

“Position paper” – Loek Zonnenberg, MSc, MBA, MEd.

Rondetafelgesprek Vaste Kamercommissie OC&W over toetsing, examens en selectie in het funderend onderwijs d.d. 24 april 2025, Blok 2 van 15.30-16.30 uur.

Introductie: Rapport Toetsen getoetst

Het rapport “Toetsen getoetst – Het begrijpen van het langzaam verlaagde niveau van het vwo Centraal Examen (CE) Wiskunde en Natuurkunde (februari 2023)” is tot stand gekomen in samenwerking tussen 2 partijen vanwege hun gedeelde passie voor onderwijs: Loek Zonnenberg, docent wiskunde en voormalig Partner bij McKinsey & Company, en Paul Rutten, Partner bij McKinsey & Company. Zij hebben hun rapport uitgebracht onder de vlag van McKinsey en het kan dus gezien worden als een **McKinsey onderzoek**. Echter Loek Zonnenberg is momenteel niet meer werkzaam bij McKinsey en kan dus niet namens McKinsey spreken en zal dit dus namens zichzelf als privépersoon doen. Zie **BIJLAGE 1** voor de achtergrond van Loek Zonnenberg.

Het rapport is onafhankelijk, reflecteert de inzichten van de auteurs, en is niet geschreven in opdracht van enig bedrijf, de overheid of ander instituut. Een divers wetenschappelijke groep van meer dan 40 experts (zie **BIJLAGE 2**) heeft de methodieken en analyses gevalideerd.

Er is sprake van een paradox in het Nederlandse onderwijs. Enerzijds is het gemiddelde cijfer voor de centrale vwo-eindexamens wiskunde (A&B) en natuurkunde in de afgelopen dertig jaar gestegen, maar de scores van Nederlandse leerlingen op andere nationale en internationale toetsen lopen terug.

Wat is nu waarheid? We hebben vastgesteld dat het niveau van vwo wiskunde (A&B) en natuurkunde de laatste 30 jaar fors is gedaald (ook voor Corona), zowel qua niveau van de Nederlandse leerlingen die het vwo verlaten als qua niveau waarop zij geëxamineerd zijn. Wij hebben dit bekeken op basis van meer dan 1500 vragen uit Centrale Examens over de laatste 30 jaar (1990-95 versus 2015-21).

Uit onze analyses blijkt dat voor wiskunde (A&B) en natuurkunde ruim 40 procent (!) van de stof van het curriculum is verdwenen uit de centrale examens. Dit is heel erg geleidelijk gegaan waardoor deze trend gemakkelijk gemist is. Door de data over een periode van 30 jaar te bekijken, is deze trend wel zichtbaar geworden. Het gaat voornamelijk om de moeilijkere stof. Wat overblijft zijn de onderwerpen waar op de examens simpelere vragen over moeten worden beantwoord en waarvoor minder stappen hoeven te worden doorlopen. Tel daar de iets soepeler geworden beoordeling bij op en je hebt het antwoord op de vraag waarom cijfers hoger uitvallen terwijl het niveau daalt.

In het onderzoek is daarnaast vergeleken hoe vwo-eindexamenleerlingen van vroeger en nu landelijk scoorden op bijna identieke vragen. De leerlingen van nu maakten deze minder goed dan de generatie van hun ouders. Op identieke vragen scoorden de huidige leerlingen 0,5-1,5 punt op het cijfer lager dan vroeger. De substantiële daling van het niveau is sluipenderwijs tot stand gekomen. Het is moeilijk om prestaties van leerlingen te kwantificeren en te vergelijken, en onderzoek hiernaar gaat met onzekerheden gepaard. Mede daardoor hebben we in Nederland pas laat zicht gekregen op dit probleem. Dit dalende niveau kan belangrijk zijn omdat we wellicht jongeren kansen ontnemen wanneer we hen niet optimaal stimuleren en ondersteunen om zich volop te ontwikkelen. Ook is het voor onze kenniseconomie, welvaart en welzijn belangrijk dat Nederland niet achterop raakt.

Zie **BIJLAGE 3** voor het onderzoek “Toetsen getoetst” in één oogopslag

De gestelde vragen en een eerste antwoord als startpunt voor de discussie

[1] *Hoe komen toetsen en examens tot stand? Hoe en door wie worden toetsen en examens gemaakt, hoe worden de referentieniveaus Nederlandse taal en rekenen hierin verwerkt? Hoe wordt de kwaliteit van de toetsen en examens gegarandeerd?*

Centrale Examens (CE): door Cito en conform de richtlijnen en procedures van het CvTE.

Schoolexamens (SE): voornamelijk door scholen conform de inhoudelijke richtlijnen van het CvTE en procedures binnen scholen (bijvoorbeeld middels lokale toets experts, examensecretarissen) met per vak de vergelijking met van het SE-cijfer met het CE-cijfer middels inspectienormen (het SE-cijfer mag niet 0,5 of meer hoger dan het CE-cijfer zijn).

Overige toetsen in het VO: voornamelijk door scholen conform het landelijke curriculum vertaald en vastgelegd in de verschillende lesmethoden die ook (voorbeeld)toetsen aanleveren.

[2] *Hoe wordt bepaald of leerlingen een toets of examen gehaald hebben? Wordt er getoetst aan een absolute of relatieve norm? Hoe kunnen we de toets- en examenresultaten over de jaren heen (en in het po: over toetsen heen) met elkaar vergelijken?*

Voor de Centrale Examens (CE) wordt een norm vastgesteld door de prestaties van leerlingen met voorgaande jaren te vergelijken. Daarbij hanteert het CvTE het uitgangspunt dat het gemiddelde cijfer bij examenpopulaties die van jaar tot jaar even vaardig zijn, gelijk is. Daarnaast is de normhandhaving op het havo en vwo CE in de praktijk niet een nauwkeurige absolute bepaling bij gebrek aan nauwkeurige absolute ijkingsinstrumenten (bron: rapport Toetsen Getoetst, zie **BIJLAGE 4**).

Naar mijn mening wordt ideaal gezien van tevoren vastgesteld wat er geleerd moet worden op basis waarvan wordt getoetst. Hierbij bepaalt de docent/instelling wat wordt gezien als beheersing (“mastery”) en welke absolute normering hierbij hoort. Toetsing met relatieve normering levert geen informatie over de mate waarin de leerdoelen zijn behaald.

Om CE-resultaten over de jaren heen te kunnen vergelijken kunnen het beste twee (of meer) perioden met vergelijkbare centrale examens met elkaar worden vergeleken. Bij voorkeur dienen deze twee perioden zo ver mogelijk uit elkaar te liggen om trends zichtbaar te maken en beslaat elke periode minimaal 5 examens (om effecten van individuele jaren uit te middelen en om voldoende statistische basis te geven). Dan onderzoek je met behulp van kwantitatieve analyses alle factoren (bijvoorbeeld de moeilijkheidsgraad van de CE-stof) die het verschil in CE-resultaat (landelijk gemiddeld cijfer) kunnen verklaren. Zie **BIJLAGE 5** voor de methodiek uit Toetsen Getoetst. Op nationaal niveau zou je een grote database met vragen kunnen aanleggen waaruit en waarmee jaarlijks een steekproef wordt genomen.

[3] *Hoe verhouden toetsen/examens en leermiddelen zich tot elkaar?*

Leerdoelen, leervormen en toetsing op één lijn liggen en zouden op elkaar moeten aansluiten (“constructive alignment”, zie het werk van prof. John B. Biggs). Het hangt mijns inziens wel van de aanbieders van lesmethodes van het vak af (die zijn voor sommige vakken beter dan voor andere) en van individuele docenten af. Een landelijke kwaliteitstoets voor lesmethodes is te overwegen.

[4] *Welke rol spelen toetsen en examens bij het selecteren van leerlingen voor het vervolgonderwijs? Wat toetsen we precies bij leerlingen? Wat gaat hierbij goed en wat kan beter?*

Een grote rol. Toetsen en examens behoren naar mijn mening te bepalen of leerlingen/studenten het ingangsniveau hebben qua kennis en vaardigheden voor het vervolgonderwijs, zeker als vervolgonopleidingen een (veel) grotere instroom dan opleidingsplekken hebben. Wij toetsen mijns inziens voornamelijk kennis en vaardigheden.

Examencijfers hebben voorspellende waarde voor het succes in vervolgstudies (zie bijvoorbeeld het werk van prof. dr. Peter van der Heijden en rapport Toetsen Getoetst schema 2).

Naar mijn mening wordt er meer (achteraf) genormeerd op de populatie dan streng geselecteerd. Bijvoorbeeld het aandeel havo / vwo leerlingen in het derde leerjaar is in 30 jaar tijd ongeveer 50 procent gestegen (van 31% naar 48%) en het aandeel vwo CE-kandidaten in 20 jaar tijd ongeveer 30 procent (van 16% naar 21%) (bron: Toetsen Getoetst, Schema 3 en 58).

De vraag is of (centrale) examens nog wel voldoende de juiste stof op de juiste moeilijkheidsgraad toetsen. Bijvoorbeeld de studie Luchtvaart- en Ruimtevaarttechniek kan ten opzichte van de instroom steeds minder Nederlandse leerlingen met een vwo-diploma een plek bieden, omdat zij niet (meer) door de toelatingstoets komen (bron: Toetsen Getoetst, Schema 14). Er lijken ook steeds meer door universiteiten georganiseerde toelatingstoetsen te komen. Verder kan worden gekeken of sommige onderdelen van een schoolexamen ook landelijk kunnen worden afgenomen (bijvoorbeeld luistervaardigheidstoetsen).

[5] *Wat zijn andere functies van toetsen en examens en in hoeverre staan deze functies ter discussie?*

Er zijn toetsvormen die formatief zijn (Hoe gaat het met het leren van de leerling?, Hoe gaat het met mijn onderwijs?) en ook toetsen die bedoeld zijn als leerstrategieën (“no stakes” toetsing). Een (centraal) examen is naar mijn mening een eindkwalificatie in hoeverre iemand geleerde kennis en vaardigheden beheerst en een “high-stakes” selectie-instrument om leerlingen op de juiste plek daarna te krijgen.

[6] *Wat zijn (on)gewenste neveneffecten van de manieren waarop we binnen ons stelsel toetsen en examineren? Hoe kan het anders?*

Ik ben van mening dat het hebben van een landelijk Centraal Examen (CE) een groot goed is dat kansengelijkheid bevordert. Een leerling moet zelf precies hetzelfde examen op hetzelfde moment onder vergelijkbare omstandigheden maken als andere leerlingen met exact hetzelfde beoordelingsinstrument en normering. Bijvoorbeeld bij toetsen middels een praktische opdracht die thuis gedaan mag worden, kunnen leerlingen met ouders met voldoende middelen en/of kennis extra hulp inschakelen. Er is natuurlijk ook een effect van bijles, maar bij een CE moet een leerling uiteindelijk wel nog zelf maken het examen maken, wat bij een praktische opdracht vermeden kan worden.

Voortgangstoetsen zijn mijns inziens een ontwikkelingsfeedback voor leerling en docent om het leren verder te optimaliseren. Hierbij kan worden gedacht dat leerlingen optimaal leren als ze op hun eigen niveau haalbaar uitgedaagd worden (zone van naaste ontwikkeling) opdat zij succeservaringen voor hun persoonlijke ontwikkeling krijgen. Doorgaans wordt stress pas ervaren als je het gevoel hebt dat je het niet aankunt. In hoeverre is dit te koppelen aan groei van het aandeel leerlingen op de havo en het vwo?

BIJLAGE 1: Achtergrond auteur Loek Zonnenberg

Gedurende mijn ongeveer 17 jaar bij McKinsey (Partner/vennoot) heb ik mij gespecialiseerd in grote complexe (internationale) transformaties en “capability building”, vaak gerelateerd aan onderwijs-, leer- en trainingssituaties (o.a. het opzetten van corporate academies, het opleiden van executives en “change agents”).

Na mijn McKinsey-tijd ben ik mij gaan richten op het verbeteren van het onderwijs in Nederland. In deze periode heb ik uit eigen beweging het initiatief genomen voor dit onderzoek en uit nieuwsgierigheid initiële analyses naar de vwo eindexamens wis- en natuurkunde uitgevoerd.

Voorafgaand aan mijn tijd bij McKinsey was ik o.a. docent wis- en natuurkunde op een middelbare school met veel leerlingen met een migratieachtergrond in een achterstandssituatie (Johan de Witt in Den Haag). Sinds de herfst van 2022 ben ik weer als vwo-bovenbouw docent wiskunde aan het werk in Den Haag op het Maerlant-Lyceum.

Mijn achtergrond is in Technische Natuurkunde aan de TU Delft (ir./MSc), Business Administration aan INSEAD in Fontainebleau/Singapore (MBA) en eerstegraads docent wiskunde aan het ICLON van de Universiteit Leiden (drs./MEd).

BIJLAGE 2: Geraadpleegde experts voor het rapport Toetsen getoetst

Een divers wetenschappelijke groep waaronder mensen met een achtergrond in wiskunde, natuurkunde, didactiek / pedagogiek, psychometrie en andere vakken die nauw verbonden zijn met dit onderzoek) geraadpleegd en zij hebben onze methodiek en analyses gevalideerd

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

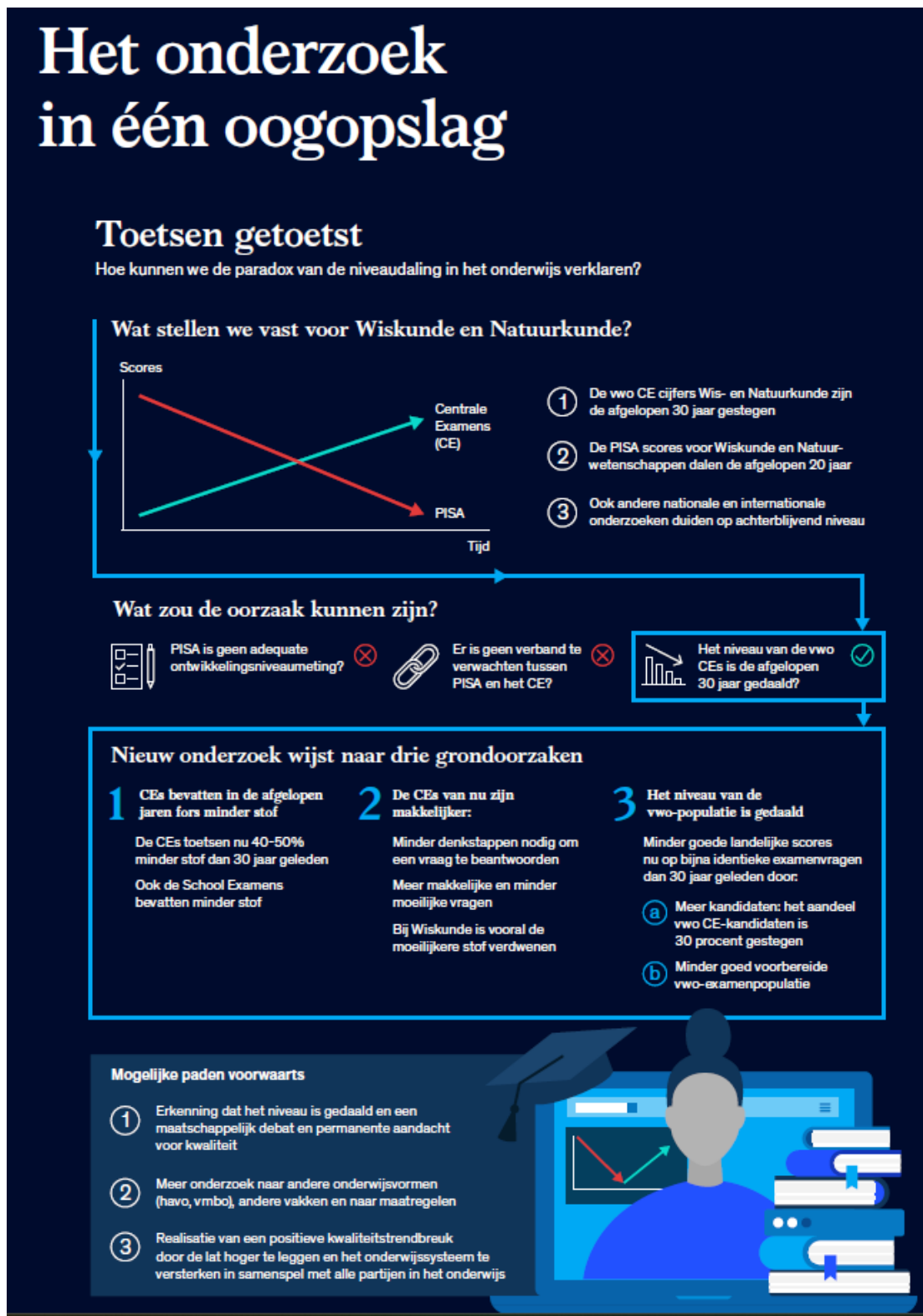
technisch natuurkundig ingenieur aan de TU Delft
en gepromoveerd theoretisch natuurkundige aan
de UvA

Jenny Lu, BSc

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

BIJLAGE 3: Het onderzoek Toetsen getoetst in één oogopslag

Deze visualisatie geeft in één oogopslag een overzicht van de paradox aan de basis van het onderzoek, de getoetste hypothesen, belangrijkste inzichten en de mogelijke paden voorwaarts.



BIJLAGE 4: Normhandhaving bij de centrale examens

Schema 57

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

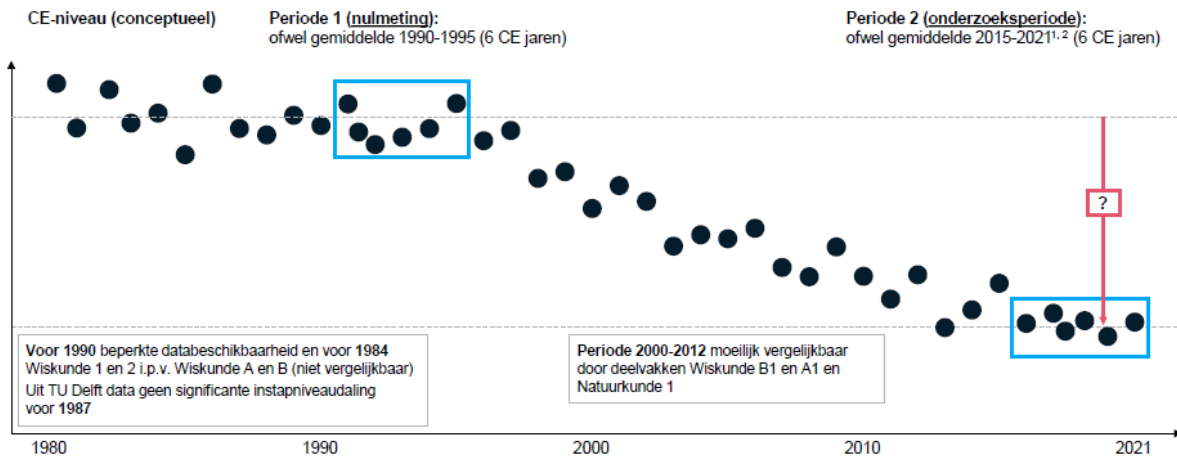
Bron: Cito rapport vaardigheidsontwikkelingen volgens pisa en examens, maart 2019 (van der molen, schouwstra, feskens, van onna); Analyse door auteurs

- 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

BIJLAGE 5: De methodiek uit Toetsen Getoetst

Schema 30

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

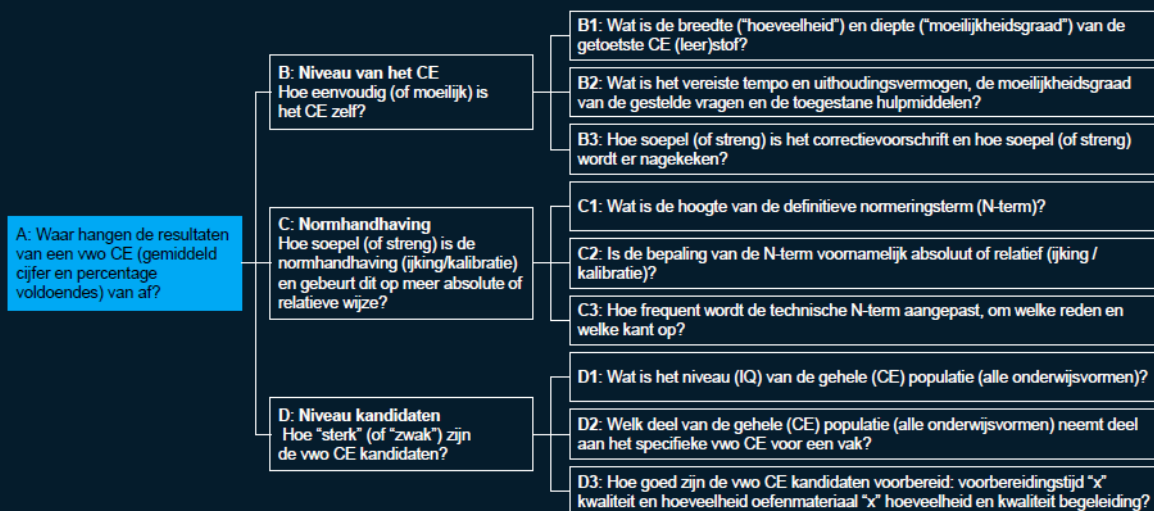


1. Inclusie of exclusie van coronajaar 2021 maakt geen significant verschil in de analyses in dit rapport; In 2020 was er geen CE
2. Voor een van de analyses met betrekking tot de stofverandering ligt de nadruk op 2018-2021 voor Wiskunde en 2018-2021 voor Natuurkunde i.v.m. de curriculumwijziging die vlak hiervoor plaatsvond

Bron: Analyse door auteurs

Schema 31

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



Bron: Analyse door auteurs