stijgende lijn wel hebben doorgezet na deze overgang

Bron: Vaardigheidsontwikkelingen volgens PISA en examens (2019) door Cito auteurs Paul van der Molen, Sanneke Schouwstra, Remco Feskens en Marieke van Onna; PISA rapportages en data rondom methodologische verantwoording, een interview met Andreas Schleicher (PISA Director for Education and Skills Special Advisor on Education Policy to the Secretary-General); Analyse door auteurs

2a: Taal – de hoeveelheid vraagstukken waarvan de stof in het nationaal curriculum zit heeft geen effect op de PISA score

PISA 2018; elk datapunt in de grafiek is een deelnemend PISA land

- Voor Taal heeft PISA een zeer hoge validiteit: nationale experts vinden slechts 3% van de vragen niet relevant (voor NL is dit 0%)
- Ook vinden zij een hoge mate van vergelijkbaarheid met nationale curricula: slechts 9% komt niet in curricula voor (voor NL is dit 1%)
- Zelfde correlatie bij alleen standaard in nationale curricula



1. Het vraagstuk komt of standaard voor in alle curricula of in sommige curricula van het land

2a: Wiskunde - geen fundamenteel verschil voor landen wanneer PISA 2009 meer vergelijkbaar gemaakt zou worden met het standaard nationale curriculum

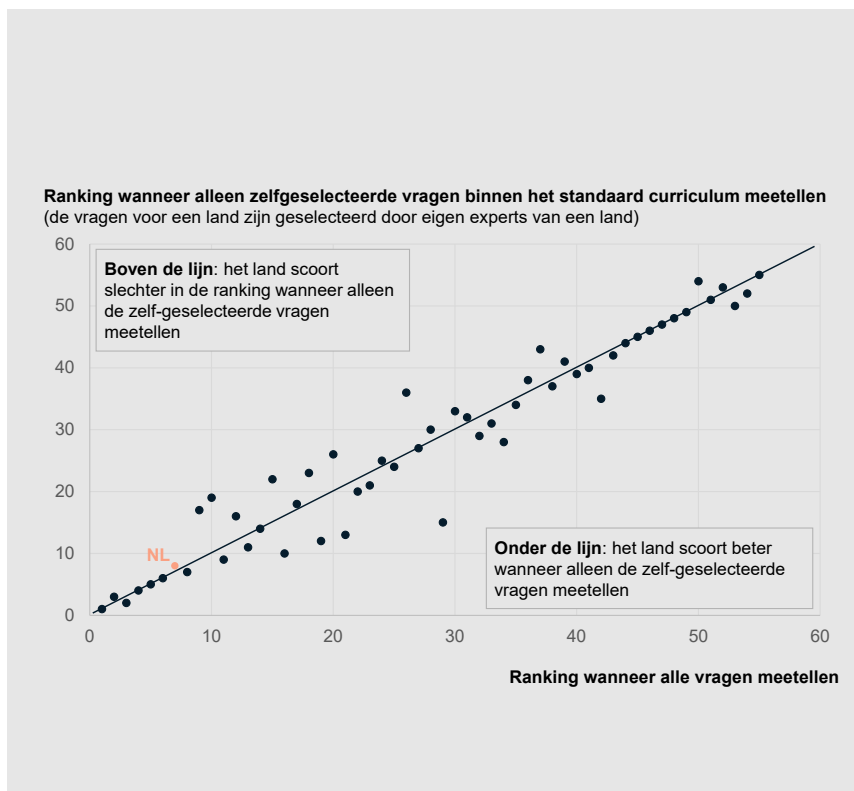
In 2009 werd het PISA domein Wiskunde vernieuwd met nieuwe vragen.

Voor deze vragen is gekeken of het voor de ranking van een land (een bolletje, NL is het oranje bolletje) uitmaakt of alle vragen meetellen (horizontaal) of alleen de door het land zelf-geselecteerde vragen (verticaal).

Elk van beide rankings is bepaald op basis van het percentage correct beantwoorde PISA vragen.

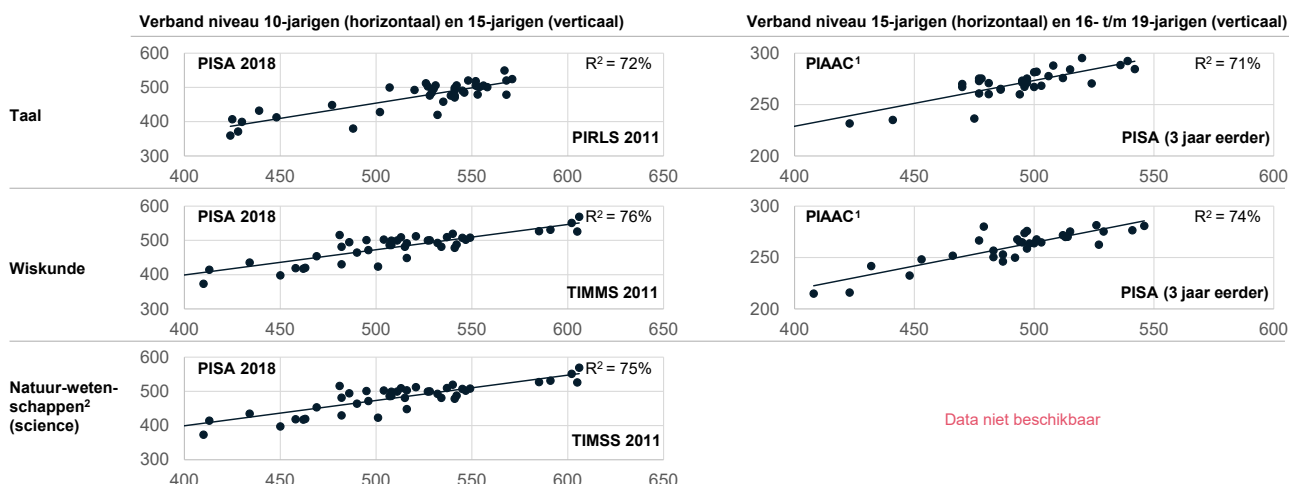
Aangezien het overgrote merendeel van de landen op of dichtbij rondom de getekende 45 graden lijn liggen, maakt het niet of nauwelijks voor een land uit of alle vragen PISA vragen meetellen of alleen de zelf-geselecteerde vragen.

Bron: PISA rapportages 2009; Analyse door auteurs



2b: Een voorsprong op 10-jarige leeftijd leidt tot een voorsprong op 15-jarige en op 16- tot en met 19-jarige leeftijd

Prestaties van "landencohorten" middels score op PIRLS/TIMMS (10-jarigen), PISA (15-jarigen) en PIAAC (16- t/m 19-jarigen) op verschillende momenten in de tijd



Noot: Door de afnememomenten van PIRLS en TIMMS enerzijds en PISA anderzijds is het verschil niet precies 5 jaar

1. Landen hebben op verschillende momenten deelgenomen aan PIAAC (Ronde 1: 2011-2012, ronde 2: 2014-2015 en ronde 3 2017-2018).

De PIAAC scores van een land zijn gematcht met de PISA scores van dat land 3 jaar eerder

2. Natuurkunde, Scheikunde, Biologie en Aardrijkskunde

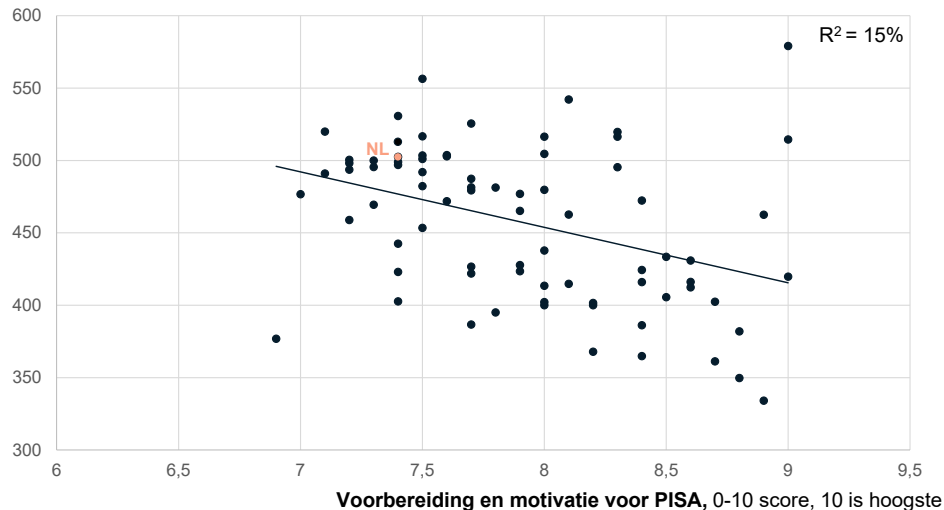
Bron: PISA rapportages 2009, 2012, 2015, 2018; PIRLS en TIMSS 2011; PIAAC 2011-2012, 2014-2015, 2017-2018. Voor deze analyse hebben de auteurs de scores van enkele jaren later van dezelfde landen gecorreleerd met de scores uit eerdere jaren. De deelnemende scholen / leerlingen zijn niet identiek; Analyse door auteurs

2c en 2d: Meer voorbereiding en motivatie voor de PISA-toets leiden niet tot een hogere behaalde PISA-score

Deelnemende PISA landen (elk datapunt in de grafiek is een deelnemend PISA land); PISA 2018

- NL scoort 7,4/10 op motivatie versus 7,6 OECD gemiddelde
- Twee datapunten wijzen erop dat NL-leerlingen de PISA toets serieus nemen
- In NL wordt slechts 1% van de toets items niet gemaakt (daarmee staan we in de top 10 beste landen)
- Daarnaast wordt over 97% van de items wordt meer dan 5 seconde nagedacht (top 20 beste landen)

Gemiddelde PISA score, Taal, Wiskunde, Natuurwetenschappen



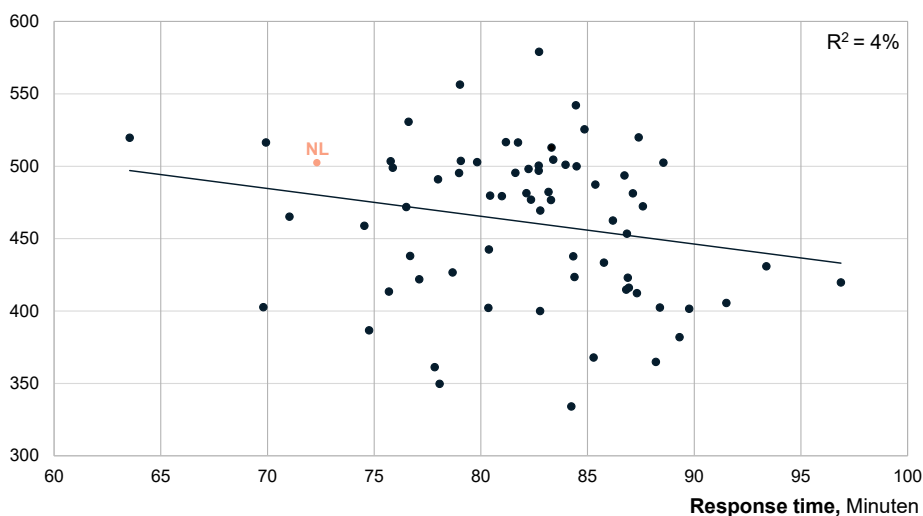
Bron: PISA rapportages 2018; Analyse door auteurs

2c en 2d: Een lagere response time (minder tijd besteed aan PISA vragen) leidt niet tot een lagere PISA score

Deelnemende PISA landen (elk datapunt in de grafiek is een deelnemend PISA land); PISA 2018

- PISA heeft in 2018 de tijd gemeten die studenten aan de PISA test besteden. Dit wordt de "response time" genoemd
- De totale hoeveelheid beschikbare tijd voor PISA is 120 minuten (2 uur)
- NL heeft in 2018 een relatief lage response time van 72 minuten
- Echter uit de grafiek blijkt dat er geen aanwijzing is dat een lagere response time tot een lagere PISA score leidt

Gemiddelde PISA score, Taal, Wiskunde, Natuurwetenschappen



Bron: PISA rapportages 2018; Analyse door auteurs

2c en 2d: Een groei-mindset leidt tot een hogere PISA-score

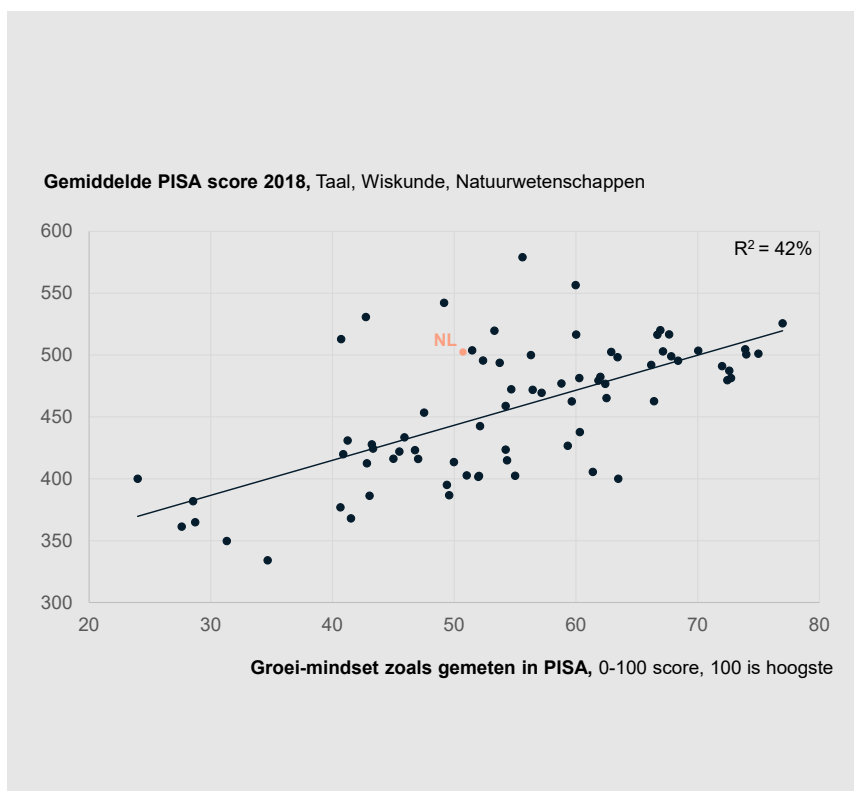
Groei-mindset is het aanhangen van de incrementele theorie van intelligentie, ofwel dat iemands vaardigheden en intelligentie over tijd kunnen ontwikkelen en toenemen door inzet

PISA heeft alle deelnemende leerlingen de vraag gesteld of zij deze theorie geloven. Vervolgens heeft PISA vastgesteld welk percentage leerlingen per deelnemend land dit gelooft.

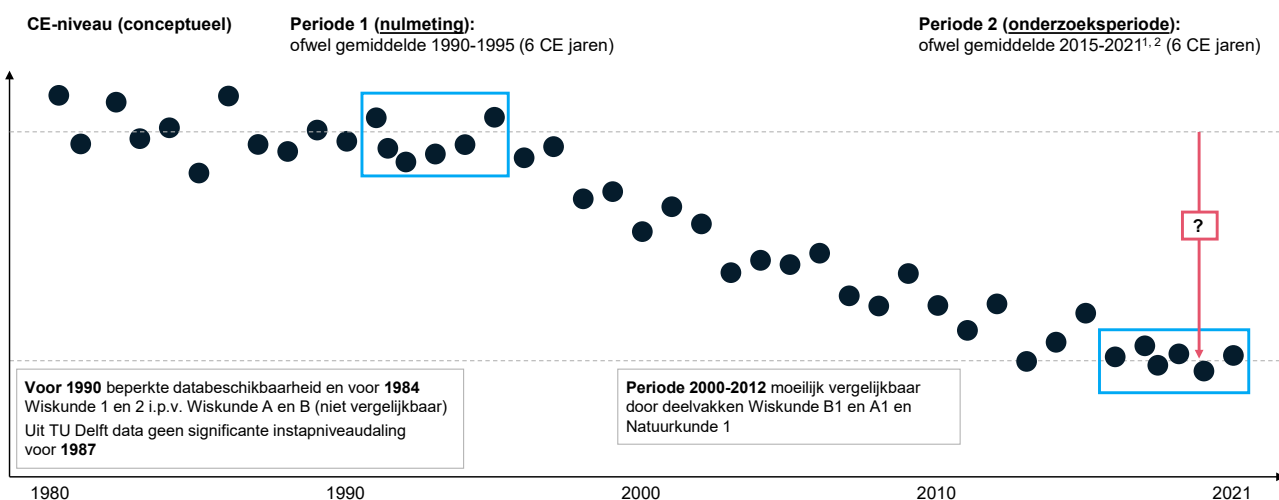
Elk datapunt in de grafiek rechts representeert een deelnemend PISA land.

Uit de grafiek blijkt des te groter het aandeel leerlingen dat in een groei-mindset gelooft, des te hoger de PISA score

Bron: PISA rapportages 2018, zie ook Carol Dweck; Analyse door auteurs



3: “Artist impression” van de onderzoeksopzet – Twee vergelijkbare perioden met goede databeschikbaarheid zijn vergeleken om de hypothese van CE-niveauperandering over een lange periode te toetsen

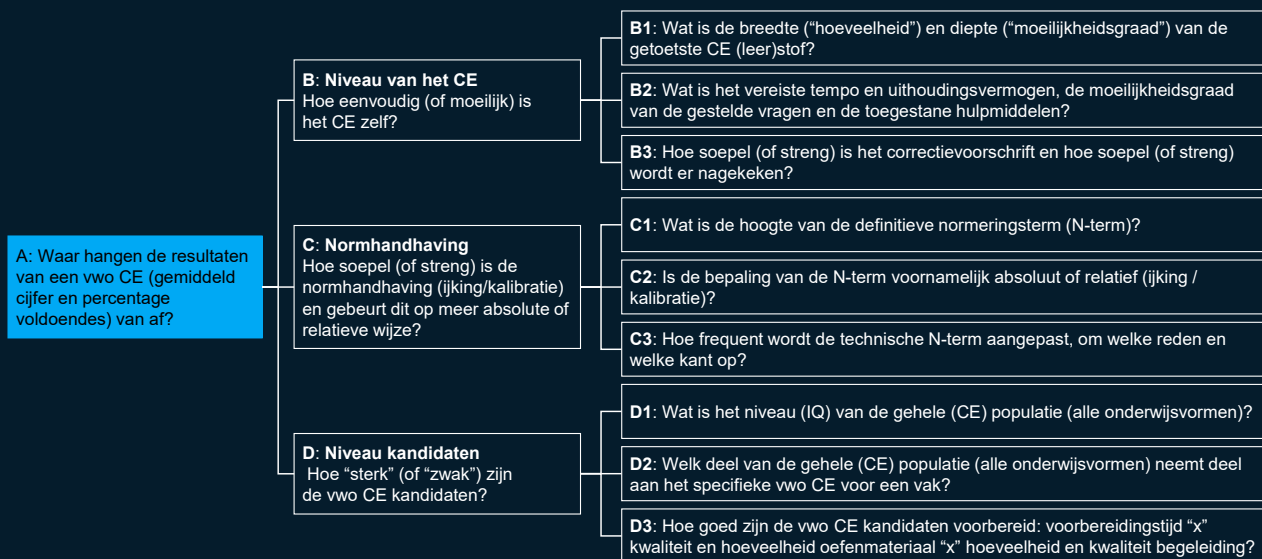


1. Inclusie of exclusie van coronajaar 2021 maakt geen significant verschil in de analyses in dit rapport; In 2020 was er geen CE

2. Voor een van de analyses met betrekking tot de stofverandering ligt de nadruk op 2018-2021 voor Wiskunde en 2016-2021 voor Natuurkunde i.v.m. de curriculumwijziging die vlak hiervoor plaatsvond

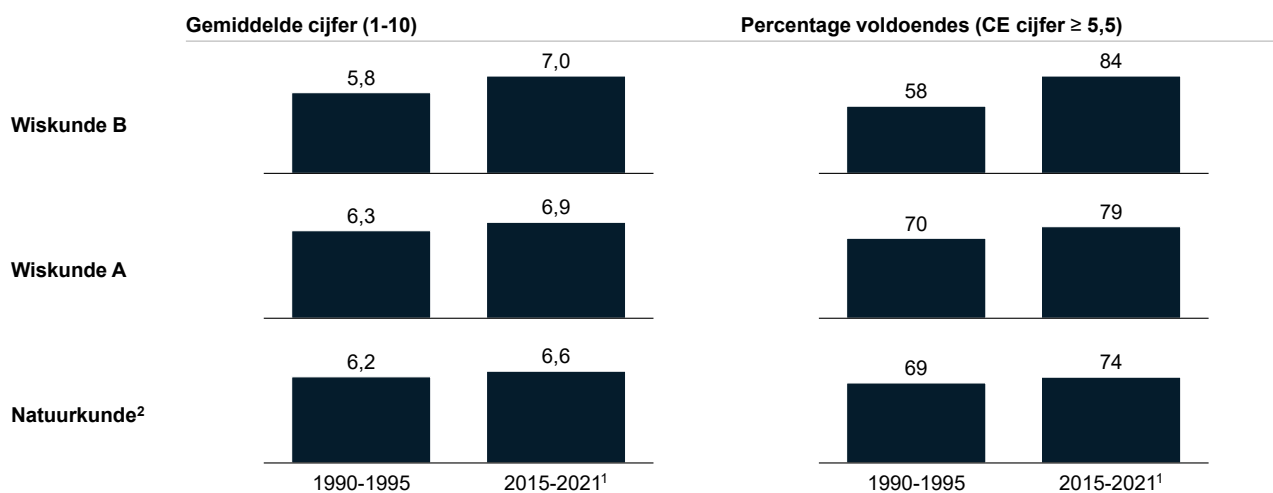
Bron: Analyse door auteurs

3: Van welke factoren hangen de resultaten van een vwo CE Wiskunde A of B of Natuurkunde af?



Bron: Analyse door auteurs

A: Significant betere CE resultaten nu dan 30 jaar geleden



Komen de significant betere resultaten voor het CE doordat het CE makkelijker is, er soepeler wordt genormeerd en / of doordat de examenkandidaten sterker zijn?

1. Exclusief 2020 (geen CE door Corona)

2. Voor de vergelijking met PISA. Naast voor Natuurkunde zijn ook de CE cijfers voor de andere Natuurwetenschappenvakken zijn afgelopen 20-30 jaar gestegen. Op basis van gevonden publiekelijk beschikbare data: Scheikunde van 6,4 in 1990-1994 naar 6,6 in 2015-2021, Biologie van 6,1 in 1990-1994 naar 6,3 in 2015-2021 en Aardrijkskunde van 6,2 in 2000-2005 naar 6,3 in 2015-2021 (voor Aardrijkskunde geen publieke data gevonden voor de periode 1990-1995)

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/95 en exa menverslagen 2015-2021; Cito TIAs; Analyse door auteurs

Schema 33

B1: Breedte - de CE-(leer)stof is netto 40-50% afgenomen voor Wiskunde A en B en voor Natuurkunde

- Onderzoeksmethodiek: 1) gestart met huidige examenprogramma's (2021) zoals gepubliceerd door het CvTE, 2) deze programma's vergeleken met de examenstof in de periode 1990-1995 middels vwo Samengevat en examenprogramma's, 3) op basis daarvan een hypothese geformuleerd waar de stof overeenkomt en verschilt, 4) de verschillen en overeenkomsten geverifieerd door de vragen uit de centrale examens van 1990-1995 en 2018-2021 voor Wiskunde en 2016-2021 voor Natuurkunde te bestuderen
- In totaliteit zijn ongeveer 1.500 centrale examenvragen geanalyseerd.
- Het totale vwo curriculum 30 jaar geleden bestond uit 7 examenvakken waarbij in de bovenbouw o.a. ook Frans, Duits, Maatschappijleer en Lichamelijke Opvoeding werden gegeven. Het totale curriculum voor Wiskunde A en B en Natuurkunde (CE en SE tezamen) bevat ook minder stof dan 30 jaar geleden. Het vermoeden bestaat bij de auteurs dat het totale curriculum voor andere vakken ook is afgenomen. Bijvoorbeeld minder verplichte boeken bij de talen. Het vwo programma bevat nu bijvoorbeeld wel een profielwerkstuk en CKV.
- Een groot deel van stof die uit het examenprogramma van Wiskunde A en B is gehaald, zit nu in het programma van Wiskunde D, maar dit is voor geen enkel profiel verplicht (wordt slechts door een paar duizend leerlingen per jaar gekozen) en is geen centraal examenvak met een onafhankelijke landelijke niveaucontrole waarin ook wordt getoetst in welke mate de stof kan worden gecombineerd met andere onderdelen van de CE stof.

Noot: Minder CE stof leidt ook tot minder vraagcombinatiemogelijkheden

1. Een deel van huidige examenstof die niet in het CE stof wordt getoetst, zit wel nog in het School Examen (SE)

Bron: Centrale Examens 1990-95 en 2015-21 tijdvak 1 en 2; Analyse door auteurs

Wiskunde B ~40-50%
netto minder CE stof
dan 30 jaar geleden¹

Wiskunde A ~45-55%
netto minder CE stof
dan 30 jaar geleden¹

Natuurkunde ~40-50%
netto minder CE stof
dan 30 jaar geleden¹

Schema 34

B1: Wiskunde B – 40-50% netto minder CE stof nu dan 30 jaar geleden

Vergelijking van de vwo CE Wiskunde B examenstof in 1990-1995 en 2018-2021; alleen weergegeven waar de CE-stof verschilt

■ Uit CE verdwenen stof ● Zeer veel / moeilijk ● Zeer weinig / makkelijk

CE domeinen	CE sub-domeinen en betreffende (delen van de) examenstof uit dat sub-domein	CE leerstof		In CE?	
		Hoeveel?	Niveau?	90-95	18-21
B: Functies, grafieken en vergelijkingen	• B2: Parate kennis hebben van en het productief kunnen toepassen van cyclometrische functies inclusief hun afgeleiden: arc sin, arc cos en arc tan	●	●	✓	✗
	• B3: Vermenigvuldigen van een willekeurige functie t.o.v. een willekeurig punt en spiegelen / aantonen van symmetrie van een willekeurige lijn	●	●	✓	✗
	• B6: Het kennen en kunnen toepassen van de standaardlimieten voor $x^a \cdot \ln(x)$, $x^a \cdot e^x$, $\sin x / x$ voor limieten naar 0 en plus of min oneindig	●	●	✓	✗
C: Differentiaal-rekening	• C1: Het kunnen analyseren en oplossen van differentiaalvergelijkingen in de vorm $dy / dx = f(x,y)$ inclusief bijvoorbeeld lijnelementenvelden en singuliere punten	●	●	✓	✗
	• C1: Het onderzoeken of een functie differentieerbaar is in een punt met behulp van de regel van De l'Hôpital	●	●	✓	✗
	• C3: Het productief kunnen toepassen van geavanceerde integratietechnieken als partieel integreren en de substitutieregel inclusief het hebben van parate kennis en vaardigheid om functies als $f(x) \cdot e^{ax+b}$ (met $f(x)$ een polynoom van orde 1 of groter), $1/(1+x^2)$ en $\pm 1/\sqrt{1-x^2}$ te integreren	●	●	✓	✗
D: Goniometrie	• D: Parate kennis hebben van dubbele hoek formules voor $\sin(2x)$ en $\cos(2x)$ (formules worden gegeven)	●	●	✓	✗
	• D: Toepassen van de som- en verschilformules voor $\sin(t \pm v)$ en $\cos(t \pm v)$, maar niet als parate kennis	●	●	✗	✓
E: Meetkunde	• E1: De gehele CE leerstof voor 3-dimensionale meetkunde inclusief kunnen bewijzen. Dit was bijvoorbeeld een van de drie boeken vwo bovenbouw boeken van de methode Getal en Ruimte en standaard een van de vier CE opgaven	●	●	✓	✗
	• E1: Het hebben als parate kennis en kunnen toepassen van 3 stellingen in de vlakke meetkunde ¹ en een zwaartepunt met behulp van vectoren bepalen	●	●	✗	✓
	• E2: Willekeurige krommen kunnen analyseren in de vorm $K(x,y) = c$ incl (a) differentiëren en maxima en minima bepalen, (b) oppervlakte en volume-integralen, (c) lijn en puntsymmetrie, en (d) asymptoten. Onderdeel (b), (c) en (d) zijn ook verdwenen voor parametervoorstellingen in de vorm $x(t)$ en $y(t)$	●	●	✓	✗

Noot: Een klein deel van de verdwenen CE-stof kan nog voorkomen in het SE middels de keuzeonderwerpen (40 van de 600 SLU ofwel 6-7% van huidige examenprogramma). Scholen zijn in de keuze echter volledig vrij en een snelle controle van PTA's van een aantal scholen liet zien dat de representatie van de verdwenen CE stof middels keuzeonderwerpen in het SE zeer beperkt is

1. (1) Van een rechthoekige driehoek is het midden van de schuine zijde het middelpunt van de omgeschreven cirkel, (2) Een driehoek waarvan een zijde de middellijn van de omgeschreven cirkel is, is rechthoekig (Stelling van Thales) en (3) Als vanuit een punt twee raaklijnen aan een cirkel getrokken worden, dan zijn de afstanden van dat punt tot de twee raakpunten gelijk

Bron: Examenprogramma's vwo 2021, vwo samengevat, vwo examens Wiskunde B 1990 -1995 en 2018-2021, Interviews met vakexperts; Analyse door auteurs

B1: Wiskunde B – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen

B3: Vermenigvuldigen t.o.v. een punt – uit CE 1992-I

Van $\mathbb{R}$ naar $\mathbb{R}$ is gegeven de functie $f: x \rightarrow \frac{1}{xe^x}$.

Ten opzichte van een rechthoekig assenstelsel Oxy is K de grafiek van f .

Op K ligt een punt P met x -coördinaat a .

Bij een vermenigvuldiging ten opzichte van O met factor -2 is Q het beeld van P .

Bereken a in het geval dat Q op K ligt.

B6: Standaardlimieten – uit CE 1995-I

De functie f met domein $\mathbb{R} \setminus \{0\}$ is gegeven door:

$$f: x \rightarrow \frac{1}{x^2} e^{\frac{1}{x}} \quad \text{7p} \quad \square \quad \text{Onderzoek} \quad \lim_{x \rightarrow 0} f(x) \quad , \quad \lim_{x \rightarrow 0} f'(x) \quad \text{en} \quad \lim_{x \rightarrow 0} f''(x)$$

C1: Differentiaalvergelijkingen – uit CE 1991-I

Gegeven is de differentiaalvergelijking D :

$$4e^y (e^y - 1) \frac{dy}{dx} = 1 - x.$$

6 $\square$ Geef door arcering het gedeelte van het Oxy -vlak aan, waar de richtingscoëfficiënten van de door D bepaalde lijnelementen positief zijn.

7 $\square$ Toon aan dat de kromme K gegeven door $x = 1 + 2 \sin t$ en $y = \ln(1 + \cos t)$ een oplossingskromme van D is.

8 $\square$ L is de oplossingskromme van D die door het punt $(-2, \ln 3)$ gaat. Stel een vergelijking van L op.

C3: Geavanceerde integratietechnieken – uit CE 1995-I

De functie f met domein $\mathbb{R} \setminus \{0\}$ is gegeven door:

$$f: x \rightarrow \frac{1}{x^2} e^{\frac{1}{x}} \quad \text{6p} \quad \square \quad \text{Bereken voor welke } a < 0 \text{ geldt:} \quad \int_a^0 f(x) dx = f(a)$$

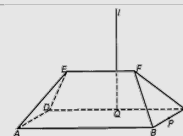
Bron: vwo Centrale Examens Wiskunde B 1990-1995; Analyse door auteurs

E1: Meetkunde – uit CE 1993-I

In figuur 3 en op de bijlage geldt voor het lichaam $ABCD.EF$:

- $ABCD$ is een rechthoek met $AB = 8$ en $AD = 4$.
- $EF \parallel AB$.
- $AE = DE = BF = CF = EF = 4$.
- P is het midden van BC .
- Q is het midden van DC .
- l is de lijn door Q loodrecht op $ABCD$.

figuur 3



7p 11 $\square$ Bereken de inhoud van het lichaam.

6p 12 $\square$ Bereken de afstand van de lijnen BF en DE .

Van een kegel K met l als as is BCF een raakvlak.

7p 13 $\square$ Bereken de halve tophoek van K in graden nauwkeurig.

Door P gaan twee raakvlakken aan K , waarvan BCF er één is.

7p 14 $\square$ Teken in een figuur van de bijlage de doorsnede van het andere raakvlak aan K met het gegeven lichaam. Licht je werkwijze toe.

E2: Willekeurige krommen – uit CE 1993-II

Ten opzichte van een assenstelsel Oxy is de kromme K gegeven door:

$$x = \ln|t+1| \quad \text{en} \quad y = \ln|t-1| \quad \text{waarbij } t \neq -1 \quad \text{en} \quad t \neq 1.$$

6p 7 $\square$ Bewijs dat K symmetrisch is ten opzichte van de lijn $y = x$.

E2: Willekeurige krommen – uit CE 1992-II

Ten opzichte van een rechthoekig assenstelsel Oxy is de kromme K gegeven door $x = 1 - \sin t$ en $y = t - \sin t$, waarbij $t \in [0, 2\pi]$.

6p 3 $\square$ Bewijs dat K symmetrisch is ten opzichte van het punt $(1, \pi)$.

B1: Wiskunde A – 45-55% netto minder CE stof nu dan 30 jaar geleden

Vergelijking van de vwo CE Wiskunde A examenstof in 1990-1995 en 2018-2021;

alleen weergegeven waar de CE-stof verschilt

■ Uit CE verdwenen stof ● Zeer veel / moeilijk ● Zeer weinig / makkelijk

CE domeinen	CE sub-domeinen en betreffende (delen van de) examenstof uit dat sub-domein	CE leerstof		In CE?	
		Hoeveel?	Niveau?	90-95	18-21
B: Algebra en tellen	• B3: Volledige Grafen en Matrices theorie inclusief overgangs-, kans- en populatie groeimatrices	●	●	✓	✗
	• C1: Goniometrie: cosinusfunctie en parate kennis en vaardigheid trendlijn (combinatie lineaire en goniometrische functie)	●	●	✓	✗
	• C2: Correlatie en regressie analyses kunnen uitvoeren (coëfficiënten uitrekenen)	●	●	✓	✗
D: Verandering	• D1: Rijen en Reeksen	●	●	✗	✓
	• D3: Afgeleide: afgeleide sinus en cosinus en tweede afgeleide	●	●	✓	✗
	• D4: (Lineair) optimaliseren / programmeren van een willekeurige doelfunctie met 2 of meer variabelen incl een domeinbepaling in 2-3 dimensies en isolijnen-analyse. Lineair optimaliseren met 2 variabelen met een domeinbepaling in 2 dimensies kan een keuzeonderwerp zijn.	●	●	✓	✗
E: Statistiek en kansrekening	• E (wel in SE): Statistiek en kansrekening (dit is 40% van de totale vwo Wiskunde A stof volgens het CvTE): — E1 Probleemstelling en onderzoeksontwerp — E2 Visualisatie van data — E3 Kwantificering — E4 Kansbegrip inclusief trekken met en zonder teruglegging — E5 Kansverdelingen inclusief binomiale en normale verdeling, som of verschil van 2 of meer onafhankelijke normale verdelingen, de normale verdeling van een gemiddelde — E6 Verklarende Statistiek inclusief het toetsen van hypothesen bij binomiale of normale verdelingen — E7 Statistiek met ICT	● ●	●	✓	✗
	• E (niet in SE): Geavanceerde analyses met binomiale en normale verdelingen: 1) onderzoeken of een verdeling normaal is, 2) een binomiale verdeling benaderen met een normale verdeling en 3) hypothesen toetsen bij binomiale of normale verdelingen inclusief rangrekeningtoetsen	●	●	✓	✗

Noot: Domein E Statistiek en kansrekening komt nog wel voor in het SE en een klein deel van de overige verdwenen CE-stof kan nog voorkomen in het SE middels de keuzeonderwerpen (40 van de 520 SLU ofwel 7-8% van huidig examenprogramma). Scholen zijn in de keuze echter volledig vrij en een snelle controle van PTA's van een aantal scholen liet zien dat de representatie van de verdwenen CE stof middels keuzeonderwerpen in het SE beperkt is

Bron: Examenprogramma's vwo 2021, vwo samengevat, vwo examens Wiskunde A 1990-1995 en 2018-2021, Interviews met vakexperts; Analyse door beide auteurs

Schema 37

B1: Wiskunde A – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (1/3)

B3: Grafen en Matrices theorie – uit CE 1994-I

In de overgangzone tussen het woestijnklimaat en het gematigde klimaat aan de westkust van Noord-Amerika treft men over een oppervlakte van ongeveer 2000 km² een vegetatie aan van groenblijvende struiken. Men spreekt daar over de Chaparral. De brandbaarheid van de planten is sterk afhankelijk van de leeftijd. Vanwege het vele dode materiaal zijn vooral de oudere planten zeer brandbaar. Brand heeft naast het gevaar voor mens en dier ook een belangrijke nuttige functie: op de plaats van de verbrande struiken komen vrijwel direct jonge en levenskrachtige planten uit de grond. Spontane branden worden daarom niet altijd gebelust. De verjonging zorgt er immers voor dat er geen grote, uitgestrekte gebieden ontstaan van dode materiaal die bij brand tot catastrofes zouden kunnen leiden.

Van deze situatie wordt een model gemaakt, waarbij men de volgende uitgangspunten hanteert.

De vegetatie wordt op grond van de leeftijd onderverdeeld in vier klassen:

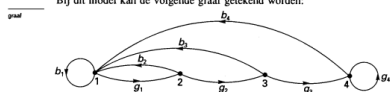
- klasse 1 0 tot 10 jaar
- klasse 2 10 tot 20 jaar
- klasse 3 20 tot 30 jaar
- klasse 4 30 jaar en ouder

Als maat voor de omvang van een klasse neemt men niet het aantal planten maar de oppervlakte van het door die klasse bedekte gebied.

Per klasse blijft het percentage dat elke 10 jaar verbrandt, constant.

De totale oppervlakte van het gebied blijft 2000 km².

Bij dit model kan de volgende graaf getekend worden:



met b_i = het gedeelte van klasse i dat verbrandt ($b_i < 1$) en
met g_i = het gedeelte van klasse i dat niet verbrandt ($g_i < 1$).
Bij deze graaf kan een overgangsmatrix M worden opgesteld, waarin b_i en g_i voorkomen.

49 11 □ Stel deze matrix M op in de vorm:

van
naar $\begin{pmatrix} & & & \\ & & & \\ & & & \\ & & & \end{pmatrix}$

Bron: vwo Centrale Examens Wiskunde A 1990-1995; Analyse door auteurs

In tabel 1 staat vermeld hoe groot de oppervlakte is die elke klasse bedekt op het tijdstip $t = 0$ en op het tijdstip $t = 1$ (10 jaar later).

klasse	Oppervlakten in km ²	
	op $t=0$	op $t=1$
1	302	462
2	284	300
3	314	278
4	1100	960

49 12 □ Bereken g_i , b_i en b_j in drie decimalen nauwkeurig.

Van de matrix M zijn met de computer de machten M^2 , M^3 , M^4 , ... uitgerekend. Men constateerde dat de matrices M^n vanaf een zekere waarde van n nauwelijks meer verschillen. Zo zijn, na afronding op twee decimalen, de matrices M^n voor $n > 20$ allemaal gelijk aan de volgende overgangsmatrix:

		van			
		1	2	3	4
naar	1	0,19	0,19	0,19	0,19
	2	0,19	0,19	0,19	0,19
	3	0,18	0,18	0,18	0,18
	4	0,44	0,44	0,44	0,44

Het blijkt dat op elke rij de getallen gelijk zijn.

49 13 □ Welke conclusies kan men uit dit alles trekken voor de samenstelling van de Chaparral-vegetatie?

In de praktijk passen de beheerders van de Chaparral ook nog gecontroleerde bewuste afbranding van gedeeltes van de vegetatie ouder dan 10 jaar toe.

In ons model nemen we ter vereenvoudiging aan dat dit *direct na* elke periode van 10 jaar in één moment plaats vindt.

Neem aan dat men steeds

2,5% van klasse 2,

1,3% van klasse 3 en

7,2% van klasse 4 verbrandt.

Dit proces van bewuste afbranding kan weergegeven worden door een vier-bij-vier matrix B , waarin de hierboven genoemde percentages zijn verwerkt.

Met behulp van het matrixproduct $B \cdot M$ kan dan het gezamenlijke proces over 10 jaar van de spontane afbranding, gevolgd door de gecontroleerde bewuste afbranding, beschreven worden.

49 14 □ Stel matrix B op.

Schema 38

B1: Wiskunde A – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (2/3)

C2: Correlatie en regressie – uit CE 1993-II

In een Amerikaans laboratorium heeft men proeven genomen waarbij gelet werd op het verband tussen de hoogte van de bewaartemperatuur (F) in graden Fahrenheit en de werkzaamheid (W) van een bepaald geneesmiddel. Bij temperaturen van 30°, 50°, 70° en 90° (Fahrenheit) werden drie porties van gelijk gewicht uit eenzelfde productie 20 dagen bewaard. Na deze periode werd op identieke wijze de werkzaamheid van de porties vastgesteld. De werkzaamheid werd uitgedrukt in percentages van de werkzaamheid zoals die was voor het bewaren.

In tabel 2 waar de twaalf meetresultaten zijn weergegeven, kan onder andere worden afgelezen dat de werkzaamheid van de drie porties die bij een temperatuur van 30° werden bewaard achtereenvolgens 39%, 42% en 35% van hun oorspronkelijke werkzaamheid hadden.

Bewaartemp. (F)	30°			50°			70°			90°		
	1	2	3	1	2	3	1	2	3	1	2	3
Werkzaamheid (W)	39	42	35	32	26	33	19	27	23	14	19	21

Indien men een rechtlijnig verband veronderstelt, levert dit voor de regressielijn van W op F de vergelijking: $W = -0,35F + 48,30$.

49 9 □ Verwerk de gegevens van tabel 2 in een spreidingsdiagram en teken daarin de regressielijn.

79 10 □ Bereken de correlatiecoëfficiënt tussen W en F .

Het verband tussen de temperatuur in graden Fahrenheit (F) en in graden Celsius (C) wordt vastgelegd door $F = 1,8C + 32$.

De regressielijn van W op C heeft de gedaante $W = aC + b$, waarbij de bewaartemperatuur in graden Celsius wordt gemeten.

39 11 □ Bereken a en b in twee decimalen nauwkeurig.

Uit andere experimenten is gebleken dat de werkzaamheid bij een vaste bewaartemperatuur exponentieel afhangt van de lengte van de bewaarijke.

59 12 □ Schat de gemiddelde werkzaamheid van porties die 40 dagen bij een temperatuur van 20° Celsius zijn bewaard.

Bron: vwo Centrale Examens Wiskunde A 1990-1995; Analyse door auteurs

D3: Afgeleide functie – uit CE 1992-II

De vorm van de grafiek laat duidelijk zien dat een model voor de gemiddelde getijdebeweging dat uitgaat van één enkele sinusoïde niet erg realistisch is. Beter is het om het stijgende deel AB en het dalende deel BC elk met een afzonderlijke sinusoïde te beschrijven. Omdat de tijdsduur 4 uur en 55 minuten ongeveer overeenkomt met 4,92 uur, geldt voor de waterhoogte (h) voor waarden van t tussen 0 en 4,92 bij benadering:

$$h = 100,5 \cdot \sin 0,64(t - 2,46) + 2,5$$

49 15 □ Bereken, uitgaande van dit model, de maximale stijgsnelheid van het water in centimeters per uur.

B1: Wiskunde A – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (3/3)

D4: (Linear) Programmeren – uit CE 1995-I

De fabrieksgedirecteerde *Reijer* moet verscheidene transportplannen maken voor het vervoer van de producten *A* en *B* naar de fabriek. De *A* is veldrijder 120 van zijn rijdt, in *B* is veldrijder 180 van zijn rijdt (1 van = 1000 kg).

Voor het vervoer van de *A* kan *Reijer* verscheidene streeklussen bepalen over veldrijden. Over de afstand en wegen worden geen bij het transport per veldrijder nog wat zijn veldrijden. De afstand van de rijdt die naar *B* veldrijden wordt, wordt de veldrijden.

De afstand van veldrijden van *B* naar de fabriek. Men kan verscheidene plannen per week veldrijden 240 van naar *A* en *B* veldrijden.

Laat met een berekening zien dat het niet mogelijk is zowel *A* als *B* veldrijden met veldrijden veldrijden van veldrijden rijdt te veldrijden.

Reijer krijgt de beschikking over een klein veldrijden. Bij het transport per veldrijden gaat geen tijd verloren. Daarnaast is het nu niet mogelijk beide veldrijden van veldrijden rijdt te veldrijden. Een veldrijden is echter dat de veldrijden per een veldrijden rijdt (de rijdt).

Transportkosten per een van *B* veldrijden rijdt

naar <i>A</i>	naar <i>B</i>
per veldrijden \$ 40	\$ 40
per veldrijden \$ 120	\$ 200

(\$ 40 = 40 dollar)

De afstand van het veldrijden naar de fabriek wordt de afstand van de rijdt die naar *B* veldrijden wordt, wordt de veldrijden.

Men kan met het transportplan van tabel 3. De veldrijden rijdt.

Transportkosten, aantal veldrijden van *A* en *B* (x, y, z, w, v, u)

naar <i>A</i>	naar <i>B</i>
per veldrijden x	y
per veldrijden z	w

Zowel tussen x en y als tussen y en z bestaat een direct verband. Met behulp hiervan kunnen de totale kosten in dollars per week (K) voor rijdt en veldrijden uitgedrukt worden in x en y.

13 a) Toon aan dat de totale kosten K = 104000 - 20x - 40y.

14 b) Waarmee geldt: x ≤ 120 en y ≤ 125?

Reijer wil het transport zo regelen dat de totale kosten zo laag mogelijk zijn. Daarbij moet ook rekening gehouden worden met de capaciteit van het veldrijden. Hiermee kan veldrijden van lengte 80 van rijdt worden veldrijden.

15 c) Bepaal bij welke waarden van x, y, z, w, v, u de totale kosten per week voor rijdt en veldrijden minimaal zijn.

Bron: vwo Centrale Examen Wiskunde A 1990-1995; Analyse door auteurs

E1 t/m E7: Statistiek – uit CE 1990-I

In 1972 spande een groep vrouwen een proces aan tegen een fabriek in Texas die apparaten voor air-conditioning produceert.

Deze fabriek nam alleen nieuwe personeelsleden in dienst die langer waren dan 170,0 cm. De vrouwen waren bij hun sollicitatie afgewezen omdat ze niet aan deze eis voldeden. De advocaat van de vrouwen benadrukte het discriminerende karakter van deze aanstellingsvoorwaarde door te stellen dat 91,0% van alle Amerikaanse vrouwen tussen 18 en 65 jaar niet lang genoeg was om aangenomen te kunnen worden. Dit percentage ontleende hij aan een onderzoek van het Amerikaanse ministerie van volkgezondheid.

Neem aan dat de lengte van de Amerikaanse vrouwen in de betreffende leeftijdsgroep normaal verdeeld is met gemiddelde μ en standaarddeviatie σ .

15 a) Stel uitgaande van het genoemde percentage een verband op tussen μ en σ .

Neem aan dat $\mu = 160,4$ cm.

16 a) Toon aan dat $\sigma \approx 7,2$ cm.

De groep Amerikaanse vrouwen tussen 18 en 65 jaar die langer zijn dan 170,0 cm noemen we *V*. Voor de mediaan (*m*) van de lengte van de vrouwen van *V* geldt dat 50% van de vrouwen uit *V* langer is dan *m*.

17 a) Toon aan dat $m \approx 172,6$ cm, uitgaande van $\mu = 160,4$ cm en $\sigma = 7,2$ cm.

De vertegenwoordiger van de fabriek bij het proces noemde het percentage van 91 sterk overdreven. Het door de tegenpartij aangehaalde onderzoek stamde uit 1948. De gemiddelde lengte van volwassenen was volgens hem in de periode 1948–1972 flink toegenomen. Hij ondersteunde zijn betoog met het resultaat van een recent onderzoek. In een aselecte steekproef van 100 vrouwen uit *V* werd bij 60 vrouwen een lengte gemeten van meer dan 172,6 cm. Volgens hem was bij een significantieniveau van 5% hieruit voor 1972 te concluderen dat *m* groter is dan 172,6 cm.

18 a) Onderzoek of het aantreffen van 60 vrouwen langer dan 172,6 cm in een aselecte steekproef van 100 vrouwen uit *V* bij een significantieniveau van 5% inderdaad leidt tot het verwerven van de hypothese $m = 172,6$ cm.

De advocaat van de vrouwen gaf toe dat het door hem aangehaalde onderzoek wat verouderd was en de gemiddelde lengte van de vrouwen waarschijnlijk wel was toegenomen. Hij bleef echter benadrukken dat ook in 1972 nog steeds een grote meerderheid van de Amerikaanse vrouwen op grond van hun lengte door het bedrijf zou worden afgewezen.

Stel dat voor 1972 geldt: $\mu = 164,0$ cm en $\sigma = 7,2$ cm.

19 a) Bereken het percentage Amerikaanse vrouwen in de genoemde leeftijdsgroep dat in 1972 niet lang genoeg was voor een functie bij de fabriek.

B1: Natuurkunde – 40-50% netto minder CE stof nu dan 30 jaar geleden

Vergelijking van de vwo CE Natuurkunde examenstof in 1990-1995 en 2016-2021; vooral weergegeven waar de CE-stof verschilt

uit CE verdwenen stof ● Zeer veel / moeilijk ● Zeer weinig / makkelijk

CE domeinen	CE sub-domeinen en betreffende (delen van de) examenstof uit dat sub-domein	CE leerstof		In CE?	
		Hoeveel?	Niveau?	90-95	18-21
B: Golven	B1: Wiskunde beschrijving golven mbv formules voor $v(t)$ en $a(t)$, berekenen van t als $u(t)$ gegeven is, begrip hoeksnelheid en radiaal, dopplereffect wiskundig, parate kennis $E_{\text{tril}} = 2\pi^2 m r^2 f^2$, fasesprongen bij reflecties, decibel, $L = 10 \log I/I_0$	●	●	✓	✗
C: Beweging en wisselwerking	C1a: bijna de gehele wiskundige beschrijving van bewegingen (dynamica): formules voor de eenparig (rechtlijnig) versnelde beweging, vrije val, horizontale en verticale worp, hellend vlak. Het aflezen en interpreteren van x-t, s-t, v-t en a-t-diagrammen komt nog wel voor C1b: het begrip stoot is verandering van impuls, wet van behoud van impuls incl kunnen toepassen bij (in)elastische botsingen C1c: gehele statica – draaipunt, arm, momentenwet $\sum M = 0$, $\sum F_{x,y} = 0$ C3: F_{mpz} berekeningen bij meer dan 1 kracht (nog wel bij 1 kracht)	●	●	✓	✗
D: Lading en veld	D1: Condensator ($I(t) = e^{-t/RC}$, $C = Q/V = \epsilon A/d$, $E = 0.5 C U^2$, $\sigma = Q/A$, begrip RC-tijd, $Q = \int I(t) dt$), Halfgeleiderdiodes (n-type, p-type), $V_{\text{eff}} = \frac{1}{2} \sqrt{2} V_{\text{max}}$ en $I_{\text{eff}} = \frac{1}{2} \sqrt{2} I_{\text{max}}$ D2: Baanberekeningen (dynamica) geladen deeltjes in E en B-velden, $E = \Delta V / \Delta x$ (elektrisch veld van condensator met 2 vlakke platen), parate kennis Hall-effect, equipotentiaallijnen, $V_{\text{ind}} = -N \cdot d\Phi/dt$, $B_{\text{max}} = \mu_0 I_{\text{max}} / (2mr^2)$	●	●	✓	✗
E: Straling en energie	E1 (wel in SE): Eigenschappen van stoffen en materialen – deeltjesmodellen, reële en ideale gassen, gaswetten, Van der Waalskracht, verdamping, (damp)druk, uitzetten, warmtecapaciteit, soortelijke warmte, warmtetransport, druk $p = F/A$ E2 (term 2): het licht van sterren analyseren (Herzsprung-Russell diagram) E3 Kern- en deeltjesprocessen – fotonen, muonen, neutrino's, positronen, quarks, $E=mc^2$, impuls, creatie/annihilatie, behoudswetten, paarvorming, spijting, fusie, massadefect, bindingsenergie, zwakke en sterke kernkracht	●	●	✓	✗
F: Quantum en relativiteit	F1: Quantumwereld (term 5) – quantum-tunneleffect kwalitatief beschrijven F1: Quantumwereld – tralie (inclusief formules) en parate kennis laser (bijv. populatie-inversie) F2: Relativiteitstheorie – gelijktijdigheid, tijdrek, Lorentztransformaties, energie, impuls, massa, lengtekrimp, tweelingparadox, ruimte-tijd-diagram	●	●	✗	✗
G: Leven en aarde	G1: Biofysica: Oog (gezichtshoek, scheidend vermogen, bril, loop, accommodatie, oogafwijkingen) en Oor (bouw oor, gehoordrempel) G2: Geofysica: weerkunde, vulkanisme, getijden, bewegende continenten, etc.	●	●	✓	✗
Z: Geschrapte domeinen	Z1: Lichtbrekingen en lenzen: o.a. wet van Snellius, $n_1/n_2 = v_1/v_2$, lenzenformule, constructie beeldvorming positieve lenzen, $S = 1/f$, Brewsterverster, Z2: Fysische Informatie – regelsystemen en signaalverwerking (binaire getallen, EN/OF-poort, invertor, geheugencel, comparator, pulssteller, etc)	●	●	✓	✗

1. Sub-domein E1 (Eigenschappen van stoffen en materialen) en twee van de vier genoemde sub-domeinen onderdelen (E3, F2, G1, G2) komt (nog) wel in het SE voor

Bron: Examenprogramma's vwo 2021, vwo samengevat, vwo examens Natuurkunde 1990-1995 en 2018-2021, Interviews met vakexperts; Analyse door beide auteurs

Schema 41

B1: Natuurkunde – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (1/5)

B1: Golven – uit CE 1994-I

Op het moment dat Adriaan zijn kreet slaakt, valt hij met een snelheid van 22 m/s . De temperatuur van de lucht is op die plaats 40°C .

- 7 ☐ Bereken de frequentie van de eerste formaat van de kreet zoals die wordt waargenomen door een toeschouwer op de brug recht boven Adriaan.

Na deze eerste "jump" heeft Adriaan dit nooit weer gedaan.

C1a: Mechanica – wiskundige beschrijving – uit CE 1992-I

De atomen komen met een snelheid van $9,6 \cdot 10^3 \text{ m/s}$ uit de opening van een oven links op de foto. Van hier uit bewegen ze langs een rechte baan tot ze bij punt X vrijwel stilstaan. De Na-atomen vormen daarna een dichte wolk rondom X; deze is rechts op de foto te zien.

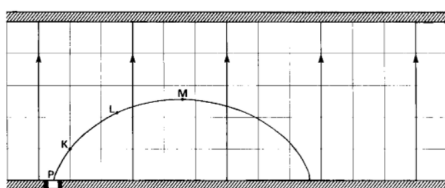
- 5 ☐ Bereken de gemiddelde vertraging van de atomen tijdens het afremmen.

C3: F_{mpz} berekenen bij 2 of meer krachten – uit CE 1990-I

In de directe omgeving van M is de Lorentzkracht op het proton vrijwel constant. Doordat ook de elektrische kracht op het proton constant is en doordat de resulterende kracht op het proton hier loodrecht op de baan staat, beschrijft het proton in de omgeving van M wel een cirkelbaan.

- 14 ☐ Het proton heeft in punt M een snelheid van $5,0 \cdot 10^7 \text{ m/s}$. Bereken de straal van deze cirkelbaan.

Bijlage:



Bron: vwo Centrale Examens Natuurkunde 1990-1995; Analyse door auteurs

Schema 42

B1: Natuurkunde – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (2/5)

D1: Elektriciteit – uit CE 1992-I

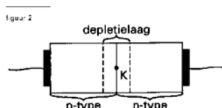
In figuur 2 is de bouw van een halfgeleiderdiode schematisch weergegeven.

Het punt K bevindt zich in de depletie laag (het barrièregebied). Er staat nog geen spanning over de diode.

- 7 ☐ Leg uit of de elektrische veldsterkte in het punt K naar links dan wel naar rechts is gericht, danwel gelijk is aan nul.

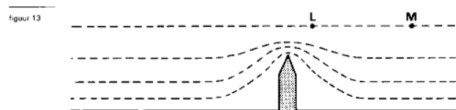
Een van de twee typen halfgeleidermateriaal waaruit de diode is opgebouwd, is vervaardigd door zuiver silicium te verontreinigen met indium.

- 8 ☐ Leg uit of dit materiaal een p-type dan wel een n-type halfgeleider is.



D2: Elektrische en magnetische velden – uit CE 1992-I

In figuur 13 zijn de vlakken van gelijke potentiaal onder de onweerswolk in de omgeving van een toren weergegeven.



- 23 ☐ Hoe blijkt uit figuur 13 dat de veldsterkte vlak boven de toren het grootst is?

Figuur 13 is ook weergegeven op de bijlage. De toren is een geleider.

- 24 ☐ Teken in de figuur op de bijlage de elektrische veldlijnen die door de punten L en M gaan.

Bron: vwo Centrale Examens Natuurkunde 1990-1995; Analyse door auteurs

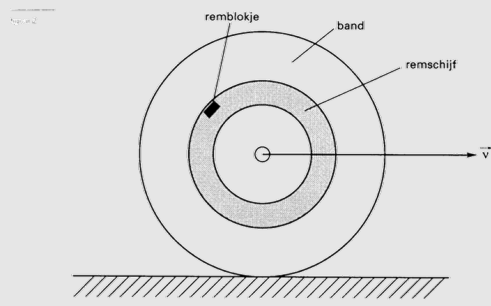
C1b: Mechanica – wet behoud impuls – uit CE 1990-II

Enkele van deze N^{2+} -ionen verkeren in een hogere energietoestand. Ze kunnen spontaan terugvallen naar de grondtoestand. De hierbij vrijkomende energie wordt dan gebruikt om het N^{2+} -ion uiteen te doen vallen in een N^+ -ion en een N-atoom. De kinetische energie van deze twee deeltjes samen is dan $0,73 \text{ eV}$ meer dan de kinetische energie van het N^{2+} -ion voor het uiteenvallen.

- 17 ☐ Bereken de grootte van de snelheid v die het zo gevormde N^+ -ion verkrijgt in het geval dat dit ontstaan is uit een stilstaand N^{2+} -ion.

C1c: Statica – uit CE 1991-I

Bij de tweede remproef wordt geremd op een glad wegdek. De wielen blokkeren dan. Dit betekent dat ze niet ronddraaien. Als de snelheid van de auto 20 m/s bedraagt, is de wrijvingskracht van de weg op één band 290 N . In figuur 2 is het wiel in zij aanzicht weergegeven.



- 5 ☐ Bereken de grootte van de wrijvingskracht die de remblokken samen op deze remschijf uitoefenen.

E1: Warmte en Gassen – uit CE 1992-I

- 16 ☐ Leg met behulp van de eerste hoofdwet van de warmteleer uit waardoor de temperatuur van het gas daalt tijdens het ontsnappen van een deel van het gas.

Zodra het kwik in de manometer aan beide zijden even hoog staat, wordt de kraan gesloten. Het nu afgesloten gas bevindt zich dan in toestand B. Hierna stroomt warmte het gas in totdat de temperatuur van het gas gelijk is aan de temperatuur van de omgeving. De gasdruk is dan gelijk aan $1,037 \cdot 10^5 \text{ Pa}$. Het gas bevindt zich dan in toestand C. Het volume van het gas in de manometer mag worden verwaarloosd.

- 17 ☐ Bereken de temperatuur van het gas in toestand B.

Men herhaalt de proef. Men laat het gas nu echter *langzaam* ontsnappen, zodat de uitzetting niet adiabatisch is.

- 18 ☐ Leg uit dat de temperatuur ook nu eerst even moet dalen.

Men creëert in een manometer als in figuur 10 een drukverschil, zodat het kwik in de daar getekende stand gaat staan. Daarop worden beide zijden van de manometer in verbinding gebracht met de buitenlucht. Het kwik in de manometer blijkt dan te gaan schommelen met trillingstijd T_0 . De demping van de beweging wordt verwaarloosd.

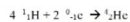
- 19 ☐ Leg uit dat deze beweging harmonisch is.

Schema 43

B1: Natuurkunde – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (3/5)

E3: Kern- en deeltjesprocessen – uit CE 1990-I

Sterren ontkenen gedurende een groot deel van hun bestaan hun energie aan de fusie van waterstof tot helium volgens de onderstaande netto kernreactievergelijking:



- 4p 19 ☐ Bereken de energie die vrijkomt bij de vorming van één heliumkern.

Als een groot deel van de waterstof in de kern van de ster is verbruikt, verliest de ster zijn stabiliteit. De kern van de ster gaat zich dan samentrekken. De temperatuur loopt daarbij zover op dat ook andere fusiereacties optreden, bijvoorbeeld de fusie van drie ^3_2He -kernen tot één $^{12}_6\text{C}$ -kern. De bindingsenergie per nucleon van deze heliumkernen is kleiner dan de bindingsenergie per nucleon van de koolstofkern.

- 3p 20 ☐ Leg met behulp van dit gegeven en met behulp van de definitie van bindingsenergie uit of bij fusie van helium tot koolstof energie vrij komt.

G1: Biofysica – uit CE 1994-II

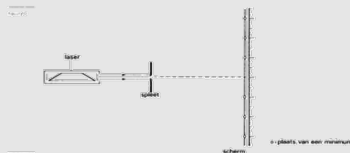
Als dezelfde persoon overdag op dezelfde afstand in dezelfde bundel kijkt, ziet hij deze veel minder helder. Er zijn twee mechanismen die voor adaptatie van het oog zorgen.

- 3p 19 ☐ Noem deze twee mechanismen en geef aan welke van deze twee het grootste bereik heeft.

Bron: vwo Centrale Examens Natuurkunde 1990-1995; Analyse door auteurs

F1: Tralie – uit CE Natuurkunde 1990-I

Men laat het licht van een dergelijke laser (golflengte 633 nm) op een spleet vallen. Figuur 10 toont een schematische weergave van de opstelling.

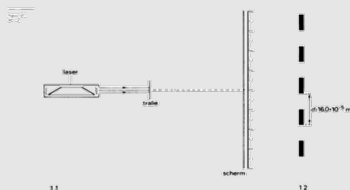


De breedte van de spleet bedraagt $0,3 \cdot 10^{-3}$ m. Op een scherm neemt men een waarnemingshoek waar. De plaatsen van de maxima op het scherm zijn in figuur 11 met een rechte weergave. De afstanden op het scherm zijn op twee groottes weergegeven.

- 2p 22 ☐ Bereken de afstand tussen de spleet en het scherm.

Men verrijkt de spleet door een tralie. Verder blijft de opstelling ongewijzigd.

Zie figuur 12. De doortrekkende delen van het tralie zijn even breed als de ruim doortrekkende delen van het tralie en ook even breed als de spleet van de opstelling van figuur 10. De tralieconstante bedraagt dus $0,3 \cdot 10^{-3}$ m. Zie figuur 12.



- 4p 24 ☐ Teken in figuur 11 op de bijlage de plaatsen van de maxima die onder deze omstandigheden op het scherm zichtbaar zijn. Leg uit hoe je de plaats van deze maxima hebt bepaald.

Schema 44

B1: Natuurkunde – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (4/5)

Z1: Licht en breking – uit CE 1990-II

Een glasvezel is een dunne glazen draad, die gebruikt wordt om lichtsignalen over te brengen. Glasvezelkabels worden bijvoorbeeld gebruikt voor telefoonverbindingen. Een glasvezelkabel bestaat uit een glasvezel, een glazen omhulling van een andere glasoort en een beschermende plastic mantel. Zie figuur 9.

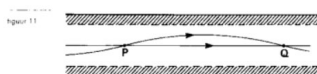


Een lichtstraal die niet evenwijdig is met de as van de glasvezel kan zich toch door de glasvezel blijven voortplanten als hij totaal reflecteert tegen de glazen omhulling. Zie figuur 10. De brekingsindex van de glasvezel in een bepaalde kabel is 1,541; die van de omhulling is 1,507.

- 3p 21 ☐ Bereken zowel de lichtsnelheid in de glasvezel als die in de omhulling.

- 3p 22 ☐ Bereken de grenshoek bij de overgang glasvezel-omhulling.

De kwaliteit van een dergelijke glasvezelkabel wordt nadelig beïnvloed doordat de lichtstralen die evenwijdig aan de as van de vezel lopen, een kortere weg afleggen dan de lichtstralen die heen en weer kaatsend verder gaan. Men kan echter de optische weglengten van deze wegen beïnvloeden door de brekingsindex vanaf de as naar de rand van de vezel geleidelijk te laten veranderen. De stralen lopen dan in de glasvezel zoals in figuur 11 is geschetst.



- 3p 23 ☐ Leg uit of men de brekingsindex van de as naar de rand van de vezel moet laten toenemen dan wel afnemen teneinde gelijke optische weglengten tussen P en Q te realiseren.

Bron: vwo Centrale Examens Natuurkunde 1990-1995; Analyse door auteurs

Z1: Licht en breking – uit CE 1993-I

Bij helder weer is de ballon tot op grote afstand met het blote oog te zien. Een waarnemer kijkt vanaf 50 km afstand naar de ballon. Het scheidend vermogen van een oog is zodanig, dat twee punten onder een gezichtshoek van $0,01^\circ$ nog juist afzonderlijk kunnen worden waargenomen.

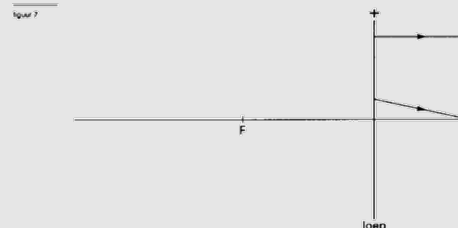
- 3p 13 ☐ Bereken of de waarnemer de linker- en rechterkant van de ballon gescheiden kan waarnemen.

De waarnemer neemt een foto van de ballon. Zijn camera heeft een telelens met een brandpuntsafstand van 250 mm.

- 3p 14 ☐ Bereken de diameter van de afbeelding van de ballon op het negatief.

De waarnemer heeft een bril met glazen waarvan de sterkte $-5,0$ dioptrie is. Zonder bril is zijn nabijheidsafstand 11 cm. Met bril kan hij op het negatief minder details waarnemen dan zonder bril. Bereken zijn nabijheidsafstand met bril. Vervaarloos hierbij de afstand tussen brillenglas en ooglenzen.

Vervolgens bekijkt hij het negatief zonder bril met een loep. In figuur 7 is een brandpunt van de loep met de letter F aangegeven. Een gedeelte van de loep van twee lichtstralen vanuit één punt van het negatief naar het oog is al op ware grootte getekend.



Deze figuur staat ook op de bijlage.

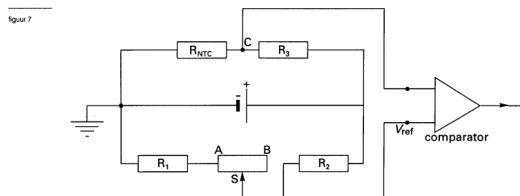
- 1p 16 ☐ Bepaal door middel van een constructie in de figuur op de bijlage de afstand tussen het negatief en de loep.

Schema 45

B1: Natuurkunde – voorbeelden van CE vragen van 30 jaar geleden die buiten de CE-stof van nu vallen (5/5)

Z2: Signaalverwerking en regelsystemen – uit CE 1994-I

De schakeling van de thermostaat staat in het schema van figuur 7.



Als het in de kamer te koud is, moet de kamerthermostaat een hoog signaal afgeven, anders een laag signaal.

Als het uitgangssignaal van de comparator hieraan niet voldoet, moet er nog een inverter achter de comparator geschakeld worden.

3p 19 ☐ Leg uit of in dit geval een inverter nodig is.

De maximale temperatuur van het water van de centrale verwarming wordt geregeld met behulp van een ketelthermostaat. Als de temperatuur van het water te hoog is, geeft de ketelthermostaat een hoog signaal af, anders een laag signaal.

Z2: Signaalverwerking en regelsystemen – uit CE 1994-I

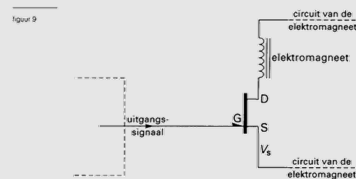
De cv-ketel moet dus het water alleen verwarmen als de kamerthermostaat een hoog signaal en de ketelthermostaat tegelijkertijd een laag signaal afgeeft. Deze twee signalen worden daartoe verwerkt in een schakeling die een hoog signaal geeft als de cv-ketel moet gaan werken. Zie figuur 8.



In deze figuur, die ook op de bijlage staat, moeten in het met een streeplijn omgeven vak nog twee verwerkers getekend worden, zodanig dat aan de gestelde eisen wordt voldaan.

3p 20 ☐ Teken op de bijlage in het daartoe bestemde vak het schema met de twee verwerkers.

Het signaal (0 V of 5 V) van deze schakeling wordt toegevoerd aan de gate van een n-junction FET. Op de drain van deze FET is een elektromagneet aangesloten. Zie figuur 9.



De elektromagneet bedient de gasklep van de cv-ketel. De gasklep wordt geopend als de FET volledig opengeschut wordt.

3p 21 ☐ Bereken hoe groot de potentiaal van de source moet zijn.

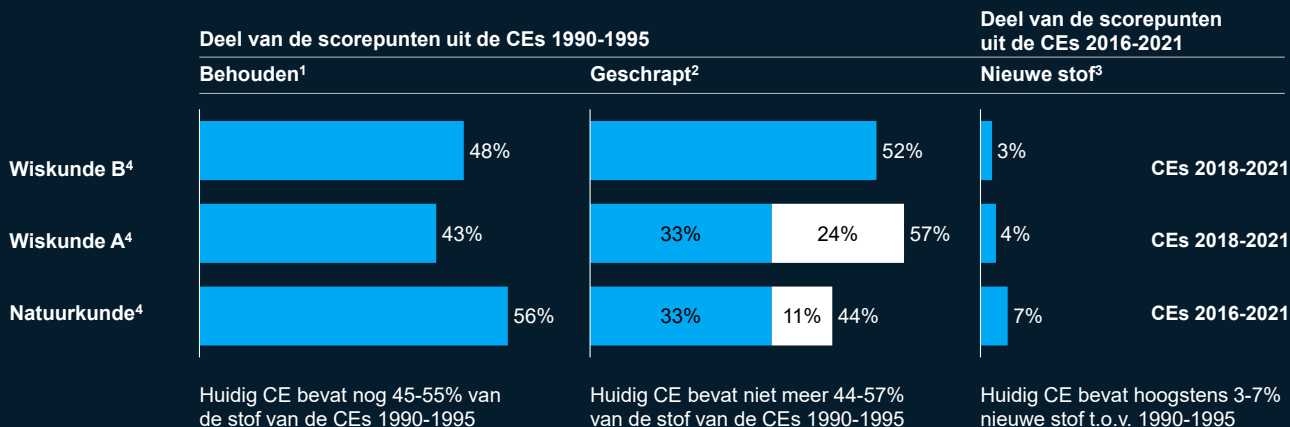
Bron: vwo Centrale Examens Natuurkunde 1990-1995; Analyse door auteurs

Schema 46

B1: nu netto 40-50% minder stof in het CE getoetst dan 30 jaar geleden

Tijdvak 1

☐ Nog wel verplicht in SE⁵



1. Deel van het CE dat beantwoord kan worden met het examenprogramma 2021

2. Deel van het CE dat niet beantwoord kan worden met het examenprogramma 2021

3. Deel van het CE dat niet beantwoord kan worden met de examenprogramma's 1990-1995

4. Bij tijdvak 1+2+3 samen zouden de resultaten zijn: Wiskunde B 49% behouden, 51% geschrapt en 2% nieuw, Wiskunde A 43% behouden, 57% geschrapt en 4% nieuw en Natuurkunde 55% behouden, 45% geschrapt en 5% nieuw

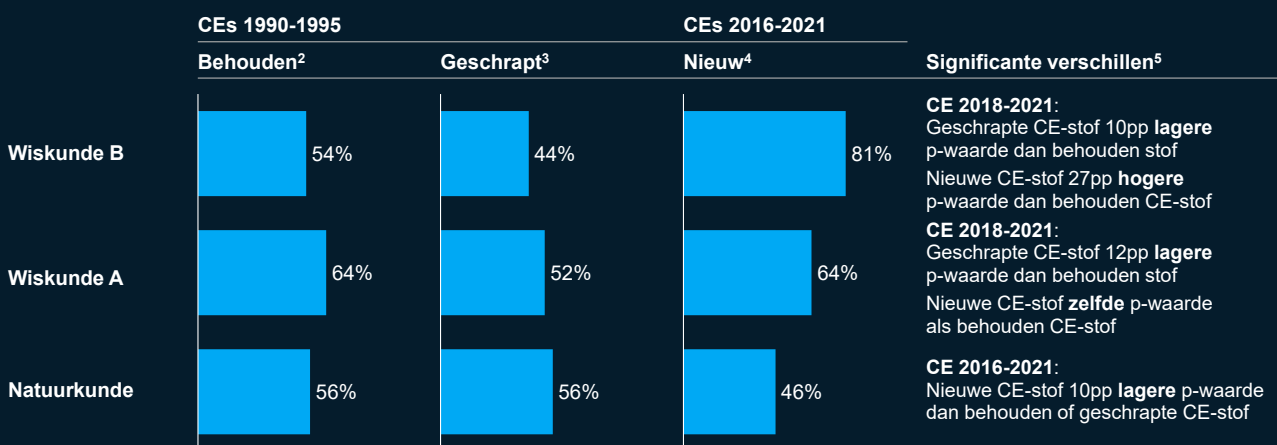
5. Bij Wiskunde A en B kan een klein deel van de verdwenen CE-stof kan nog voorkomen in het SE middels keuzeonderwerpen (40 van de 600 SLU voor Wiskunde B, ofwel 6-7% van huidig examenprogramma en 40 SLU van de 520 voor Wiskunde A, ofwel 7-8%). Scholen zijn in de keuze echter volledig vrij en een snelle controle van PTA's van een aantal scholen liet zien dat de representatie van de verdwenen CE stof middels keuzeonderwerpen in het SE zeer beperkt is.

Bron: Examenprogramma's vwo, vwo samengevat, vwo examenbundels, Vakbladen Euclides en NVOX, Analyse door auteurs

Schema 47

B1: bij Wiskunde A en B zijn vooral de onderdelen waarover moeilijkere vragen werden gesteld verdwenen

P-waarde¹ (% van de vraag goed beantwoord); Tijdvak 1



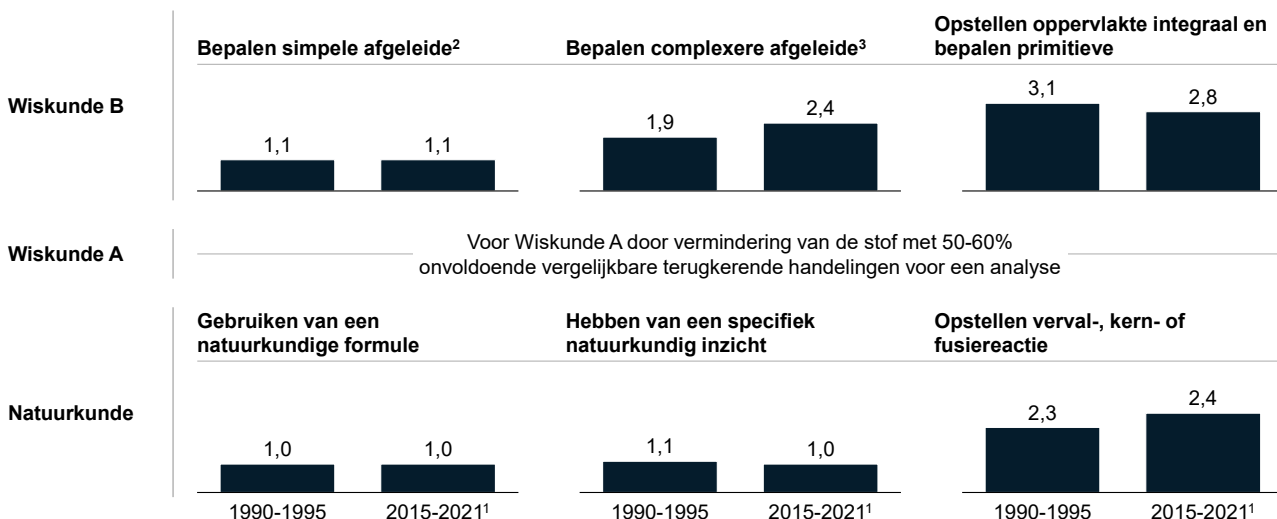
1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is. Bij p = 0 hebben alle kandidaten de vraag helemaal fout, bij p = 100% alle kandidaten helemaal goed
2. Deel van het CE dat beantwoord kan worden met het examenprogramma 2021
3. Deel van het CE dat niet beantwoord kan worden met het examenprogramma 2021
4. Deel van het CE dat niet beantwoord kan worden met de examenprogramma's 1990-1995
5. De betrouwbaarheid (1-significantieniveau) voor deze gevonden verschillen varieert van 85 tot 99%

Bron: Examenprogramma's vwo, vwo samengevat, vwo examenbundels, Vakbladen Euclides en NVOX, Cito TIA's; Analyse door auteurs

Schema 48

B2: het verkrijgen van 1 scorepunt vereist nu ongeveer evenveel inhoudelijk wiskundig en natuurkundig werk als 30 jaar geleden

Scorepunt voor vaker voorkomende handelingen; Eerste, tweede en derde tijdvakken

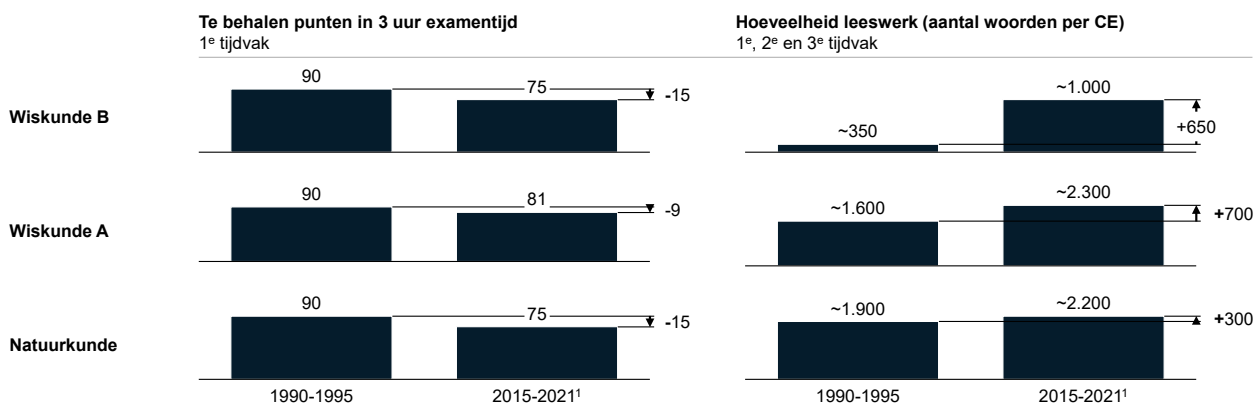


1. Exclusief 2020 (geen CE door Corona)
2. Afgeleide van standaardfuncties of een eenvoudige (lineaire) combinatie daarvan
3. Afgeleide van combinatie van standaardfuncties waarbij de product-, quotiënt- en kettingregel moet worden toegepast

Bron: CE's en correctievoorschriften 1990-1995 en 2015-2021; Analyse door auteurs

B2: nu een lager vereist tempo (5-10% minder) dan 30 jaar geleden

Vereist tempo en uithoudingsvermogen



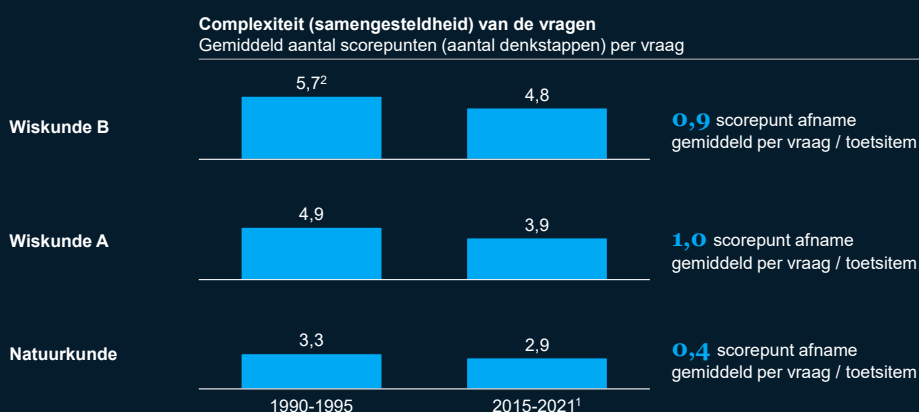
- Te behalen punten: 9-15 scorepunten minder inhoudelijk "werk" dan 25-30 jaar geleden, ofwel 10-17% minder inhoudelijk werk voor dezelfde tijd (3 uur)
- De hoeveelheid "leeswerk" is wel toegenomen waardoor 2-5% van de examentijd meer aan lezen moet worden besteed
 - Leeswerk: Het aantal te lezen woorden in een CE is toegenomen - Wiskunde B ~650 woorden meer, Wiskunde A: ~700 en Natuurkunde: ~300
 - Leestijd: De gemiddelde leessnelheid is ongeveer 200-250 woorden per minuut, maar aangezien het een CE is waarbij ook de woorden geïnterpreteerd en begrepen moeten worden is uitgegaan van ongeveer 100 woorden per minuut. Dit leidt tot de volgende extra leestijd per vak: Wiskunde B 6,5 minuut ofwel 3-4% van de tijd, Wiskunde A ongeveer 7 minuten, ofwel ongeveer 4% van de tijd, Natuurkunde 3 minuten, ofwel ongeveer 1-2% van de tijd
 - Schrijfwerk: het aantal woorden dat moet worden opgeschreven om de vragen te kunnen beantwoorden is niet onderzocht

1. Exclusief 2020 (geen CE door Corona)

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Wikipedia "leessnelheid"; Analyse door auteurs

B2: nu minder complexe (samengestelde) vragen in het CE dan 30 jaar geleden

Tijdvak 1



1. Exclusief 2020 (geen CE door Corona)

2. Wiskunde B had in 1990-1995 veel de vraag "onderzoek de functie en teken de grafiek", dit was een bewerkelijke vraag maar een standaarduitvoering van een aantal kleinere deelvragen. Deze vragen zijn niet meegenomen. Als deze vragen wel waren meegenomen, dan was het gemiddeld 5,9 scorepunt per vraag

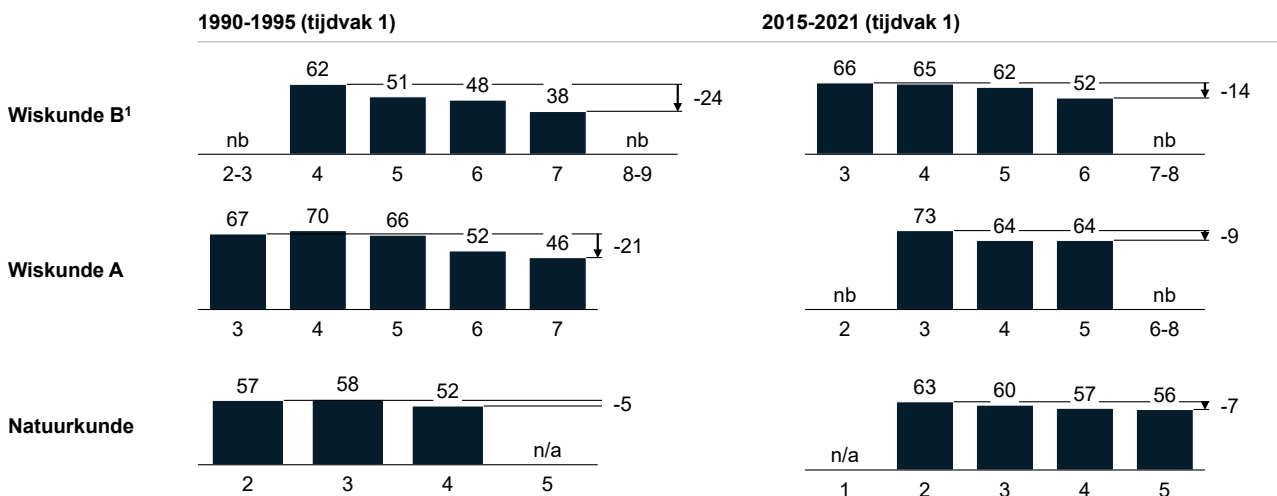
Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Het verkrijgen van 1 scorepunt vereist nu ongeveer evenveel inhoudelijk wiskundig en natuurkundig werk als 30 jaar geleden (zie **schema 48**)

Als er minder scorepunten zijn in een vraag / toets-item, dan is het aantal denkstappen van zo'n vraag typisch minder aangezien er minder inhoudelijk wiskundig/natuurkundig werk voor die vraag is

B2: hoe complexer (meer samengesteld) een vraag is, des te moeilijker is de vraag (lagere p-waarde)

P-waarde (%) versus aantal punten per vraag; alleen weergegeven als $n > 10$ per categorie²



Noot: "nb" (niet beschikbaar) wanneer er 10 of minder datapunten zijn. De gevonden resultaten zijn geverifieerd op significantie

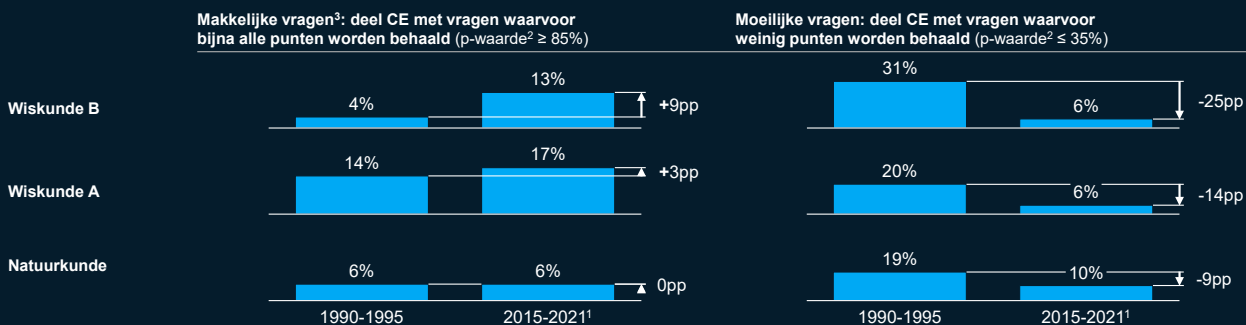
1. Wiskunde B had in 1990-1995 veel de vraag "onderzoek de functie en teken de grafiek", dit was een bewerkelijke vraag maar een standaarduitvoering van een aantal kleinere deelvragen. Deze vragen zijn niet meegenomen

2. Vooral bij Natuurkunde is de p-waarde erg afhankelijk van het CE-domein waarvan er relatief veel zijn, dus is het noodzakelijk alleen $n > 10$ categorieën mee te nemen

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021, Cito TIA's; Analyse door auteurs

B2: nu meer makkelijke en minder moeilijke vragen op het CE dan 30 jaar geleden

Eerste tijdvakken



• Als het niveau van de kandidaten toe zou nemen dan zouden ook minder vragen als moeilijk en meer vragen als makkelijk worden geclassificeerd. In onderdeel D van dit rapport laten we zien dat er geen sprake is van een significante niveautoename van de kandidaten. Voor Wiskunde A en Natuurkunde, maar ook voor Wiskunde B zou je verwachten een grotere stijging te zien van het aantal makkelijke vragen. Dit is echter niet het geval. Op basis hiervan kunnen we aannemen dat dit effect een ontwerpkeuze van het CE is.

• De gemiddelde p-waarde van vragen met een p-waarde tussen de 35% en 85% is niet substantieel veranderd: Wiskunde B (van 58% in 1990-1995 naar 58% in 2018-2021), Wiskunde A (van 60% in 1990-95 naar 63% in 2018-2021) en Natuurkunde (van 61% in 1990-1995 naar 62% in 2016-2021)

• Een sensitiviteitsanalyse is uitgevoerd en de resultaten veranderen niet significant als de p-waardegrens voor een moeilijke of makkelijke vraag bijvoorbeeld ± 5 procentpunt wordt gewijzigd

1. Exclusief 2020 (geen CSE door Corona)

2. p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde % punten dat behaald is.

Bij $p = 0$ hebben alle kandidaten de vraag helemaal fout, bij $p = 100\%$ hebben alle kandidaten de vraag helemaal goed

3. Exclusief onjuist gestelde vragen waarvoor achteraf alle punten moesten worden toegekend ($p = 100\%$). Voor 2015-2021 zou Wiskunde A 18% makkelijke vragen hebben en Natuurkunde 7%. De rest ongewijzigd.

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1990-1995; Cito website; Analyse door auteurs

B2: Nu meer hulpmiddelen toegestaan dan 30 jaar geleden

Hulpmiddel	Toelichting
 Grafische Rekenmachine	- Wiskunde: Voor een deel van de vragen kan “vooraf” het antwoord eenvoudig worden bepaald zonder (complexe) berekening. Daarna kan je naar het antwoord “toewerken”. Dit is bijvoorbeeld het geval bij het algebraïsch oplossen van een vergelijking, het bepalen van het snijpunt van 2 grafieken of de oppervlakte onder een grafiek - Wiskunde: Standaard grafieken kunnen eenvoudig worden geschetst waar voorheen dit als parate kennis nodig was
 Formules vooraf gegeven	Wiskunde A en B: een deel van de formules wordt vooraf gegeven waar dit 25-30 jaar geleden als parate kennis was vereist, bijvoorbeeld bij Wiskunde B een aantal goniometrische formules voor en bij Wiskunde A de regels voor differentiëren en rekenregels voor logaritmen
 Hints in het CE	- Nu bestaat een groter deel van het CE uit vragen waarbij het antwoord gegeven wordt - Wiskunde B: 3% in 1990-1995 versus 12% per examen in 2015-2021 (wel een significante stijging) - Wiskunde A: 10% in 1990-1995 versus 12% per examen in 2015-2021 (geen significante stijging, dus nagenoeg gelijk) - Natuurkunde: 1% in 1990-1995 versus 10% per examen in 2015-2021 (wel een significante stijging) - Dit zijn ofwel losstaande vragen (berekening/afleidingen) of vragen waarbij de leerling in een vervolgvraag het antwoord weer moet gebruiken. Dit werd 30 jaar geleden meer als één open vraag gesteld¹ - Bij Wiskunde B is het aantal hulpplaatjes en grafieken gestegen van gemiddeld ongeveer 2 per examen in 1990-1995 naar ongeveer 10 per examen in 2015-2021 - Bij Natuurkunde worden minder vragen gesteld waarbij een beroep op wiskundige vaardigheden wordt gedaan - Er worden meer hints gegeven op het CE en de denkstrategie (deelvragen van een toetsvraag/-item) worden meer expliciet genoemd in plaats van dat de kandidaat geacht wordt deze zelf te bedenken

1. Exclusief echte bewijs / onderzoeksvragen waarbij het antwoord gegeven wordt. Bijvoorbeeld “leg uit dat iets zo is” in “onderzoek of iets zo is”. Het aandeel van deze vragen is ook toegenomen bij Wiskunde B van 7% in 1990-1995 naar 19% in 2015-2021 en bij Natuurkunde van 4% in 1990-95 naar 13% in 2015-21. Bij Wiskunde A geen significante verandering: van 8% in 1990-1995 naar 5% in 2015-2021

Bron: Examenprogramma's; vwo CE examens Wiskunde A en B en Natuurkunde; Analyse door auteurs

B3: Het correctievoorschrift is nu iets minder streng

Voorbeelden van versoepelingen

Bij Wiskunde A en B en bij Natuurkunde wordt het CE nu iets minder streng (soepeler) nagekeken dan 30 jaar geleden

Voorbeelden van versoepelingen bij **Wiskunde A en/of B**

- Bij de introductie “Gelijke Monniken, gelijke Kappen” in 2015 heeft een aantal versoepelingen plaatsgevonden bijvoorbeeld rond notatiefouten (bijvoorbeeld niet aanrekenen doorrekenen: $2 \times 6 = 12 + 3 = 15$), afronden en het gebruik van eenheden
- Bij Wiskunde A mogen in totaal maar 2 punten voor het gehele CE worden afgetrokken voor afrondfouten
- Bij Wiskunde B zijn de versoepelingen niet van toepassing met betrekking tot maximeren rekenfouten en notatiefouten
- De inschatting van docenten tijdens de landelijke examenbesprekingen was dat dit **gemiddeld 1 à 2 extra scorepunten per kandidaat** oplevert (Bron: Nederlandse Vereniging Van Wiskundeleraren Vakblad Euclides 2015 nr 1 blz 15)

Voorbeelden van versoepelingen bij **Natuurkunde**

- Meer en gemakkelijker deelscores te verdienen, bijvoorbeeld voor een vraag van 4 punten konden in 1990-95 geregeld maximaal 2 deelpunten worden verdiend en dan 4 punten bij een geheel juiste beantwoording. In 2015-21 wordt er nu ook een 3^e deelpunt toegekend voor een handeling als bijvoorbeeld het opzoeken van de massa van een elektron
- Bij een aantal vragen hoeven fouten in aantal significante cijfers niet meer aangerekend te worden (bijvoorbeeld 2021 tijdvak 1 vraag 2, 11 en 18 en tijdvak 3 vraag 10 en 16) en vanaf 2022 in zijn geheel niet tenzij expliciet naar significantie wordt gevraagd (bij een aantal vragen)
- De inschatting van de auteurs en geraadpleegde vakexperts is dit **gemiddeld 1 extra scorepunten per kandidaat** oplevert

C1-3: De N-termen (minus 1) en cesuur(verschuivingen) duiden op het aantal tienden dat het cijfer wordt aangepast in beide perioden

Normhandhaving in 1990-1995 versus 2015-2021

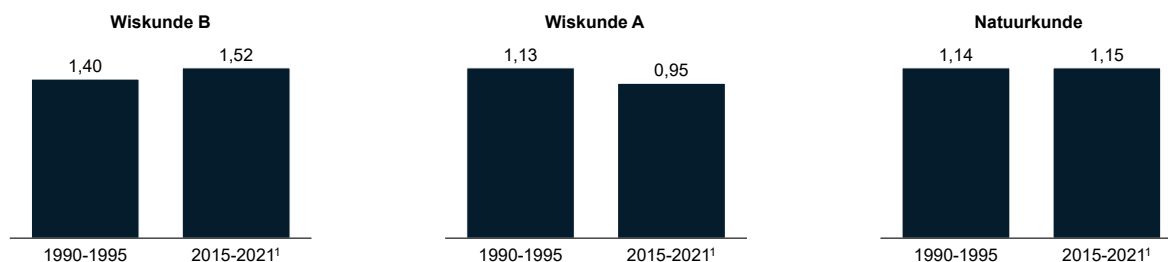
	In 1990-1995...	In 2015-2021...
Norm-handhaving	Middels een cesuur(verschuiving) die minstens 0 en hoogstens 1,0 is	Middels een normeringsterm (N) die minstens 0 en hoogstens 2,0 is
Formule voor cijfer (1 t/m 10)¹	$\text{Cijfer} = 9 \cdot \frac{\text{Behaalde score (S)}}{\text{Lengte Scoreschaal (L)}} + 1 + \text{cesuur}$	$\text{Cijfer} = 9 \cdot \frac{\text{Behaalde score (S)}}{\text{Lengte Scoreschaal (L)}} + N$
Interpretatie cijferformule	De cesuur(verschuiving) bepaalt het aantal tienden dat het cijfer wordt opgehoogd (maximaal 1,0 erbij)	- De N-term is standaard 1 - De N-term kan groter dan 1 zijn (maximaal 2,0, maar dit was bijvoorbeeld bij CE Wiskunde B 2019 2^e tijdvak 2,2) Dit duidt op het aantal tienden dat het cijfer wordt opgehoogd (maximaal 1,0 erbij) - De N-term kan lager dan 1 zijn (minimaal 0,0). Dit duidt op het aantal tienden dat het cijfer wordt verlaagd (maximaal 1,0 eraf)
Gehanteerde N-term in dit onderzoek	- Een berekende N-term op basis van de officieel gepubliceerde cesuur. Hierbij is de berekende N-term gelijk aan 1 + cesuur - Het aantal tienden dat het cijfer wordt aangepast is gelijk aan deze berekende N-term minus 1²	- N-termen zoals officieel gepubliceerd - Het aantal tienden dat het cijfer wordt aangepast is gelijk aan de N-term minus 1²

1. Geldt in het algemeen, maar bij zeer hoge en lage cijfers vindt afvlakking plaats

2. Een N-term van 1,6 betekent dat 0,6 bij het cijfer wordt opgeteld en een N-term van 0,4 betekent dat 0,4 van het cijfer wordt afgetrokken

Bron: Analyse door auteurs

C1-3: De N-termen nu zijn vergelijkbaar met 30 jaar geleden



- De N-termen voor beide perioden zijn vergelijkbaar, maar zijn tussentijds wel veranderd:
 - Wiskunde B – N-term was gemiddeld lager: 1,1 in 1996-1999, 1,0 in 2000-2004, 1,0 in 2005-2009 en 1,1 in 2010-2014
 - Wiskunde A – N-term steeg eerst en daalde toen weer: 1,2 in 1996-1999, 1,3 in 2000-2004, 0,9 in 2005-2009 en 0,9 in 2010-2014
 - Natuurkunde – N-term steeg eerst en daalde toen weer: 1,2 in 1996-1999, 1,3 in 2000-2004, 0,9 in 2005-2009 en 0,9 in 2010-2014
- Op het totaal vond dus een aanpassing van de technische N-term slechts in een beperkt deel van de gevallen plaats en was de aanpassing gering.
 - Het Cito geeft aan de technische N-term in 81% van de gevallen niet aan te passen en bij 19% wel
 - Als het wordt aangepast dan voornamelijk door correctie van een negatieve N-term die minimaal 0 moet zijn (4% van alle N-termen, en de correctie is +0,31) en door onvolkomen vragen waarvoor achteraf alle punten moeten worden toegekend (9% van alle N-termen). In beide gevallen ging de lat omlaag
 - Dus slechts in 6% van alle gevallen wordt de technische N-term aangepast ten behoeve van een daadwerkelijke extra correctie en die extra correctie is klein (veelal + 0,1 of -0,1 punt)
- Het kijken naar alleen de N-term is onvoldoende, omdat je als je makkelijkere vragen stelt en het niveau van de examenkandidaten is afgenomen, je dezelfde N-term kan krijgen

1. Exclusief 2020 (geen CE door Corona)

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Vaardigheidsontwikkelingen volgens PISA en examens (2019) door Cito auteurs Paul van der Molen, Sanneke Schouwstra, Remco Feskens en Marieke van Onna; Analyse door auteurs

C2: De normhandhaving op het vwo CE is in de praktijk niet een nauwkeurige absolute bepaling bij gebrek aan nauwkeurige absolute ijkingsinstrumenten

Aanvullende activiteiten voor de normering van de centrale examens (2015)

Vak	VMBO BB CBT	VMBO KB CBT	VMBO kb cse	VMBO gl/gt	Havo	Vwo
1 Nederlands	Aip	sst	po	po	sst	sst
2 Engels	Aip	sst	afst GT	po	po	po
3 Wiskunde B	Aip	sst	sst	pr	pr	sst/pr
4 Wiskunde A					pr	pr
5 Wiskunde C						sst
6 Duits	Aip	sst	afst GT	po	po	po
7 Biologie	Aip	sst	afst GT	sst	pr	sst/pr
8 Economie	Aip	sst	afst GT	pr		sst
9 Aardrijkskunde					sst	sst
10 M&O					sst	sst
11 Frans			afst GT	po	po	pr
12 Geschiedenis						
13 Natuurkunde/nask 1	Aip		afst GT		pr	pr
14 Scheikunde/nask 2					pr	pr
15 Kunst/beeldend				sst	sst	sst

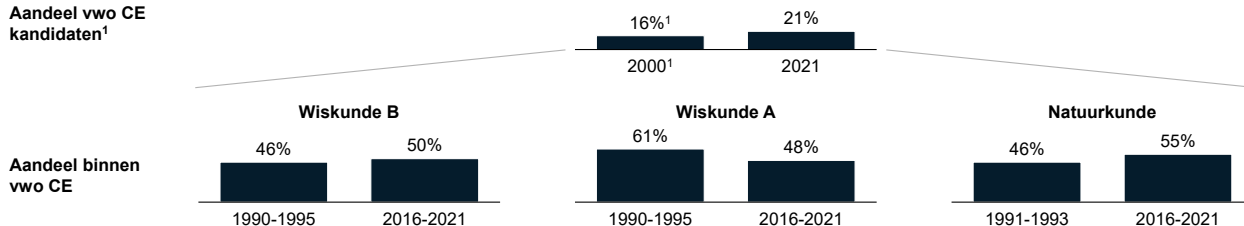
CBT = Digitale examens (computer based testing), CE = Centraal Examen, Aip = Anchor in Package, po = posttest, pr = pretest, afstGT = afstand in vaardigheid van de populaties in vmbo kb en gt, sst = standaardbepaling

Bron: Cito rapport vaardigheidsontwikkelingen volgens pisa en examens, maart 2019 (van der molen, schouwstra, feskens, van onna); Analyse door auteurs

- Het Cito geeft aan dat Anchor in Package (AIP, ankervragen) het meest nauwkeurige normhandhavingsinstrument is
- De andere normhandhavingsinstrumenten pre- en posttest (pr en po) 3x minder nauwkeurig en de standaardbepaling (sst) 9x
- Voor het vwo CE wordt AIP niet gebruikt en dus is de normhandhaving op het vwo CE in de praktijk voornamelijk niet een nauwkeurige absolute meting
- Deze normhandhavingstechnieken zijn rond 1985-1990 gestart op het vwo, en waren tot 2000 initieel op basis van de deelnemende populatie. Een moeilijkheid is dat dit in een periode was waarin het aandeel vwo leerlingen is gestegen
- Vanaf 2012 Fisher methode waar andere normhandhaving niet mogelijk was

D1-3: Het niveau van de vwo examenpopulatie is naar verwachting lager dan 30 jaar geleden

Aandeel vwo CE kandidaten¹



D1: Het gemiddelde IQ is afgelopen 40 jaar nagenoeg constant gebleven, of zeer licht afgenomen (Woodley en Meisenberg, 2013)

D2: Aan het CE Wiskunde A en B en Natuurkunde nemen nu relatief meer kandidaten deel dan 25 jaar geleden, dus is deze CE populatie relatief zwakker geworden:

- Het aandeel vwo CE kandidaten is ongeveer 30% gestegen
- Binnen het vwo CE is het aandeel Wiskunde B 4pp gestegen en Natuurkunde 9pp. Het aandeel Wiskunde A 13pp gedaald, maar dit leidt netto toch tot relatief meer kandidaten

D3: Het is niet aannemelijk dat huidige CE populatie beter is voorbereid: (1) de voorbereidingstijd: de huidige generatie heeft meer afleidingen tot zijn/haar beschikking (quote van een ervaren geraadpleegde wiskundeleraar: nu meer bijbaantjes en elke euro wordt 3x uitgegeven: "1x verdienen, 1x uitgeven en 1x herstellen", (2) de kwaliteit en hoeveelheid van het oefenmateriaal: 30 jaar geleden was ook 15-20 jaar aan oefenexamens beschikbaar en (3) de hoeveelheid en de kwaliteit van de begeleiding: het aandeel academische vwo bèta-docenten is afgenomen en bijlessen waren ook beschikbaar (weliswaar minder formeel en geïnstitutionaliseerd dan nu). Het totale effect (niet voor een individuele leerling) van de bijlesindustrie wordt beperkt geschat, omdat de omzet van de bijlesindustrie t.o.v. regulier vo-lerarensalarissen minder dan 2% is.

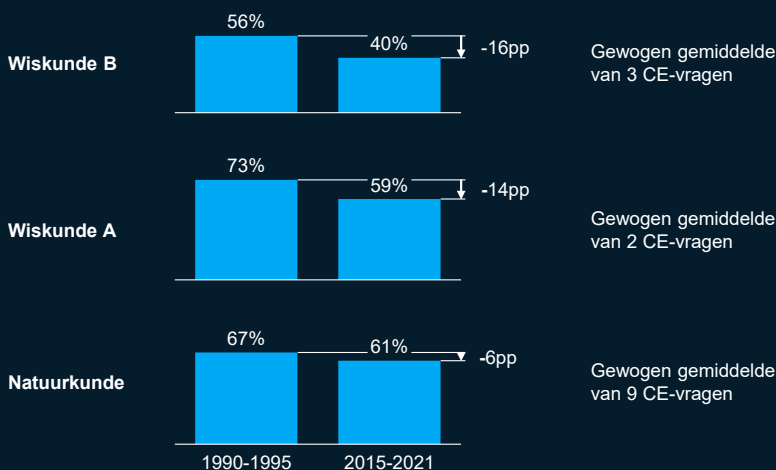
1. Uitgaande van volgende aantallen CE kandidaten in 2000: vwo 33.192, Havo 50.239, VMBO GL/TL 83.436 en VMBO KB+BB 43.735. Het aantal VMBO KB+BB is geschat op basis van het aantal in 2003 (35.596) door aan te nemen dat het aandeel VMBO BB CE kandidaten in 2000 en 2003 hetzelfde was. In 2003 was het aandeel vwo CE kandidaten 17,0%.

Bron: Examenverslagen; Vakbladen Euclides en NVOX; In the Netherlands the anti-Flynn effect is a Jensen effect (Woodley and Meisenberg, 2013), betadecanen.nl/projecten

Schema 59

D1-3: nu minder goede scores op (bijna) identieke vragen

p-waarde²; vragen uit eerste tijdvakken



1. Exclusief 2020 (geen CSE door Corona)
2. p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde % punten dat behaald is. Bij p = 0 hebben alle kandidaten de vraag helemaal fout, bij p = 100% hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1990-1995; Cito website; Analyse door auteurs

Een systematische matching van alle CE-vragen van tijdvak 1 tussen beide perioden heeft plaatsgevonden. Elke individuele vraag is eerst toegekend aan het desbetreffende examendomein. Vervolgens zijn de gevonden vragen per domein beoordeeld op vergelijkbaarheid. Dit heeft voor Wiskunde B 3 matches, voor Wiskunde A 2 matches en voor Natuurkunde 9 matches opgeleverd. Op de volgende 3 pagina's zijn de gevonden matches getoond.

Er is gecontroleerd of er een significant verschil zou zijn als het CE van 2021 weggelaten zou worden. Dit is niet het geval en bij Wiskunde A zou er zelfs een grotere niveaudaling zichtbaar zijn

De gevonden resultaten duiden op een niveaudaling onder de kandidaten. De grootte hiervan is uiteraard slechts indicatief aangezien er een beperkt aantal vergelijkbare vragen gevonden is. Het is in ieder geval niet aannemelijk dat er een (sterke) niveaustijging onder de vwo CE-kandidaten voor Wiskunde A en B en Natuurkunde heeft plaatsgevonden die de cijferstijging verklaart

Schema 60

D1-3: Wiskunde B – scores op (bijna) identieke vragen

Eerste tijdvakken

1990-1995

Vraagpaar 1: uit CE 1995 vraag 1

53%

Opgave 1

De functie f met domein $\mathbb{R}$ is gegeven door:

$$f: x \rightarrow 4 - x^2$$

Punt A ligt zo op de y -as, dat de raaklijnen door A aan de grafiek van f onderling loodrecht zijn. Bereken de y -coördinaat van A .

2015-2021

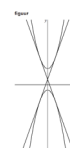
Vraagpaar 1: uit CE 2016 vraag 9

30%

Gegeven zijn de twee parabolen met vergelijkingen $y = x^2 + 3$ en $y = -x^2 - 1$.

Er zijn twee lijnen die aan beide parabolen raken. Deze twee raaklijnen snijden elkaar in het punt dat midden tussen de toppen van de beide parabolen ligt. Zie de figuur.

Stel met behulp van exacte berekeningen van beide raaklijnen een vergelijking op.



Vraagpaar 2: uit CE 1990 vraag 3

48%

Gegeven is de functie $f: x \rightarrow \frac{3 \cos x}{2 + \sin x}$, met domein $[-\pi, \pi]$.

Ten opzichte van een rechthoekig assenstelsel Oxy is K de grafiek van f .

Los op: $f(x) \cdot f(-x) = \frac{9}{5}$.

Vraagpaar 2: uit CE 2019 vraag 10

42%

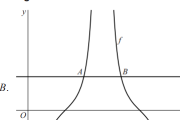
De functie f is gegeven door:

$$f(x) = \frac{\cos(x)}{-\sin^2(x)}$$

Lijn k is de lijn met vergelijking $y = \sqrt{2}$.

Lijn k en de grafiek van f hebben oneindig veel snijpunten. De punten A en B zijn de twee snijpunten met de kleinste positieve x -coördinaten. Deze zijn in figuur 1 aangegeven. Bereken exact de x -coördinaten van A en B .

figuur 1



Vraagpaar 3: uit CE 1993 vraag 2

69%

Gegeven is de functie $f: x \rightarrow \frac{(x+3)^3}{3x^2}$, met domein $\mathbb{R} \setminus \{0\}$.

Ten opzichte van een assenstelsel Oxy is K de grafiek van f .

2. Stel een vergelijking op van de scheve asymptoot van K .

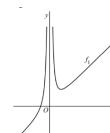
Vraagpaar 3: uit CE 2021 vraag 11

56%

Voor $p > 0$ wordt de functie f_p gegeven door:

$$f_p(x) = \frac{x^3 + 4p}{x^2}$$

In de figuur is de grafiek van f_1 weergegeven.



De grafiek van f_1 heeft een schuiv asymptoot. Bewijs dat de grafiek van f_1 boven deze schuiv asymptoot ligt.

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is. Bij p = 0 hebben alle kandidaten de vraag helemaal fout, bij p = 100% hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 61

D1-3: Wiskunde A – scores op (bijna) identieke vragen

Eerste tijdvakken

..% p-waarde (%)¹

1990-1995

2015-2021

Vraagpaar 1: uit CE 1994 vraag 18

76%

Soms raakt drinkwater door menselijke of dierlijke afvalstoffen besmet met colibacteriën. Omdat het drinken van zulk besmet water uiterst gevaarlijk is, zal een drinkwaterbedrijf in zo'n geval het water extra gaan zuiveren. In het geval er bij het begin van de extra zuivering per liter water 1800 colibacteriën zijn, geldt de formule:
 $N(t) = 1800 \cdot e^{-0,15t}$
Hierbij is $N(t)$ het aantal colibacteriën per liter water na t uren.
Bereken de groeisnelheid $N'(t)$ op het moment dat $t = 2$.

Vraagpaar 1: uit CE 2019 vraag 16

54%

Het feit dat de grafieken in figuur 2 evenwijdige rechte lijnen zijn, betekent dat het aantal aardbevingen van elke klasse exponentieel toeneemt met dezelfde groeifactor. Het totaal aantal aardbevingen A voor magnitudes $\geq 1,5$ is te beschrijven met de volgende formule:
 $A = 12 \cdot e^{0,012t}$ met $t = 0$ voor april 1994 en t in maanden.

Bereken door middel van differentiëren de waarde van de afgeleide van A voor $t = 117$. Geef je antwoord in één decimaal en leg uit wat de betekenis van deze waarde is in deze situatie.

Vraagpaar 2: uit CE 1994 vraag 19

70%

Soms raakt drinkwater door menselijke of dierlijke afvalstoffen besmet met colibacteriën. Omdat het drinken van zulk besmet water uiterst gevaarlijk is, zal een drinkwaterbedrijf in zo'n geval het water extra gaan zuiveren. In het geval er bij het begin van de extra zuivering per liter water 1800 colibacteriën zijn, geldt de formule:
 $N(t) = 1800 \cdot e^{-0,15t}$
Hierbij is $N(t)$ het aantal colibacteriën per liter water na t uren.

Zodra het aantal colibacteriën met 99% is afgenomen, stopt men de extra zuivering. Bereken in uren nauwkeurig hoe lang de extra zuivering zal duren.

Vraagpaar 2: uit CE 2018 vraag 11

64%

De formule kan zo worden herschreven dat je een moeilijkheidsgraad kunt invullen en zo de tijd in maanden kunt berekenen die nodig is om die moeilijkheidsgraad te bereiken.
Herschrijf de formule $D = 3,65 \cdot e^{0,533t}$ zo dat t wordt uitgedrukt in D .

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is.
Bij $p = 0$ hebben alle kandidaten de vraag helemaal fout, bij $p = 100\%$ hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 62

D1-3: Natuurkunde – scores op (bijna) identieke vragen (1/4)

CE Domein B1 Golven – Informatieoverdracht; eerste tijdvak

..% p-waarde (%)¹

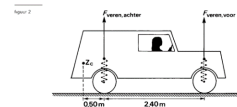
1990-1995

2015-2021

Vraagpaar 1: uit CE 1990 vraag 6

73%

De auto met inzichtende wordt nu stil gezet op een horizontaal vlak. Het kostwerk van de auto is met vier veren met elk een veerconstante van $7,5 \cdot 10^3 \text{ N/m}$ aan de wielen bevestigd. Ter vereenvoudiging is aangenomen, dat de veren verticaal staan. Zie figuur 2.



De rijk cement wordt verwijderd. Het voorste gedeelte van de auto ligt dan midden tussen de vier veren. Als de auto in zijn geheel verticaal omhoog wordt gedrukt en vervolgens wordt losgelaten, gaat de auto op en neer bewegen met een trillingstijd van 0,95 s. De massa van het trillende geheel is van de auto heeft de volgende massa.
Bereken deze afgeleide massa.

Vraagpaar 1: uit CE 2016 vraag 16

64%

Een model om het molecuul te beschrijven is een massa-veersysteem, waarbij het waterstofatoom trilt, het jodiumatoom stilstaat en de binding beschouwd wordt als een veer.

De trillingsfrequentie f van dit massa-veersysteem is $6,92 \cdot 10^{13} \text{ Hz}$. Hiermee kan met het klassieke model de veerconstante berekend worden. Voer de berekening uit.

Vraagpaar 2: uit CE 1994 vraag 5

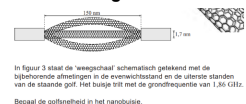
78%

De uitkomst van de berekeningen vindt Adriaan bemoedigend. Toch is hij zeer gespannen als hij de sprong waagt. Deze spanning ontlaat zich als hij tijdens zijn val uitschreeuwt: "geeseeeebblbl!". De frequentie van de eerste formant van deze kreet is 600 Hz. Beschouw zijn mondholte als een aan één zijde gesloten buis, waarbij de eerste formant de grondtoon van deze buis is. De geluidssnelheid in zijn mond is $3,5 \cdot 10^2 \text{ m/s}$. Bereken de lengte van zijn mondholte.

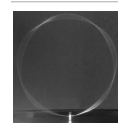
Vraagpaar 2: uit CE 2017 vraag 16 en CE 2021 vraag 9

79%

Onderzoekers beweren dat ze een 'weegschaal' hebben ontwikkeld, die een enkel proton kan wegen. De weegschaal bestaat uit een nanobuisje dat aan twee zijden is vastgeklemd en trilt als een staande golf.



In figuur 3 staat de 'weegschaal' schematisch getekend met de bijbehorende afmetingen in de evenwichtsstand en de uiterste standen van de staande golf. Het bulpje trilt met de grondfrequentie van $1,36 \cdot 10^{12} \text{ Hz}$. Bepaal de golfnelheid in het nanobuisje.



Roland en Arno bouwen de opstelling zoals schematisch weergegeven in figuur 1. In deze opstelling is een toongenerator verbonden met een trillingsapparaat. Aan het trillingsapparaat is een cirkelvormige metaal lus gekoppeld. Deze lus heeft een diameter van 24,5 cm. Bij bepaalde ingestelde frequenties op de toongenerator ontstaat er een staande golf in de cirkelvormige lus. In figuur 2 is een voorbeeld van zo'n staande golf te zien.

Op het moment dat de foto van figuur 2 gemaakt werd, stond de toongenerator ingesteld op 69 Hz. Bepaal de gemiddelde golfnelheid in de lus bij deze frequentie.

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is.
Bij $p = 0$ hebben alle kandidaten de vraag helemaal fout, bij $p = 100\%$ hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 63

D1-3: Natuurkunde – scores op (bijna) identieke vragen (2/4)

CE Domein B2 Golven – Medische beeldvorming; eerste tijdvak

..% p-waarde (%)¹

1990-1995

Vraagpaar 3: uit CE 1995 vraag 2

76%

Het technetium ontstaat in het laboratorium van het ziekenhuis als **vervalproduct** van een β^- -straler. Geef de reactievergelijking van dit verval.

Vraagpaar 4: uit CE 1993 vraag 26

61%

In Nederland is de activiteit van deze radonisotoop per m³ buitenlucht gemiddeld 3,0 Bq. De activiteit wordt gegeven door de formule:

$$A = \lambda \cdot N$$

Hierin is:

A de activiteit;
 λ de vervalconstante, die omgekeerd evenredig is met de halveringstijd;
 N het aantal radioactieve kernen.
Bereken het gemiddelde aantal ²²²Rn-kernen per kubieke meter buitenlucht.

Vraagpaar 5: uit CE 1992 vraag 1 en CE 1995 vraag 1

82%

Wanneer een foton van licht met een golflengte van 589,6 nm op een stilstaand natriumatoom in de grondtoestand valt, wordt dit foton geabsorbeerd. Het natriumatoom komt hierdoor in de eerste aangeslagen toestand. Bereken de energie van dit foton.

In ziekenhuizen wordt gebruik gemaakt van de technetium-isotoop ^{99m}Tc voor radiodiagnostiek. De 'm' achter het massagetal betekent dat de kern in een aangeslagen toestand is. Bij terugvallen uit deze aangeslagen toestand komt γ -straling vrij. De energie per γ -foton is 340 keV. Bereken de frequentie van deze γ -straling.

2015-2021

Vraagpaar 3: uit CE 2019 vraag 7

83%

Kankercellen nemen meer glucose op dan gewone cellen. Om een beeld van kankercellen te vormen, worden radioactieve isotopen ingebouwd in moleculen die sterk op glucose lijken. Een veelgebruikte isotoop is F-18, dat vervalt onder uitzending van een positron.

Geef de vervalvergelijking van F-18.

Vraagpaar 4: uit CE 2016 vraag 21 en CE 2017 vraag 22

59%

Bij het begin van het onderzoek krijgt de patiënt een hoeveelheid calcium-47 toegediend met een activiteit van 2,5 MBq. Bereken de massa in kg van het calcium-47.

De naaldjes worden tijdens een operatie in de tumor geplaatst. Ze worden 10 dagen vóór de operatie door een bedrijf in de Verenigde Staten geproduceerd. Op het moment van de operatie moet de activiteit van het I-125 in één naaldje 17 MBq zijn. Bereken de massa I-125 in kg die daartoe tijdens de productie in één naaldje moet worden aangebracht.

Vraagpaar 5: uit CE 2021 vraag 24

72%

materiaal	bandgap in eV
CdTe	1,28
Cu	0,72
GeSi	0,23
PbSe	0,27
Si	1,10
ZnS	3,60
ZnSe	2,70

In figuur 5 staat voor verschillende materialen de bandgap weergegeven. Fotonen en Lichte geïrradieert een zonnecel gemaakt van silicium. Een bepaald foton brengt een elektron in de geleidingsband van deze zonnecel.
Bereken met behulp van figuur 5 wat de golflengte van dit foton maximaal kan zijn.

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is.

Bij p = 0 hebben alle kandidaten de vraag helemaal fout, bij p = 100% hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 64

D1-3: Natuurkunde – scores op (bijna) identieke vragen (3/4)

CE Domein C3 – Beweging en wisselwerking: Gravitatie; eerste tijdvak

..% p-waarde (%)¹

1990-1995

Vraagpaar 6: uit CE 1991 vraag 24

60%

In 1982 is in het sterrenbeeld **Vulpecula** ook een neutronenster ontdekt. Deze ster heeft een massa van $1,2 \cdot 10^{30}$ kg en een straal van 13 km. Neem aan dat de ster bolvormig is.

De omwentelingstijd van deze ster is zeer klein. Als de ster echter een te kleine omwentelingstijd had, zou de gravitatiekracht niet sterk genoeg zijn om de materie aan de evenaar van de ster vast te houden. Bereken de omwentelingstijd van de ster waarbij de gravitatiekracht hiervoor nog juist sterk genoeg is. Neem hierbij aan dat de ster bolvormig blijft.

Vraagpaar 7: uit CE 1994 vraag 19

63%

Kort nadat de zon in het oosten is opgekomen, wordt vanaf de evenaar een satelliet in een 500 km hoge cirkelvormige baan rond de aarde gebracht. Op deze hoogte is de gravitatie-energie van de satelliet -17,393 GJ. De massa van de satelliet is 300 kg. Bereken de toename van de gravitatie-energie vanaf de lancering.

2015-2021

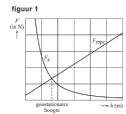
Vraagpaar 6: uit CE 2016 vraag 5

59%

Geostationaire satellieten bevinden zich namelijk op een vaste plaats boven de evenaar vanaf de aarde gezien. Een kabel tussen de aarde en een geostationaire satelliet kan niet, omdat de satelliet dan door de kabel naar beneden getrokken wordt. Maar zou een langere kabel met een contra-gewicht wel kunnen?
Hierover gaat de haalbaarheidsstudie naar de 'ruimtelift': lange en lange kabel duizenden kilometers omhoog klimmen. Wat je nodig hebt is een strakke kabel en een slimme manier van klimmen.



kabel
In figuur 1 is de gravitatiekracht op een voorwerp als functie van de hoogte boven het aardoppervlak weergegeven. Ook is de middelpuntzoekende kracht weergegeven die nodig is voor dat voorwerp als het beweegt met dezelfde omloopsnelheid als de aarde. Bereken de geostationaire hoogte.



Vraagpaar 7: uit CE 2018 vraag 6

58%

Een komeet met een massa van $12 \cdot 10^7$ kg beweegt op een hoogte van 100 km boven het aardoppervlak met een snelheid van 50 km s^{-1} richting de aarde. Bij aankomst op de aarde is de massa van de komeet afgenomen tot $6,0 \cdot 10^7$ kg. Het totaal van de kinetische energie en de gravitatie-energie is dan nog slechts 0,20% van de oorspronkelijke totale energie. (Het verschil in gravitatie-energie ten opzichte van de zon is bij deze overgang verwaarloosbaar.) Bereken met welke snelheid de komeet op de aarde aankomt.

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is.

Bij p = 0 hebben alle kandidaten de vraag helemaal fout, bij p = 100% hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 65

D1-3: Natuurkunde – scores op (bijna) identieke vragen (4/4)

CE Domein D1 en D2 – Lading en veld; eerste tijdvak

..% p-waarde (%)¹

1990-1995

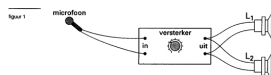
2015-2021

Vraagpaar 8: uit CE 1995 vraag 7

57%

De uitgang van de versterker is op te vatten als een spanningsbron met een bronspanning van 17 V en een inwendige weerstand van 4,0 Ω. De luidsprekers (zie figuur 1) hebben elk een weerstand van 8,0 Ω. Bereken het elektrische vermogen dat per luidspreker wordt opgenomen.

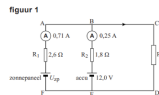
In een grote vergaderzaal wordt het stemgeluid van de spreker versterkt door een geluidsinstallatie. Zie figuur 1.



Vraagpaar 8: uit CE 2018 vraag 23

48%

Overnemen wordt om naast de accu een zonnepaneel in de schakeling op te nemen. Hiertoe wordt een proefopstelling gebouwd. Zie figuur 1.



De installatie van het weerstand is weergegeven door een gewone weerstand, aangegeven met R_x . Het zonnepaneel staat in de schakeling parallel aan de accu. Bij een bepaalde lichtsterkte op het zonnepaneel leveren beide spanningsbronnen stroom aan de weerstand R_5 en worden de in figuur 1 weergegeven stroomsterktes gemeten.

Bereken het vermogen dat in dat geval aan weerstand R_5 geleverd wordt.

Vraagpaar 9: Uit CE 1995 vraag 25

60%

De sontrilaser is een modern beveiligingssysteem met als doel kop-staat-lichtingen te voorkomen. Het systeem bestaat uit een laser die samen met een ontvanger achter de voorruit van de auto is gemonteerd. De laser zendt per seconde 1000 zeer korte elektromagnetische stralingspulsen naar voren toe uit. Na reflectie op een auto worden de pulsen door de ontvanger weer opgevangen. De straling van de laser heeft een golflengte van 859 nm. Per seconde zendt de laser 0,50 mJ stralingsenergie uit. Bereken het aantal fotonen dat per puls door de laser wordt uitgezonden.

Vraagpaar 9: uit CE 2016 vraag 4

46%

Een bepaalde, blauwe, LED zendt fotonen uit met een golflengte van 470 nm en zendt een vermogen van 0,075 W aan licht uit. De stroomsterkte door de LED bedraagt 50 mA. Uit deze gegevens volgt dat niet bij alle geleidings-elektronen die door deze LED gaan, een 'blauw' foton vrijkomt. Dit gebeurt slechts bij een bepaald percentage van de geleidings-elektronen. Bereken dat percentage.

1. De p-waarde is een indicatie voor de moeilijkheidsgraad van een vraag, ofwel het gemiddelde percentage punten dat behaald is. Bij $p = 0$ hebben alle kandidaten de vraag helemaal fout, bij $p = 100\%$ hebben alle kandidaten de vraag helemaal goed

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021; Analyse door auteurs

Schema 66

D1-3: Natuurkunde - Een extra meting geeft geen aanleiding om waargenomen niveaudaling van leerlingen bij Natuurkunde te herzien

Opzet extra test

Omdat de scoredaling op (bijna) identieke vragen bij Natuurkunde kleiner is dan bij Wiskunde A en B is extra getest of deze scoredaling toe te schrijven is aan een niveaudaling van leerlingen en niet bijvoorbeeld veroorzaakt is doordat de vragen van nu (2015-2021) uit de vraagparen moeilijker zijn dan die van toen (1990-1995).

Hiertoe is een aantal van de vraagparen getest bij een aantal leerlingen in een statistische toets. Het opzet waarbij dezelfde leerling zowel de vraag van toen als nu maakt. Het voordeel van een dergelijke opzet is dat niet een groot aantal leerlingen nodig is om enigszins statistisch betrouwbare resultaten te verkrijgen.

Hierbij wordt aangenomen dat de kans dat een leerling per toeval een vraag van nu beter maakt 50% en een vraag van toen ook 50%. De vragen die even goed worden gemaakt worden niet meegenomen.

De week voor de toets hebben de leerlingen een korte herhaal-les gekregen over de stof om ervoor te zorgen dat alle leerlingen een gelijk minimaal benodigd kennisniveau hebben. Een serieuze deelname werd beloofd met een gratis lunch.

De leerlingen hebben aangegeven een serieuze poging tot beantwoording van de vragen te hebben gedaan en dat de vragen niet bij hen bekend waren.

Resultaten en conclusies

Vraagpaar uit onderzoek (zie Schema's 62 t/m 65)	1	7	8	
• Vraag van toen (1990-1995)	1990-I vraag 6	1995-I vraag 18	1995-I vraag 7	
• Vraag van nu (2015-2021)	2016-I vraag 16	2018-I vraag 6	2018-I vraag 23	
Aantal scorepunten op vraag van...				Gemiddeld
• Toen (1990-1995)	3	3	4	3,3
• Nu (2015-2021)	3	5	4	4,0
Landelijke scores (p-waarde) op CE's van...				Gemiddeld
• Toen (1990-1995)	73%	63%	57%	64%
• Nu (2015-2021)	64%	58%	48%	56%
p-waarde van deelnemende leerlingen van op...				Gemiddeld
• ...vragen van toen (1990-1995)	33%	38%	39%	37%
• ...vragen van nu (2015-2021)	59%	66%	56%	61%
Welk aantal huidige leerlingen maakt welke vraag beter?	1	7	8	
• De vraag van toen beter gemaakt	1	2	1	
• Beide vragen even goed gemaakt	2	0	4	
• De vraag van nu beter gemaakt	6	5	4	

- De deelnemende leerlingen zijn qua niveau goed gekalibreerd met huidig landelijk niveau (van alle leerlingen): op exact dezelfde CE vragen van nu scoren zij 61% goed versus 56% landelijk
- Echter lijkt hun niveau beduidend onder dat van de leerlingen van toen te liggen: op exact dezelfde CE vragen van toen scoren zij 37% goed versus 64% landelijk toen
- De vragen van toen uit deze drie vraagparen zijn even moeilijk of moeilijker dan de vragen van nu
- (1) Dezelfde leerlingen van nu scoren 37% goed op de vragen van toen versus 61% op de vragen van nu
- (2) Meer leerlingen maken de vraag van nu beter dan de vraag van toen:
 - Vraagpaar 1: De kans dat 6 of meer leerlingen uit 7 de vraag van nu beter maken is ongeveer 6%
 - Vraagpaar 7: De kans dat 5 of meer leerlingen uit 7 de vraag van nu beter maken is ongeveer 23%
 - Vraagpaar 8: De kans dat 4 of meer leerlingen uit 5 de vraag van nu beter maken is ongeveer 19%
- De 8pp landelijke scoredaling (64% versus 54%) komt goed overeen met de 6pp scoredaling voor alle vragen

Bron: Vakblad voor Wiskundeleraren Euclides jaargang 66 t/m 71, vakblad voor o.a. Natuurkundeleraren NVOX jaar 1994/1995 en examenverslagen 2015-2021, metingen onder bereidwillige leerlingen uit 6 vwo van het Maerlant-Lyceum te Den Haag; Analyse door auteurs

Schema 67

3: Ondanks het niveau van de vwo CE populatie voor Wiskunde en Natuurkunde nu lager is, zijn de CE-resultaten gestegen; dit duidt hoe zeer deze CE's makkelijker geworden zijn

↑ Toename = Constant ↓ Afname

	Verandering	Componenten en rationale voor
A CE resultaat	=	↑ ↑ Gemiddelde cijfer: zeer sterk toegenomen Wiskunde B (1,2 gestegen), Wiskunde A (0,6 gestegen) en Natuurkunde (0,4 gestegen) ↑ ↑ Percentage voldoende: zeer sterk toegenomen Wiskunde B (58% → 84%), Wiskunde A (70% → 79%) en Natuurkunde (69% → 74%)
B Niveau van het CE	↑	↑ B1 Breedte en diepte CE stof: eenvoudiger geworden doordat de CE stof netto (zeer) fors afgenomen is: Wiskunde B (40-50% afname), Wiskunde A (45-55% afname) en Natuurkunde (40-50% afname). Bij Wiskunde A en B is voornamelijk de moeilijkere stof verdwenen ↑ ↑ B2 Eenvoud vragen: is aanzienlijk toegenomen. Evenveel vakinhoudelijk werk om een scorepunt te verdienen, maar het vereiste tempo is 5-10% lager (lengte scoreschaal is met 9-15 punten afgenomen) en makkelijkere vragen (minder denkstappen/punten per vraag en meer makkelijke en minder moeilijke vragen) ↑ B3 Souplese correctievoorschrift: is iets toegenomen, bijvoorbeeld minder eisen aan notatie, significante getallen en meer en eenvoudiger deelpunten verdienen
C Mate van norm-handhaving	=	= C1 Definitieve N-term: nagenoeg hetzelfde in 1990-1995 vergeleken met 2015-2021 = C2 Absolute of relatieve normhandhaving: de methodiek is nagenoeg onveranderd en is voor Wiskunde A en B en Natuurkunde in de praktijk niet een nauwkeurige absolute bepaling door gebrek aan nauwkeurige absolute ijkingsinstrumenten. De gehanteerde pre- en post-test zijn 3x en de standaardbepaling 9x onnauwkeuriger dan ankervraagstukken (AIP) = C3: Aanpassingen technische N-term: de N-term wordt beperkt aangepast en het is een beperkte aanpassing in beide richtingen als het gebeurt
D Niveau CE kandidaten	↓	= D1 Niveau (IQ) gehele CE populatie (alle onderwijsvormen): is (nagenoeg) niet veranderd ↓ D2 Niveau (IQ) vwo CE populatie voor een vak: relatief een zwakkere populatie doordat het percentage vwo CE is toegenomen en er relatief kandidaten meer Wiskunde B en Natuurkunde kiezen en er minder goede scores op (bijna) identieke vragen zijn ↓ D3: Hoe goed zijn de vwo CE kandidaten voorbereid (voorbereidingstijd x kwaliteit en hoeveelheid oefenmateriaal x hoeveelheid en kwaliteit van de begeleiding): Minder goed de scores op (bijna) identieke vragen

Bron: Analyses uit dit onderzoek; Analyse door auteurs

Schema 68

3: Simpele doorrekening om de cijferverandering (deels) proberen te verklaren

	Componenten	Wiskunde B	Wiskunde A	Natuurkunde
A CE resultaat	Gemiddelde cijfer (1-10)	+1,2	+0,6	+0,4
	B1 Makkelijkere CE stof (minder stof niet meegerekend)	+0,55	+0,55	-0,05
	B2 Eenvoud vragen: minder complexe (samengestelde vragen)	+0,5	+0,45	+0,1
B Niveau van het CE	B2 Eenvoud vragen: meer makkelijke en minder moeilijke	+1,1	+0,85 ¹	+0,5 ²
	B3 Soepelere correctie	+0,2	+0,2	+0,15
	B1-3 Subtotaal	+2,35	+2,05	+0,70
C Mate van norm-handhaving	C1 Definitieve N-term			
	C2 Absolute of relatieve normhandhaving	+0,1	-0,2	+0,0
	C3 Aanpassingen technische N-term			
D Niveau CE kandidaten	D1 Niveau (IQ) gehele CE populatie			
	D2 Niveau (IQ) CE populatie voor een vak	-1,4	-1,25	-0,5
	D3: Hoe goed zijn de CE kandidaten voorbereid			
Onverklaard		+0,15	+0,00	+0,20

1. 0,05 van de 0,85 cijferstijging wordt veroorzaakt door onjuist gestelde vragen waarvoor achteraf alle punten gegeven werden (p-waarde = 100%)
2. 0,15 van de 0,50 cijferstijging wordt veroorzaakt door onjuist gestelde vragen waarvoor achteraf alle punten gegeven werden (p-waarde = 100%)

Bron: Analyses uit dit onderzoek; Analyse door auteurs

3: Wiskunde B – simpele doorrekening om de cijferverandering (deels) proberen te verklaren

Factoren	Verandering	Berekening
Gemiddelde cijfer (1-10) =	+1,2	Volgt direct uit Schema 1: 7,0 gemiddeld cijfer in 2015-21 minus 5,8 in 1990-1995
B1 Makkelijkere CE stof (minder stof niet meegerekend)	+0,55	- Alleen het effect van CE-stofverandering meegenomen tegen moeilijkheidsgraad (p-waarde) van 1990-1995 - De CEs 1990-1995 hadden een gemiddelde p-waarde van 49. - Van deze CEs is 48% behouden (p-waarde 54) en 52% geschrapt (p-waarde 44) - De CEs 2018-2021 bestonden 3% uit nieuwe stof (p-waarde 81) en dus 97% uit oude behouden stof (p-waarde 54) - Dus door alleen de stofverandering zou de p-waarde dus toenemen van 49 naar 55 (6pp stijging). - Deze p-waardestijging levert een 0,55 hoger cijfer (uit 10) op
B2 Eenvoud vragen: minder complexe (samengestelde vragen)	+0,5	- Uit een regressieanalyse op de data uit Schema 51 volgt dat 1 scorepunt meer per vraag een daling van de p-waarde oplevert van ~6pp (gemiddelde van 7,5 in 1990-1995 en 4,5 in 2015-2021) - Uit Schema 50 volgt dat het aantal te behalen punten per vraag is afgenomen van 5,7 naar 4,8, dus een daling van 0,9 scorepunten. Deze daling impliceert een stijging van de p-waarde met 5-6pp (= 0,9 x 6pp) - Deze p-waardestijging levert een 0,5 hoger cijfer (uit 10) op
B2 Eenvoud vragen: meer makkelijke en minder moeilijke	+1,1	- In 1990-1995 4% makkelijke (p-waarde 90) en 31% moeilijke (p-waarde 25) en dus 65% "gewone vragen (p-waarde 58). De gemiddelde p-waarde was 49 (Schema 52) - In 2015-2021 13% makkelijke (p-waarde 89), 6% moeilijke (p-waarde 32) en dus 81% "gewone" vragen (p-waarde 58). Uitgaande van de p-waarden van 2015-2021 zou de gemiddelde p-waarde 61 zijn (Schema 52) - Deze p-waardestijging van 12pp levert een 1,1 hoger cijfer (uit 10) op - In beide perioden 1990-1995 en 2015-2021 zijn er geen onjuist gestelde vragen waarvoor na afloop alle punten moesten worden toegekend (p-waarde = 100), dus hoeft hiervoor niet gecorrigeerd te worden
B3 Soepelere correctie	+0,2	Uit Schema 54 volgt dat de soepelere correctie 1 à 2 scorepunten extra oplevert door een recente vernieuwing
C1-3 Hogere N-term	+0,1	De N-term was in 1990-1995 1,4 en in 2015-2021 1,5 (Schema 56)
D1-3 lager niveau CE kandidaten	-1,4	- Uit Schema's 59 en 60 volgt dat voor 3 (bijna) identieke vragen de p-waarde 56 was in 1990-1995 en 40 in 2015-2021 - Dit betekent een niveaudaling in p-waarde termen van 16pp over het gehele CE - Deze p-waardedaling levert een 1,4 lager cijfer (uit 10) op
Onverklaard	+0,15	

Noot: De getallen in dit schema zijn afgerond. In het model is met niet-afgeronde getallen gerekend

Bron: Analyses uit dit onderzoek; Analyse door auteurs

3: Wiskunde A – simpele doorrekening om de cijferverandering (deels) proberen te verklaren

Factoren	Verandering	Berekening
Gemiddelde cijfer (1-10) =	+0,6	Volgt direct uit Schema 1: 6,9 gemiddeld cijfer in 2015-21 minus 6,3 in 1990-1995
B1 Makkelijkere CE stof (minder stof niet meegerekend)	+0,55	- Alleen het effect van CE-stofverandering meegenomen tegen moeilijkheidsgraad (p-waarde) van 1990-1995 - De CEs 1990-1995 hadden een gemiddelde p-waarde van 57. - Van deze CEs is 43% behouden (p-waarde 64) en 57% geschrapt (p-waarde 52) - De CEs 2018-2021 bestonden 4% uit nieuwe stof (p-waarde 64) en dus 96% uit oude behouden stof (p-waarde 64) - Dus door alleen de stofverandering zou de p-waarde dus toenemen van 57 naar 64 (7pp stijging). - Deze p-waardestijging levert een 0,55 hoger cijfer (uit 10) op
B2 Eenvoud vragen: minder complexe (samengestelde vragen)	+0,45	- Uit een regressieanalyse op de data uit Schema 51 volgt dat 1 scorepunt meer per vraag een daling van een p-waarde oplevert van 5pp (gemiddelde van 5,9 in 1990-1995 en 4,5 in 2015-2021) - Uit Schema 50 volgt dat het aantal te behalen punten per vraag is afgenomen van 4,9 naar 3,9, dus een daling van 1,0 scorepunten. Deze daling impliceert een stijging van de p-waarde met 5pp (= 1,0 x 5pp) - Deze p-waardestijging levert een 0,45 hoger cijfer (uit 10) op
B2 Eenvoud vragen: meer makkelijke en minder moeilijke	+0,85	- In 1990-1995 14% makkelijke (p-waarde 91) en 20% makkelijke (p-waarde 25) en dus 66% "gewone vragen (p-waarde 60). De gemiddelde p-waarde was 57 (Schema 52) - In 2015-2021 17% makkelijke (p-waarde 90), 6% moeilijke (p-waarde 29) en dus 77% "gewone" vragen (p-waarde 90). Uitgaande van de p-waarden van 2015-21 zou de gemiddelde p-waarde 66 zijn (Schema 52) - Deze p-waardestijging van 9pp levert een 0,8 hoger cijfer (uit 10) op - In 2015-2021 was er 1 vraag (4 scorepunten) waarvoor na afloop alle punten moesten worden toegekend (p-waarde = 100). Dit levert een extra cijferstijging van 0,05 op
B3 Soepelere correctie	+0,2	Uit Schema 54 volgt dat de soepelere correctie 1 à 2 scorepunten extra oplevert door een recente vernieuwing
C1-3 Hogere N-term	-0,2	De N-term was in 1990-1995 1,13 en in 2015-2021 0,95 (Schema 56)
D1-3 lager niveau CE kandidaten	-1,25	- Uit Schema's 59 en 61 volgt dat voor twee (bijna) identieke vragen de p-waarde in 1990-1995 73 was en 59 in 2015-2021 - Dit betekent een niveaudaling in p-waarde termen van 14pp over het gehele CE - Deze p-waardedaling levert een 1,3 lager cijfer (uit 10) op
Onverklaard	+0,00	

Noot: De getallen in dit schema zijn afgerond. In het model is met niet-afgeronde getallen gerekend

Bron: Analyses uit dit onderzoek; Analyse door auteurs

3: Natuurkunde – simpele doorrekening om de cijferverandering (deels) proberen te verklaren

Factoren	Verandering	Berekening
Gemiddelde cijfer (1-10)	+0,4	Volgt direct uit Schema 1: 6,6 gemiddeld cijfer in 2015-21 minus 6,2 in 1990-1995
B1 Makkelijkere CE stof (minder stof niet meegerekend)	-0,05	- Alleen het effect van CE-stofverandering meegenomen tegen moeilijkheidsgraad (p-waarde) van 1990-1995 - De CE's 1990-1995 hadden een gemiddelde p-waarde van 56. - Van deze CE's is 56% behouden (p-waarde 56) en 44% geschrapt (p-waarde 56) - De CE's 2018-2021 bestonden 7% uit nieuwe stof (p-waarde 46) en dus 93% uit oude behouden stof (p-waarde 56) - Dus door alleen de stofverandering zou de p-waarde dus afnemen van 56 naar 55 (1pp daling). - Deze p-waardedaling levert een 0,05 lager cijfer (uit 10) op
B2 Eenvoud vragen: minder complexe (samengestelde vragen)	+0,1	- Uit een regressieanalyse op de data uit Schema 51 volgt dat 1 scorepunt meer per vraag een daling van een p-waarde oplevert van 2,5pp (gemiddelde van 2,6 in 1990-1995 en 2,4 in 2015-2021) - Uit Schema 50 volgt dat het aantal te behalen punten per vraag is afgenomen van 3,3 naar 2,9, dus een daling van 0,4 scorepunten. Deze daling impliceert een stijging van de p-waarde met 1pp (= 0,4 x 2,5pp) - Deze p-waardestijging levert een 0,1 hoger cijfer (uit 10) op
B2 Eenvoud vragen: meer makkelijke en minder moeilijke	+0,50	- In 1990-1995 6% makkelijke (p-waarde 88) en 19% moeilijke (p-waarde 27) en dus 76% "gewone vragen (p-waarde 61). De gemiddelde p-waarde was 56 (Schema 52) - In 2018-2021 6% makkelijke (p-waarde 91), 10% moeilijke (p-waarde 26) en dus 84% "gewone" vragen (p-waarde 62). (Schema 43). Uitgaande van de p-waarden van 2015-2021 zou de gemiddelde p-waarde 60 zijn - Deze p-waardestijging van 4pp levert een 0,35 hoger cijfer (uit 10) op - In 2015-2021 waren er 3 vragen (7 scorepunten) waarvoor na afloop alle punten moesten worden toegekend (p-waarde = 100). Dit levert een extra cijferstijging van 0,15 op
B3 Soepelere correctie	+0,15	Uit Schema 54 volgt dat de soepelere correctie 2 à 3 scorepunten extra oplevert door een recente vernieuwing
C1-3 Hogere N-term	+0,0	De N-term was in 1990-1995 1,14 en in 2015-2021 1,15 (Schema 56)
D1-3 lager niveau CE kandidaten	-0,5	- Uit Schema 59 en 62 t/m 66 volgt dat voor een (bijna) identieke vraag de p-waarde in 1990-1995 67 was en 61 in 2015-2021 - Dit betekent een niveaudaling in p-waarde termen van 6pp over het gehele CE - Deze p-waardedaling levert een 0,5 lager cijfer (uit 10) op
Onverklaard	+0,2	

Noot: De getallen in dit schema zijn afgerond. In het model is met niet-afgeronde getallen gerekend

Bron: Analyses uit dit onderzoek; Analyse door auteurs

De stof voor het School Examen (SE) voor Wiskunde A en B en Natuurkunde is minder dan 30 jaar geleden

Wiskunde B

~45%

minder SE stof dan 30 jaar geleden

Wiskunde A

~25%

minder SE stof dan 30 jaar geleden

Natuurkunde

~20-25%

minder SE stof dan 30 jaar geleden

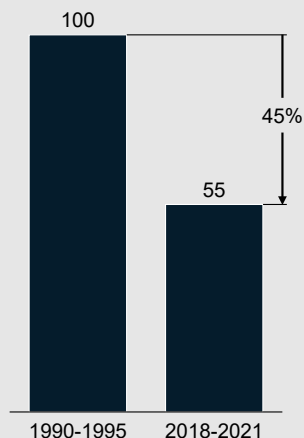
Het School Examen (SE) Wiskunde B is netto met ongeveer 45% afgenomen

Afgelopen 30 jaar

Noot: Aangenomen is dat alles wat in het SE mag, er ook inzit

Bron: Analyses uit dit onderzoek; Analyse door auteurs

Ontwikkeling SE stof Wiskunde B



Rationaal

In 1990-1995 was de stof voor het SE en CE hetzelfde

Ongeveer 45% afname van de SE stof door

- ⊖ 52% geschrapte CE stof (zie **schema 46**)
- ⊕ 3% nieuwe CE stof (zie **schema 46**)
- ⊕ 3-4% keuze onderwerpen
 - Ongeveer 6-7% van huidig programma is keuzeonderwerpen (aangenomen 40 van de 600 SLU)
 - Dat is ongeveer 3-4% van toenmalig SE programma

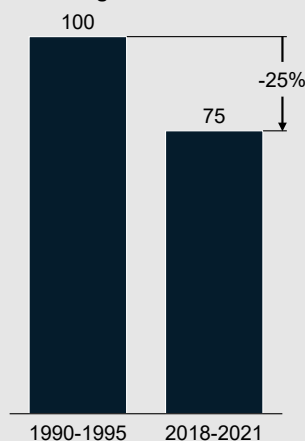
Het School Examen (SE) Wiskunde A is netto met ongeveer 25% afgenomen

Afgelopen 30 jaar

Noot: Aangenomen is dat alles wat in het SE mag, er ook inzit

Bron: Analyses uit dit onderzoek; Analyse door auteurs

Ontwikkeling SE stof Wiskunde A



Rationaal

In 1990-1995 was de stof voor het SE en CE hetzelfde

Ongeveer 25% afname van de SE stof door

- ⊖ 33% afname van de SE stof (zie **schema 46**)
- ⊕ 4% nieuwe stof in het CE, die dus ook in het SE mag
- ⊕ 5% keuze onderwerpen
 - Ongeveer 7.5% van huidig examenprogramma is keuzeonderwerpen (aangenomen 40 van de 570 SLU)
 - Dit is ongeveer 5% van toenmalig SE programma

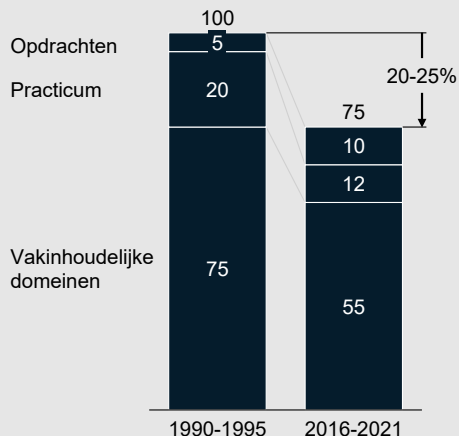
Het School Examen (SE) Natuurkunde is netto met ongeveer 20-25% afgenomen

Afgelopen 30 jaar

Noot: Aangenomen is dat alles wat in het SE mag, er ook inzit

Bron: Analyses uit dit onderzoek; Analyse door auteurs

Ontwikkeling SE stof Natuurkunde



Rationaal

In 1990-1995 bestond het SE uit (1) vakinhoudelijke domeinen (was ook de CE stof), (2) een practicum over de CE stof en (3) een modelleeropdracht

Vakinhoudelijke domeinen in het SE zijn afgenomen met 26% (33% - 7% zie **schema 46**). Dus neemt het aandeel af van 75% naar 55%

Het practicum behelst de CE stof in beide perioden. De CE stof is met 37% afgenomen (44% - 7%, zie **schema 46**). Dus neemt het aandeel af van 20% naar 12%

Opdrachten verdubbeld door naast een modelleeropdracht, nu ook een ontwerpdracht. Dus neemt het aandeel toe van 5% naar 10%

Positionering van Wiskunde en Natuurkunde in de profielen

● CE vak

	Natuur & Techniek	Natuur & Gezondheid	Economie & Maatschappij	Cultuur & Maatschappij
Algemeen verplicht	☐ Nederlands ☐ Engels ☒ Moderne vreemde taal ¹	☐ Nederlands ☐ Engels ☒ Moderne vreemde taal ¹	☐ Nederlands ☐ Engels ☒ Moderne vreemde taal ¹	☐ Nederlands ☐ Engels ☒ Moderne vreemde taal ¹
	Maatschappijleer	Maatschappijleer	Maatschappijleer	Maatschappijleer
	CKV	CKV	CKV	CKV
Profielvakken	☐ Wiskunde B ☐ Lichamelijke Opvoeding ☐ Natuurkunde ☐ Scheikunde	☐ Wiskunde A of B ☐ Lichamelijke Opvoeding ☐ Biologie ☐ Scheikunde	☒ Wiskunde A of B ☐ Lichamelijke Opvoeding ☐ Economie ☐ Geschiedenis	☐ Wiskunde C ☐ Lichamelijke Opvoeding ☐ Geschiedenis
	☒ Wiskunde D ☐ Natuur, Leven en Technologi	☒ Natuurkunde ☐ Natuur, Leven en Technologie	☐ Moderne vreemde taal ☐ Aardrijkskunde	☐ Moderne Vreemde Taal ☐ Aardrijkskunde
	☐ Onderzoeken & Ontwerpen ☐ Informatica	☐ Onderzoeken & Ontwerpen ☐ Aardrijkskunde	☐ Maatschappijwetenschappen ☐ Management & Organisatie	☐ Maatschappijwetenschappen ☐ GTC / LTC (kunstvak) / Kunstvak ²
	☐ Biologie	☒ ? ← ----- Geen Wiskunde D? ----- → ?		☐ Filosofie
Profiel-keuzevakken				

1. Gymnasiumleerlingen mogen invullen met Grieks of Latijn
 2. Het betreft het CE: kunst (algemeen)

Bron: SLO; Analyse door auteurs

Wat valt op met betrekking tot de positionering van Wiskunde en Natuurkunde in de profielen

	Natuur & Techniek	Natuur & Gezondheid	Economie & Maatschappij	Cultuur & Maatschappij
Positionering Wiskunde in profiel	Wiskunde D is niet verplicht en een 2^e moderne vreemde taal wel, terwijl Wiskunde D relevanter is voor een veel groter aantal vervolgstudies. Daarnaast kan een 2^e moderne vreemde taal een blokkade vormen voor een deel van de getalenteerde bèta's om een vwo diploma te halen	Slechts 1 Wiskunde A of B (niet beide) kan gekozen worden en Wiskunde D is geen keuzevak. Een 2^e moderne vreemde taal is wel verplicht Wiskunde is relevanter voor een veel groter aantal vervolgstudies. Daarnaast kan een 2^e moderne vreemde taal een blokkade vormen voor een deel van de getalenteerde bèta's om een vwo diploma te halen	Slechts Wiskunde A of B (niet beide) kan gekozen worden en Wiskunde D is geen keuzevak, terwijl uitgebreide wiskunde relevant is voor een aantal vervolgstudies in dit profiel (bijvoorbeeld econometrie)	Niet van toepassing
Positionering Natuurkunde in profiel	Natuurkunde is verplicht wat een logische keuze is	Natuurkunde is niet verplicht Natuurkunde en uitgebreidere Wiskunde relevanter zijn voor een veel groter aantal vervolgstudies. Daarnaast kan een 3^e moderne vreemde taal of CKV een blokkade vormen voor een deel van de getalenteerde bèta's om een vwo diploma te halen	Niet van toepassing	Niet van toepassing

Bron: Analyse door auteurs

06

Begrippenlijst

Begrip	Definitie
CE	Centraal Examen, in het vwo het landelijke eindexamen in de 6 ^e klas. In dit rapport refereren we met CE altijd aan de toetsen zoals gemaakt door Cito
Cesuur	Het aantal punten wat nodig is voor een voldoende
Cito	Stichting Cito maakt onder andere de toetsen voor het Centraal Examen op het vwo
Cito TIA	De toets- en itemanalyse. Hierin staat de moeilijkheidsgraad van een opgave, de scoreverhouding ten opzichte van de andere opgaven uit het betreffende examen en (bij meerkeuzevragen) het percentage eindexamenkandidaten dat koos voor een fout antwoord
Commissie Steur	De commissie-Steur heeft in 2019 aan minister Slob de antwoorden aangeboden op vragen van de Tweede Kamer over de ontwikkeling van de PISA- en examenresultaten in het voortgezet onderwijs in de voorgaande tien jaar
Examen	De combinatie van CE en SE
Natuurwetenschappen	In PISA en TIMMS worden Natuurwetenschappen getoetst. Dit omvat onderwerpen die we in Nederland kennen als onderdeel van natuurkunde, maar ook scheikunde, biologie en aardrijkskunde
Normhandhaving	Het aantal tienden dat bij het gemiddelde examencijfer opgeteld wordt, nadat het examen nagekeken is. Dit wordt tegenwoordig de N-term genoemd, en in het verleden de cesuur
N-term	De aanpassing van het cijfer van het CE (zie normhandhaving)
OESO / OECD	Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) / Organization for Economical Cooperation and Development (OECD)
PIAAC	PIAAC staat voor Programme for the international Assessment of Adult Competencies, en is het grootste internationale onderzoek naar kennis van volwassenen (in 32 landen)
PIRLS	PIRLS is een internationaal onderzoek dat de trends in de prestaties van leerlingen volgt op het gebied van lezen. TIMMS richt zich op groep 4
PISA	PISA is een afkorting voor Programme for International Student Assessment. Het is een internationaal vergelijkend onderzoek dat de vaardigheden en kennis in wiskunde, natuurwetenschappen en lezen van 15-jarigen test. Ongeveer 80 landen doen mee met PISA. Het onderzoek staat onder toezicht van de OESO
Procent (%)	Een eenheid voor het relatieve verschil voor een waarde (dus 20% is honderd procent hoger dan 10%)
Procentpunt (pp)	Een eenheid voor het absolute verschil voor een waarde die in procenten worden uitgedrukt (dus 20% is tien procentpunten hoger dan 10%)
p-waarde	De p-waarde betreft de moeilijkheid van een vraag. De p-waarde geeft weer hoeveel procent van de studenten de vraag goed heeft gemaakt (bij een meerkeuzevraag) of wat de gemiddelde score van alle studenten was op een open vraag. De p-waarde ligt tussen 0 en 1. Bij 0 hebben alle studenten de vraag fout beantwoord. Bij 1 hebben alle studenten de vraag goed beantwoord. Dus hoe makkelijker de vraag hoe hoger de p-waarde
Scorepunten	Het aantal punten die het gewicht van een individuele examenvraag (toetsitem) weergeeft
SE	School Examen, het deel van het eindexamen gemaakt door de school
SLO	De Stichting LeerplanOntwikkeling. Het Landelijk expertisecentrum voor het curriculum
TIMMS	TIMMS is een internationaal onderzoek dat de trends in de prestaties van leerlingen volgt op het gebied van wiskunde en natuurwetenschappen. TIMMS richt zich op groep 4 en groep 8
Toetsitem	Afgebakende vraag of opgave met een toegekend aantal scorepunten dat bij de vraag staat

Achtergrond auteurs:



Loek Zonnenberg

Loek is docent Wiskunde (zij-instromer), en een voormalige partner (alumnus) van McKinsey. Na zijn McKinsey-tijd is Loek zich gaan richten op het verbeteren van het onderwijs in Nederland. In deze periode heeft hij uit eigen beweging het initiatief genomen voor dit onderzoek en uit nieuwsgierigheid initiële analyses naar de vwo eindexamens Wis- en Natuurkunde uitgevoerd waar vanuit het onderwijsveld een positieve reactie op kwam. Voorafgaand aan zijn tijd bij McKinsey was hij tijdens zijn studententijd o.a. docent wis- en natuurkunde. Sinds deze herfst is ook Loek weer als vwo-bovenbouw docent wiskunde in Den Haag aan het werk.

Loek heeft Technische Natuurkunde (ir) gestudeerd aan de TU Delft (Quantum Fysica / Computing bij de vakgroep Quantum Transport, voorloper van QuTech) en bedrijfskunde (MBA) bij INSEAD (Singapore/Frankrijk) en heeft daarnaast een propedeuse Economie aan de Erasmus Universiteit Rotterdam.

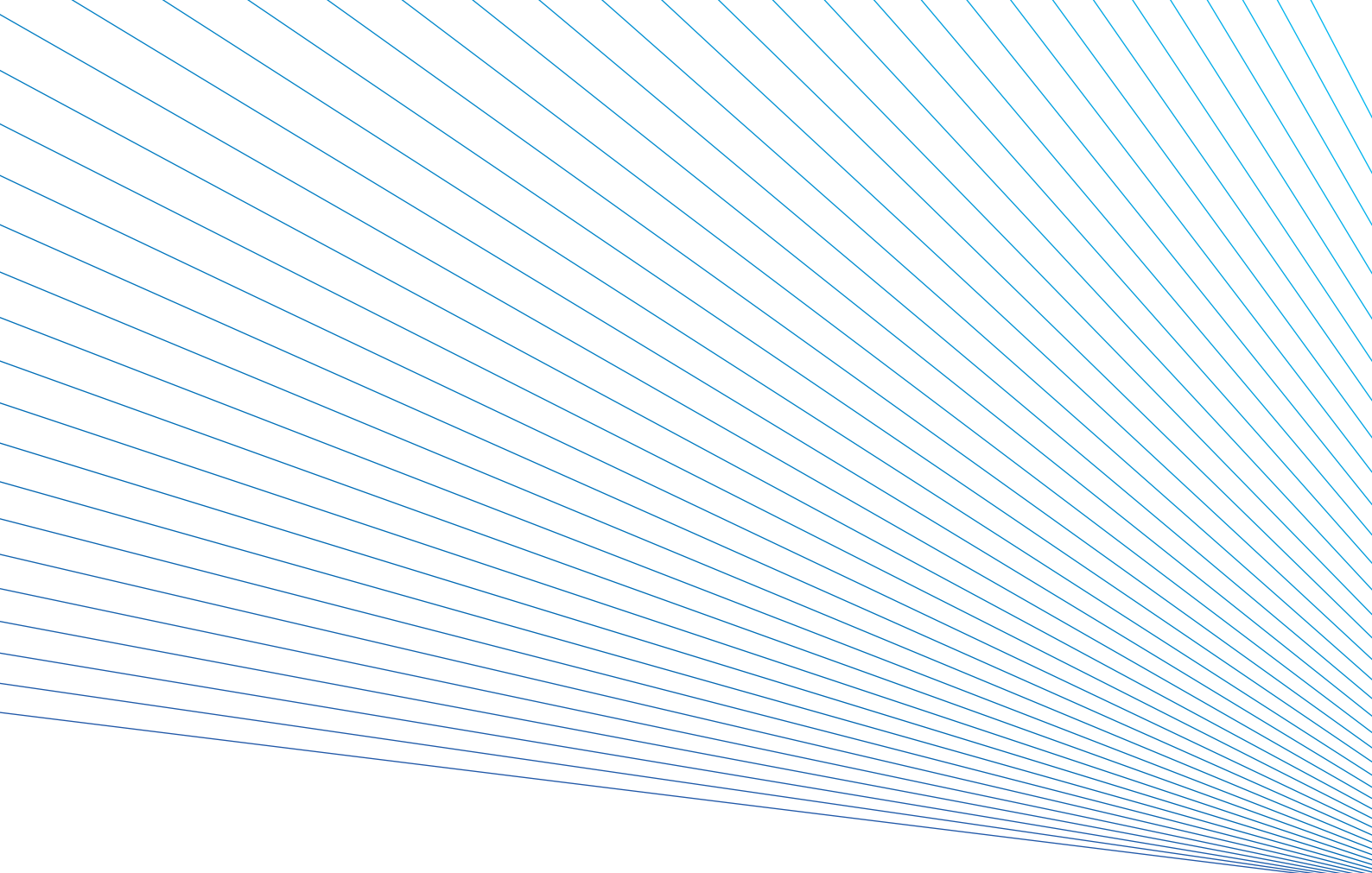


Paul Rutten

Paul Rutten is partner bij McKinsey en leidt de Nederlandse onderwijspraktijk. Hij was verantwoordelijk voor het eerdere McKinsey onderzoek *Een verstevigd Fundament* dat in 2020 is uitgekomen. Dit rapport is een logisch vervolg daarop. In *Een verstevigd Fundament* werd al gesignaleerd dat het niveau van ons onderwijs volgens PISA is afgenomen bij 15-jarigen, maar omdat eindexamencijfers hoog zijn, werden er nog vragen gesteld of er bij 18-jarigen iets aan de hand is.

Paul is daarnaast medeoprichter van Stichting leerKRACHT, Stichting de Nationale DenkTank en is voormalig bestuurslid van de Stichting Avond van Wetenschap en Maatschappij.

Paul heeft Technische Natuurkunde (ir) gestudeerd aan de TU Delft waar hij afgestudeerd is bij Quantum Transport en bedrijfskunde (MBA) bij Kellogg aan de Northwestern University in Chicago (cum laude).



Februari 2023

Copyright © Loek Zonnenberg en McKinsey & Company

Designed door Visual Media Europe

www.mckinsey.com

 @McKinsey

 @McKinsey