

VO

Vaardigheid
Examen-
kandidaten 2021



Vaardigheid

Examenkandidaten 2021

Cito

**Marieke van Onna, Sanneke Schouwstra,
Paul van der Molen, Maarten de Groot
en Wilco Emons**

Colofon

Opmaak: Media Support

Foto omslag: Gijs Versteeg

© Stichting Cito Instituut voor Toetsontwikkeling Arnhem (2021)

Alle rechten voorbehouden. Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van Stichting Cito Instituut voor Toetsontwikkeling worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

Inhoud

1	Inleiding	8
1.1	Aanleiding	8
1.2	Normering 2021	9
1.3	Onderzoeksvragen	9
1.4	Indeling van dit rapport	10
2	Methoden van onderzoek	12
2.1	Gegevens voor het onderzoek	12
2.2	Schatting van vaardigheidsontwikkeling	12
2.3	Berekeningen deelvragen	15
2.3.1	Deelvraag 1	15
2.3.2	Deelvraag 2	15
2.3.3	Deelvraag 3	16
2.3.4	Deelvraag 4	16
3	Resultaten	18
3.1	Vaardigheidsontwikkeling	18
3.2	Resultaten per deelvraag	19
3.2.1	Herkansers in 2019 en 2021	19
3.2.2	Kandidaten met lage scores	20
3.2.3	Eerstekansers in tweede tijdvak en hun gemiddelde cijfer	22
3.2.4	Spreiding tussen scholen	25
4	Conclusie	28
5	Context bij de resultaten	32
6	Literatuur	36
	Appendix I Context van het Centraal Examen 2021	42
	Appendix II Berekeningswijze vaardigheidsverschil tussen 2021 en pre-corona	47

Samenvatting

Kandidaten die in 2021 eindexamen in het voortgezet onderwijs deden, hebben een andere voorbereiding gehad op het eindexamen dan in 2019 en eerder omdat de corona-pandemie zorgde voor lockdowns, online onderwijs en quarantainesituaties. Bij het centrale examen in 2021 werden er meerdere maatregelen genomen om kandidaten een haalbare afsluiting van hun voortgezet onderwijs te bieden. Zo konden kandidaten hun centrale examens spreiden over twee tijdvakken, een extra herkansing doen en een vak, niet zijnde een kernvak, wegstrepen bij de uitslagbepaling.

Bij elk vak afzonderlijk zou de prestatie-eis zo goed mogelijk gehandhaafd blijven. Vooraf was echter al duidelijk dat het technisch niet mogelijk was om op de reguliere wijze te normeren. Hierdoor was het mogelijk dat de hoogte van de cijfers niet precies de vaardigheid van de kandidaten zou reflecteren. De minister heeft het CvTE verzocht om na afloop van de examens onderzoek te laten uitvoeren naar de daadwerkelijke vaardigheid van de kandidaten bij de centrale examens in het voortgezet onderwijs. Dit onderzoek is in opdracht van het CvTE door Stichting Cito uitgevoerd en dit rapport is het resultaat van dat onderzoek.

Om over jaren heen de prestatie-eisen zo goed mogelijk te handhaven, zijn extra onderzoeken nodig, de zogenaamde normhandhavingsonderzoeken. Deze onderzoeken geven inzicht in de ontwikkeling van de vaardigheid van cohorten kandidaten uit verschillende jaren. De uitkomsten van deze onderzoeken kennen een bepaalde mate van onzekerheid, die dit jaar als gevolg van de omstandigheden nog groter was dan in andere jaren. Deze onzekerheid wordt verkleind wanneer meerdere onderzoeken worden samen genomen. In dit rapport worden daarom alleen uitspraken gedaan over de vaardigheidsontwikkeling bij groepen van centrale examens. In dit onderzoek zijn twee manieren van clustering toegepast. Enerzijds zijn centrale examens geclusterd die inhoudelijk gerelateerd zijn, en anderzijds zijn centrale examens geclusterd per onderwijsniveau.

In dit onderzoek is de gemiddelde vaardigheid van 2021 vergeleken met de gemiddelde vaardigheid voor de corona-pandemie (periode 2015-2019, met nadruk op 2019). Bij clustering van de centrale examens op onderwijsniveau is er een daling van de gemiddelde vaardigheid te zien bij havo, vmbo-gl/tl en vmbo-bb. Bij vwo is niet zeker dat de gemiddelde vaardigheid is gedaald, bij vmbo-kb lijkt er geen wijziging in gemiddelde vaardigheid. Bij Engels is sprake van een stijging met 0,1 tot 0,2 cijferpunt. De overige inhoudelijke clusters van centrale examens laten een gemiddelde vaardigheidsdaling zien in de orde grootte van 0,2 tot 0,5 cijferpunt. Merk op dat dit gemiddelde ontwikkelingen beschrijft. Binnen de clusters van centrale examens bestond nog spreiding in vaardigheidsontwikkeling tussen de examens.

Naast de hoofdvraag van de vaardigheidsontwikkeling is een aantal deelvragen onderzocht. Door herkansingen zou het gemiddelde cijfer van een cluster van centrale examens als geheel nauwelijks stijgen. De daling in vaardigheid bij met name het cluster 'wiskunde' zou deels verklaard kunnen worden door de hoeveelheid zeer lage scores. Gezien de meestal beperkte percentages eerstekansers in het tweede tijdvak zijn de effecten daarvan op de vaardigheids-schattingen per niveau of inhoudelijk cluster van examens klein. De relatieve verschillen tussen scholen in vaardigheid zijn gemiddeld niet toegenomen.

Al met al toont dit onderzoek, met uitzondering van Engels, gemiddelde dalingen in vaardigheid. Deze dalingen mogen volgens statistische criteria 'lichte dalingen' genoemd worden. Ook in verhouding tot de jaarlijkse fluctuaties in gemiddelde vaardigheid per examen in pre-corona-omstandigheden zijn het geen grote dalingen. Dit is in lijn met bevindingen in ander nationaal en internationaal onderzoek dat leerlingen door de schoolsluitingen een leerachterstand hebben opgelopen.

1 Inleiding

1 Inleiding

1.1 Aanleiding

Begin 2020 werd Nederland, net zoals de rest van de wereld, getroffen door de coronapandemie. Het Nederlands onderwijs kwam in een ongekende situatie terecht. Op 15 maart 2020 sloten de scholen in Nederland en op 24 maart kwam het bericht dat de centrale examens in het voortgezet onderwijs in 2020 geen doorgang zouden vinden. Ook in schooljaar 2020-2021 werd er geen onderwijs gegeven op de reguliere manier. Online onderwijs, maatregelen om besmetting binnen de school te voorkomen en slechts gedeeltelijke fysieke aanwezigheid waren gebruikelijk. De centrale examens van 2021 vonden wel plaats, met uitzondering van de centraal schriftelijke en praktische examens (cspé's) in het vmbo, die als schoolexamen konden worden afgenomen. Daarbij werden er meerdere maatregelen genomen om kandidaten een haalbare afsluiting van hun voortgezet onderwijs te bieden. Zo konden kandidaten hun centrale examens spreiden over twee tijdvakken, een extra herkansing doen en een vak, zijnde een niet-kernvak, wegstrepen bij de uitslagbepaling. Details over de context van de centrale examens staan in Appendix I.

De verwachting was dat de vaardigheid van de eindexamenkandidaten van 2021 zou verschillen van die van de kandidaten die in de jaren voor de corona-pandemie eindexamen deden. De gebruikelijke normeringswijze van de centrale examens kon niet zonder aanpassingen gehanteerd worden. Daarvoor waren drie hoofdredenen:

- 1 Het was onduidelijk of wel een representatief deel van de kandidaten in het eerste tijdvak zou deelnemen.
- 2 Het was niet duidelijk of specifieke onderdelen in de voorbereiding wel evenveel aandacht hadden als in voorgaande jaren.
- 3 Het was onduidelijk of de centrale examens waarbij voorheen normhandavingsonderzoek werd gedaan in 2021 wel representatief zouden zijn voor alle andere centrale examens.

Vanwege deze redenen was een aangepaste systematiek voor de normering noodzakelijk. Bij deze aangepaste systematiek kon niet worden gegarandeerd dat de prestatie-eisen voor een voldoende (hierna: 'de lat') gelijk zouden blijven aan de eisen in voorgaande jaren. Daarmee kon ook niet worden gegarandeerd dat de cijfers zonder meer te vergelijken zijn met die van voorgaande jaren, ofwel er kon niet worden gegarandeerd dat de cijfers de vaardigheid van de populatie getrouw zouden weergeven. De minister van Basis- en Voortgezet Onderwijs en Media heeft het CvTE verzocht om na afloop van het eindexamen 2021 onderzoek te laten uitvoeren naar de daadwerkelijke vaardigheid van de examenkandidaten bij de centrale examens in het voortgezet onderwijs. Het CvTE heeft vervolgens Stichting Cito opdracht gegeven om dit onderzoek uit te voeren. Dit rapport is het resultaat van dat onderzoek. Het brengt de vaardigheid van de examenkandidaten bij de centrale examens van 2021 in beeld, in vergelijking met de jaren 2015-2019.

1.2 Normering 2021

De normering van de centrale examens kende een andere systematiek¹ dan in voorgaande jaren². Er werd meer informatie opgevraagd bij docenten dan gebruikelijk. Docenten werd bijvoorbeeld gevraagd of hun kandidaten, die deelnamen aan het eerste tijdvak, qua vaardigheid gelijk waren aan de groep die in het tweede tijdvak zou gaan deelnemen, en daarmee representatief waren voor al hun kandidaten. Met behulp van deze antwoorden werd een representatieve subpopulatie van scholen gevormd. Een relatieve N-term werd bepaald waarbij de representatieve scholen gemiddeld een gelijk cijfer zouden krijgen als in voorgaande jaren.

Wanneer alleen deze relatieve gegevens zouden worden gebruikt om te normeren, dan zou de lat niet automatisch op dezelfde hoogte liggen als in voorgaande jaren. De examenkandidaten van 2021 moesten wel allemaal over dezelfde lat springen om een voldoende te halen. Om te zorgen dat de lat ook in 2021 zoveel mogelijk op dezelfde hoogte lag als in voorgaande jaren, werd docenten gevraagd om ook een oordeel over de moeilijkheidsgraad van het examen te geven. Daarnaast werd de gebruikelijke moeilijkheid van het examen van dat vak, de zogenaamde historische N-term, beschouwd. Het docent-oordeel en de historische N-term vormden zo twee instrumenten om te zorgen dat de lat zoveel mogelijk op dezelfde hoogte zou komen te liggen als in de jaren voor 2020.

1.3 Onderzoeksvragen

Stichting Cito verricht jaarlijks het zogenaamde normhandhavingsonderzoek. In dit onderzoek worden de verschillen in moeilijkheidsgraad van de examens uit opeenvolgende jaren en de vaardigheidsontwikkelingen van de examenpopulatie in kaart gebracht. Deze gegevens zijn in dit onderzoek gebruikt om in alle vijf onderwijsniveaus de vaardigheidsontwikkeling te bestuderen bij minimaal zes centrale examens per niveau.

De hoofdvraag van dit onderzoek is hoe de vaardigheid van de examenkandidaten in 2021 zich verhoudt tot de vaardigheid van de examenpopulaties voor de corona-pandemie.

Om deze vraag goed te kunnen beantwoorden, zijn de volgende deelvragen ook onderzocht:

- 1 In welke mate is er gebruik gemaakt van herkansingen, in vergelijking met voorgaande jaren? En welke invloed had dit op de vaardigheidsschatting? De gemiddelde vaardigheid van kandidaten, en daarmee hun cijfer, kan na herkansing anders uitvallen. Herkansingen kunnen in 2021 een andere invloed op de uiteindelijke vaardigheid van de populatie hebben gehad dan in voorgaande jaren, vanwege de verruimde mogelijkheden daartoe in 2021.
- 2 In welke mate is er sprake van extreem lage scores, in vergelijking met voorgaande jaren? En in hoeverre heeft dit de gemiddelde vaardigheid beïnvloed? De gemiddelde vaardigheid van examenkandidaten kan vertekend zijn door kandidaten met extreem lage scores. Er zijn meerdere mogelijke verklaringen voor een lage score: kandidaten zijn bijvoorbeeld onvoldoende voorbereid op het examen, zij kunnen een black-out tijdens het examen hebben gekregen, of zij hebben geen serieuze poging gedaan met onderpresteren tot gevolg. Omdat kandidaten in 2021 de mogelijkheid kregen om één niet-kernvak weg te strepen bij de uitslagbepaling, zou onderpresteren in 2021 vaker voor kunnen komen. De geschatte vaardigheid is in die gevallen een onderschatting van de werkelijke vaardigheid.

1 Normering centrale examens 2021 - Examenblad

2 N-term - Examenblad 2019

- 3 In hoeverre was de schatting van de vaardigheid van de kandidaten in het eerste tijdvak representatief voor die van alle kandidaten? De gemiddelde vaardigheid wordt berekend op de kandidaten in het eerste tijdvak. Kandidaten konden in 2021 ook in het tweede tijdvak voor het eerst deelnemen aan een examen.
- 4 Is de vaardigheidsontwikkeling gelijk voor alle scholen? Of zijn de verschillen tussen scholen relatief groter geworden door de corona-pandemie?

1.4 Indeling van dit rapport

In hoofdstuk 2 zullen de methoden worden toegelicht waarmee de onderzoeksvragen beantwoord worden. Daarna volgen in hoofdstuk 3 de resultaten van de vaardigheids-schattingen en van de berekeningen naar aanleiding van de deelvragen. In de conclusie (hoofdstuk 4) worden de resultaten samengevat en in hoofdstuk 5 worden de resultaten van context voorzien en vergeleken met ander onderzoek naar het effect van de maatregelen ten gevolge van de corona-pandemie op onderwijsprestaties.

Het rapport kent twee appendices. Eén met een uitgebreide beschrijving van de maatregelen naar aanleiding van de corona-pandemie in het Nederlandse voortgezet onderwijs. Dit biedt een context waarbinnen de centrale examens zijn afgenomen. De tweede appendix gaat in meer detail in op de gebruikte methode om vaardigheidsontwikkeling te schatten.

2 Methoden van onderzoek

2 Methoden van onderzoek

2.1 Gegevens voor het onderzoek

Om de vaardigheid van de examenpopulatie 2021 te onderzoeken, worden de gegevens gebruikt die Stichting Cito jaarlijks verzamelt voor het normhandavingsonderzoek. Het normeringsonderzoek is een schakel in het hele normeringsproces³ dat jaarlijks wordt uitgevoerd onder leiding van het CvTE. De gegevens betreffen de scores op de centraal schriftelijke examens die door docenten zijn ingevoerd in Wolf (uitwisselprogramma voor examengegevens⁴), de scores op de digitale centrale examens in vmbo-bb/kb die afkomstig zijn uit Facet (een digitaal systeem voor het afnemen van toetsen en examens⁵), en gegevens die vooraf en tijdens de examencampagne worden verzameld om de moeilijkheidsgraad van centrale examens te onderzoeken. De gegevens van centrale examens met minder dan 50 kandidaten zijn niet meegenomen in het onderzoek. Ook de gegevens van de centraal schriftelijke examens in vmbo-bb en vmbo-kb zijn niet meegenomen, aangezien deze examens eveneens door zeer weinig kandidaten zijn gemaakt.

2.2 Schatting van vaardigheidsontwikkeling

Om de hoofdvraag te beantwoorden is voor clusters van centrale examens, die deel uitmaken van het normhandavingsonderzoek, nagegaan of de gemiddelde vaardigheid van de examenpopulatie in 2021 verschilt van de gemiddelde vaardigheid van de examenpopulaties voor de corona-pandemie. Er zijn drie methodes gehanteerd om het verschil in vaardigheid van de examenkandidaten uit opeenvolgende jaren te onderzoeken, waarbij rekening wordt gehouden met het verschil in moeilijkheid tussen centrale examens⁶:

- Anchor in Package (AIP): deze methode wordt ingezet bij de digitale centrale examens vmbo-bb en kb. Als kandidaten in verschillende jaren dezelfde set opgaven voorgelegd krijgen, dan is een verschil in prestatie op deze set een gevolg van een verschil in vaardigheid van de kandidaten uit de verschillende jaren.
- Pretest en posttest: volgens deze methodes krijgt een groep leerlingen in één schoolexamen zowel opgaven uit het actuele centrale examen voorgelegd als opgaven uit een ankerexamen. Hierdoor kan de moeilijkheidsgraad van de centrale examens van verschillende jaren vergeleken worden. Een vergelijking van de prestaties van examenpopulaties op de verschillende centrale examens, gecorrigeerd voor een verschil in moeilijkheidsgraad, levert informatie over de vaardigheidsontwikkeling.

3 College voor Toetsen en Examens (30 november 2016). De normeringssystematiek van de Centrale Examens VO. De normeringssystematiek van de centrale examens vo (examenblad.nl), College voor Toetsen en Examens (geen datum). Examenblad.nl: Normering centrale examens.

4 Cito (geen datum). Tools en informatie voor docenten: Wolf.

5 College voor Toetsen en Examens (2017). Examenblad.nl: Wat is Facet?

6 Van der Molen et al. (2019). Vaardigheidsontwikkelingen volgens PISA en examens. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.

- Standaardbepaling: hierbij schatten experts (docenten) de moeilijkheid van de vragen van twee centrale examens. Ook daarmee kan vergeleken worden hoe de moeilijkheidsgraad van deze centrale examens zich verhoudt. Door de geschatte moeilijkheidsgraad te combineren met de scores op het centrale examen kan geschat worden hoe vaardig de examenpopulatie is.

De drie methodes zijn niet even nauwkeurig⁷. Anchor in Package (AIP) levert de meest nauwkeurige schatting op, en standaardbepaling is het minst nauwkeurig. Helaas kan AIP slechts beperkt ingezet worden en zijn ook pre- en posttesten niet altijd mogelijk. De AIP-, pre- en posttest-methoden werken niet bij regelmatige wisselingen van onderwerpen in het examenprogramma, of bij de introductie van een nieuw examenprogramma. Ook kunnen deze methoden niet toegepast worden bij centrale examens met weinig examenkandidaten, zoals bijvoorbeeld aardrijkskunde en geschiedenis in vmbo-bb en vmbo-kb. Daarom zijn bij een aantal examens standaardbepalingen uitgevoerd in dit onderzoek. In Tabel 1 is te zien welke methoden zijn ingezet bij de verschillende examens. Tevens is te zien welke vakken in dit rapport zijn samengenomen in een cluster. Bij de vakken die niet in dit overzicht voorkomen, werd op geen enkel niveau één van deze methodes toegepast.

⁷ Van der Molen et al. (2019). Vaardigheidsontwikkelingen volgens PISA en examens. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.

Tabel 1 Gehanteerde methodes om vaardigheids- en moeilijkheidsverschillen te onderzoeken in 2021

Vak	Cluster	BB-CBT	KB-CBT	GL/TL	HAVO	VWO
Nederlands	Nederlands	AIP		sbep	sbep	sbep
Duits	Moderne vreemde talen (excl. Engels)		AIP		pre/posttest	pre/posttest
Frans	Moderne vreemde talen (excl. Engels)			pre/posttest	pre/posttest	pre/posttest
Engels	Engels	AIP	AIP	pre/posttest	pre/posttest	pre/posttest
Kunst/beeldend	Mens & Maatschappij			sbep	sbep	sbep
Aardrijkskunde	Mens & Maatschappij				sbep	sbep
Geschiedenis	Mens & Maatschappij			sbep		
Economie	Mens & Maatschappij	AIP	AIP	pre/posttest	sbep	sbep
Scheikunde/nask2	Exact			sbep	sbep	pre/posttest
Natuurkunde/nask1	Exact	AIP	AIP	sbep	pre/posttest	
Biologie	Exact	AIP	AIP	pre/posttest	pre/posttest	pre/posttest
Wiskunde A	Wiskunde				sbep	sbep
Wiskunde (B)*	Wiskunde	AIP	AIP	pre/posttest	pre/posttest	sbep

Opmerking: AIP = Anchor in Package; sbep = standaardbepaling

*In vmbo bestaat het vak wiskunde zonder letter, in havo en vwo zowel wiskunde A als wiskunde B. In vwo bestaat ook wiskunde C, dit vak is niet meegenomen in deze tabel.

De ontwikkeling in vaardigheid bij de examenpopulatie van 2021 wordt geschat door prestaties op de centrale examens van 2021 te vergelijken met die op eerdere centrale examens rekening houdend met verschillen in moeilijkheid. Om dezelfde redenen waarom het niet mogelijk was om de gangbare normeringssystematiek te hanteren in 2021, is het niet mogelijk een schatting van de vaardigheidsontwikkeling te doen voor ieder centraal examen afzonderlijk. De schatting van vaardigheid bevat onzekerheden. Door de informatie van meerdere centrale examens te combineren wordt deze onzekerheid verkleind en is het wél mogelijk een schatting van de vaardigheid te geven. In dit onderzoek zijn twee manieren van clustering toegepast. Enerzijds zijn centrale examens geclusterd die inhoudelijk gerelateerd zijn, en anderzijds zijn examens geclusterd per onderwijsniveau. Meer details omtrent deze berekeningen staan in Appendix II.

2.3 Berekeningen deelvragen

De deelvragen worden besproken aan de hand van de scores die door de eerste corrector zijn doorgegeven. De bijbehorende cijfers zijn gebaseerd op de vastgestelde N-termen per centraal examen. Bij de centrale examens krijgen kandidaten scorepunten voor correcte antwoorden. Deze scorepunten worden omgezet naar cijfers. De basisregels voor het omzetten van scores in cijfers zijn vastgelegd in de 'Regeling omzetting scores in cijfers VO'⁸.

In reguliere examenjaren leggen kandidaten een centraal schriftelijk examen af in het eerste tijdvak, het tweede tijdvak is voor herkansingen. In 2021 mochten kandidaten kiezen in welk tijdvak ze een examen voor de eerste keer afleggen. In dit rapport heten kandidaten die een examen voor de eerste keer maken eerstekansers. Kandidaten die een herexamen maken, heten herkansers. Bij de digitale centrale examens (vaak aangeduid met CBT-examens, waarbij CBT staat voor Computer Based Testing) in vmbo-bb en vmbo-kb zijn de afnamegegevens tot 10 juni meegenomen. Deze worden allen als eerstekansers beschouwd. Het is dezelfde steekproef van afnames die ook bij de normering op 15 juni werd beschouwd. De afnameperiode van de digitale centrale examens liep door tot en met 23 juli 2021.

2.3.1 Deelvraag 1

Voor een antwoord op de eerste deelvraag, naar de omvang en het effect van herkansingen, worden voor alle centrale examens in vmbo-gl/tl, havo en vwo de herkansingspercentages vergeleken met die in 2019. Dit betreft dan per centraal examen het percentage van alle herkansers ten opzichte van alle eerstekansers. Bij de CBT-examens waren herkansers en de daarbij horende eerste afname moeilijk eenduidig te identificeren, omdat de afnames flexibel konden plaatsvinden. Daarom is deze deelvraag niet bestudeerd in vmbo-bb en vmbo-kb. In vmbo-gl/tl, havo en vwo is een grove inschatting gedaan van de toename in cijferpunt van de herkansende kandidaten om de invloed van de herkansingen op de vaardigheid te schetsen.

De herkansingspercentages zijn steeds berekend ten opzichte van het aantal kandidaten dat het vak als eerstekanser maakte, en niet ten opzichte van het totaal aantal kandidaten per onderwijsniveau. Centrale examens met relatief weinig kandidaten kunnen daarmee toch een hoog percentage herkansers hebben. Schommelingen tussen 2021 en 2019 in aantallen eerstekansers bij een vak hebben met deze berekeningswijze geen invloed op de resultaten.

2.3.2 Deelvraag 2

Om de tweede deelvraag naar de invloed van extreem lage scores op de vaardigheidsschatting te beantwoorden worden voor alle vakken uit het centraal examen 2021 van de vijf onderwijsniveaus de percentages lage scores vergeleken met die in 2019. Een lage score wordt daarbij gedefinieerd als een score lager dan 25% van de maximale score. Het percentage werd weer per centraal examen uitgerekend, waardoor het totaal aantal kandidaten dat aan het examen deelnam geen invloed had op de resultaten.

Een gemiddelde vaardigheid wordt beïnvloed door de verdeling van vaardigheid in de populatie. Als deze disproportioneel meer lage scores krijgt, zakt de gemiddelde vaardigheid. Scores die lager zijn dan 25% van de maximale score, resulteren in cijfers lager dan een 4. Als er in 2021

⁸ Stcrt, 2016, 4817; Stcrt, 2021,18168: wetten.nl - Regeling - Regeling omzetting scores in cijfers VO - BWBR0037590 (overheid.nl)

andere percentages extreem lage scores voorkomen dan in 2019, kan hiermee de geschatte vaardigheidsontwikkeling tussen 2019 en 2021 deels statistisch verklaard worden.

2.3.3 Deelvraag 3

Om de derde deelvraag te beantwoorden, of de vaardigheid van kandidaten in het eerste tijdvak representatief is voor die van alle kandidaten, wordt gekeken of er sprake is van selectieve deelname in het eerste tijdvak. Enerzijds wordt gekeken hoeveel kandidaten centrale examens hebben uitgesteld tot het tweede tijdvak, anderzijds wat de gemiddelde cijfers in het eerste tijdvak waren in vergelijking met de gemiddelde cijfers van eerstekansers in het tweede tijdvak. Ook dit onderwerp is niet voor vmbo-bb en vmbo-kb bestudeerd, omdat herkansers bij de CBT-examens moeilijk te identificeren waren.

2.3.4 Deelvraag 4

De vierde deelvraag is of de vaardigheidsontwikkeling gelijk is voor alle scholen. Om deze vraag te kunnen beantwoorden wordt per centraal examen een meerniveau-model⁹ toegepast. In een meerniveau-model zijn de scholen en de kandidaten daarbinnen de twee analyseniveaus. Per analyseniveau kijkt een meerniveau-model hoe groot de spreiding in cijfers is. Dus hoeveel gemiddelde cijfers van scholen onderling verschillen (schoolniveau), en hoeveel de cijfers van kandidaten binnen scholen van elkaar verschillen (leerlingniveau). De totale spreiding in de cijfers is een optelsom van de spreiding tussen scholen en de spreiding binnen scholen.

De intraklasse correlatie (ICC) is een gangbare maat voor de relatieve omvang van de verschillen tussen scholen ten opzichte van de totale spreiding¹⁰. Deze ICC zegt iets over het landelijke beeld van de invloed van scholen op de cijfers van hun kandidaten, en kun je berekenen met een meerniveau-model. Om de vierde onderzoeksvraag te beantwoorden zijn voor de 45 examens in Tabel 1 ICC's berekend van zowel 2019 als 2021. Als de ICC van een vak tussen deze twee jaren gelijk zou zijn, dan is de vaardigheidsontwikkeling tussen 2019 en 2021 min of meer gelijk geweest voor alle scholen.

9 Zie bijvoorbeeld Snijders, T.A.B. & Bosker, R.J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (sec. ed.). London, Sage.

10 R-library lme4 is gebruikt om het meerniveau-model te schatten; de R-library performance om de icc te schatten.

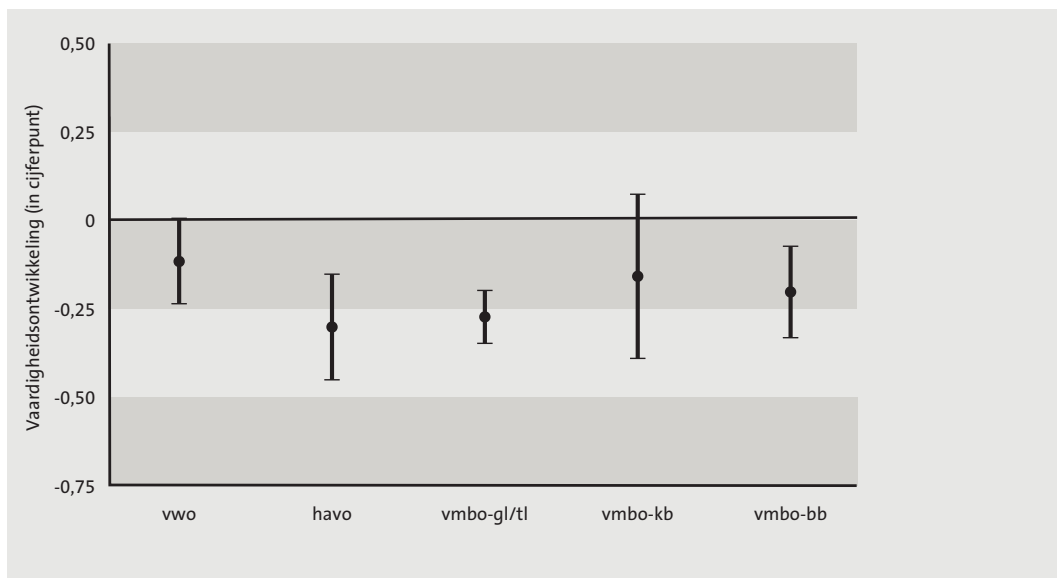
3 Resultaten

3 Resultaten

3.1 Vaardigheidsontwikkeling

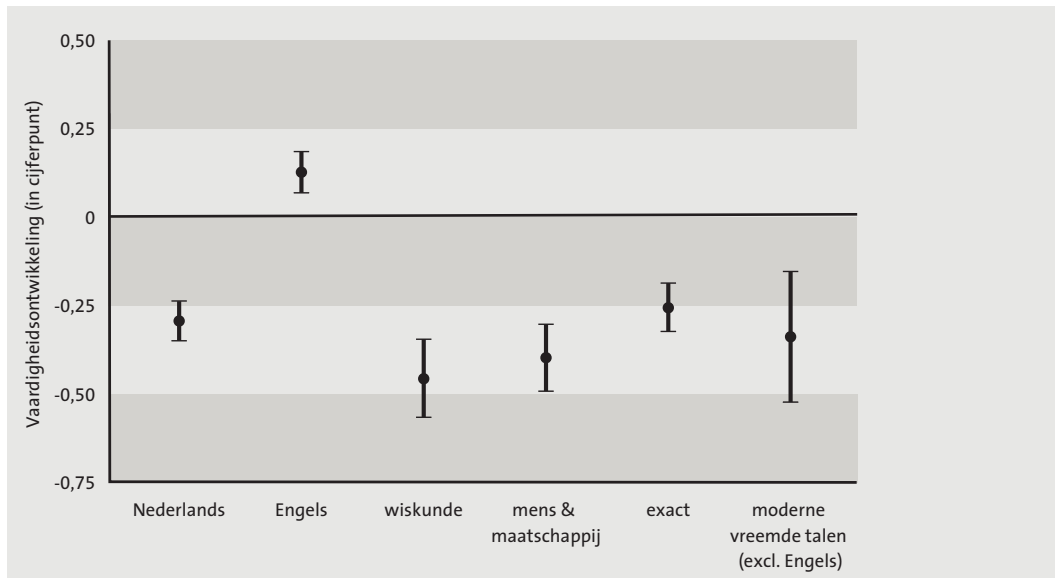
Figuur 1 en Figuur 2 tonen op de horizontale as clusters van centrale examens, zoals in Tabel 1 gespecificeerd. Per cluster van centrale examens geeft de verticale lijn het 90%-betrouwbaarheidsinterval weer waarbinnen de gemiddelde vaardigheidsontwikkeling van dat cluster valt. Als de vaardigheid in 2021 over het algemeen lager was dan voor de coronapandemie, dan ligt het volledige interval onder 0. De grootte van de vaardigheidsverschillen is uitgedrukt in cijferpunten. Dus een verschil van -0.3 betekent dat de 2021-populatie gemiddeld een vaardigheid had, die op de oude examens had geresulteerd in cijfers die gemiddeld 0,3 cijferpunt lager zouden zijn dan die van de pre-corona-populatie. De lengte van een interval zegt enerzijds iets over de verschillen in vaardigheidsontwikkeling van de examens binnen een cluster. Hoe meer examens daarin verschillen binnen een cluster, hoe groter het interval. Anderzijds zegt het iets over de onzekerheid van de onderliggende schattingen per examen: bij gelijke vaardigheidsontwikkeling van alle onderliggende examens verwacht je bij een set standaardsettingen een groter interval dan bij veel AIP-examens.

Bij clustering van de centrale examens op onderwijsniveau is er bij alle onderwijsniveaus een daling (Figuur 1), maar in vmbo-kb is de gemiddelde vaardigheid van de examenkandidaten in het eerste tijdvak in 2021 niet significant lager dan in de pre-corona-jaren. Bij vwo is de daling op het randje van significantie, omdat het 90%-betrouwbaarheidsinterval de waarde 0 net omvat. Bij vmbo-kb zijn de vaardigheidsontwikkelingen bij de onderliggende examens vrij nauwkeurig geschat omdat het alle AIP-examens betrof. Het grote interval duidt erop dat de vaardigheidsontwikkelingen divers waren tussen de vmbo-kb-examens.



Figuur 1 Geschat vaardigheidsverschil tussen 2021 en voor corona per onderwijsniveau

Ook bij clustering van de centrale examens op inhoudelijke overeenkomst is de gemiddelde vaardigheid in alle clusters gedaald ten opzichte van voor corona met uitzondering van Engels, waar de vaardigheid is toegenomen (Figuur 2).



Figuur 2 Geschat vaardigheidsverschil tussen 2021 voor corona per cluster

In dit onderzoek kijken we met name naar de gemiddelde vaardigheidsontwikkeling. Een clustering van centrale examens helpt om te kijken of deze vaardigheidsontwikkeling op alle onderwijsniveaus of inhoudelijke clusters gelijk is. De verschillen tussen de clusters willen niet zeggen dat de examens binnen één cluster allemaal dezelfde vaardigheidsontwikkeling laten zien. Er is een diversiteit in vaardigheidsontwikkeling bij de examens binnen de clusters, al is dit deels beïnvloed door de onzekerheid waarmee vaardigheidsontwikkelingen geschat kunnen worden. In de analyses per deelvraag, kijken we wat gericht naar de verschillen tussen centrale examens.

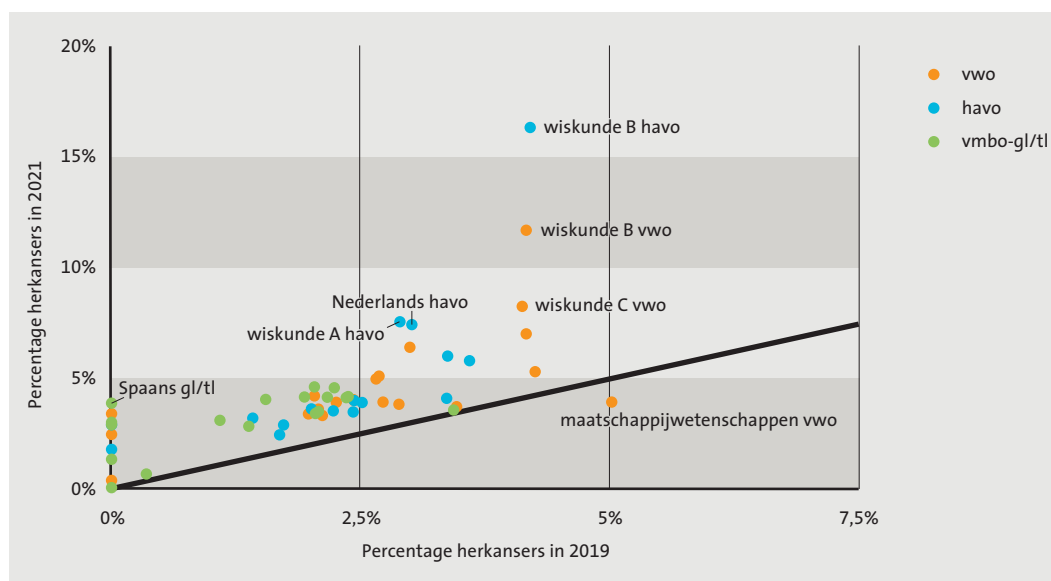
3.2 Resultaten per deelvraag

3.2.1 Herkansers in 2019 en 2021

Na herkansing was de eindvaardigheid van kandidaten veelal hoger dan daarvoor. Gemiddeld hebben herkansende examenkandidaten in vmbo-gl/tl, havo en vwo zich +0,6 cijferpunt verbeterd bij herkansing. Dit was gemiddeld over alle herkansende kandidaten van deze onderwijsniveaus en alle vakken. In Figuur 3 staat op de horizontale as het percentage herkansers per centraal examen in 2019, op de verticale as het percentage in 2021. De schuine lijn geeft aan wanneer het percentage herkansers in 2021 gelijk was aan dat in 2019. Het eerste dat opvalt in Figuur 3 is dat slechts bij één centraal examen (maatschappijwetenschappen in VWO) het percentage herkansers in 2021 lager is dan in 2019. Gemiddeld is het percentage herkansers per centraal examen in 2021 twee procentpunten hoger dan in 2019, een stijging van gemiddeld 2% naar 4%. Er zijn echter behoorlijke uitschieters als het gaat om het percentage herkansers in 2021. Terwijl in 2019 het percentage tussen de 0 en 5% lag, ligt het percentage in 2021 tussen de 0 en 16%.

In vmbo-gl/tl liggen de percentages herkansers in 2021 weliswaar iets hoger dan in 2019, maar nog steeds zijn dit relatief lage percentages. Bij Spaans is de stijging van het percentage het hoogst: in 2021 4% herkansers terwijl dit in 2019 nog 0% was.

Bij havo en vwo zijn er vooral bij de wiskunde-examens hoge percentages herkansers in 2021. Wiskunde B in havo kent het hoogste percentage herkansers, namelijk 16%, dat is een toename van 12% ten opzichte van 2019.



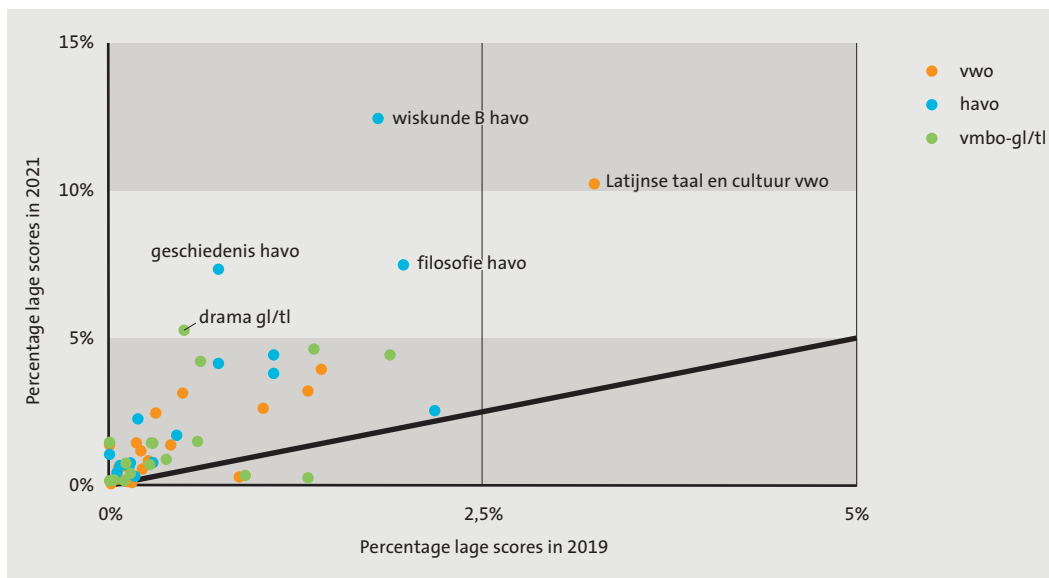
Figuur 3 Percentage herkansers per centraal examen in 2019 en 2021

De invloed van deze cijfers op de vaardigheidsschattingen in Figuur 1 en Figuur 2 is niet direct te voorzien. De vaardigheid van de examenpopulatie 2021 zou na herkansing gemiddeld zo'n $4\% \cdot 0,6$ cijferpunt = 0,024 cijferpunt hoger zijn. Omdat de vaardigheid van de examenpopulatie 2021 geschat wordt ten opzichte van die in de jaren voor de corona-pandemie, moet ook het effect van herkansingen in die jaren worden meegenomen. De vaardigheid van de populatie van 2019 zou ook stijgen na herkansing, gemiddeld zo'n $2\% \cdot 0,6$ cijferpunt = 0,012 cijferpunt. Het verschil met het herkansingseffect in 2021 is 0,012 cijferpunt. Merk op dat dit veel kleiner is dan 0,1 cijferpunt. Daarmee zou het vaardigheidsverschil zoals gepresenteerd in Figuur 1 en Figuur 2 niet veel anders worden wanneer het effect van de herkansingen meegerekend zou worden.

3.2.2 Kandidaten met lage scores

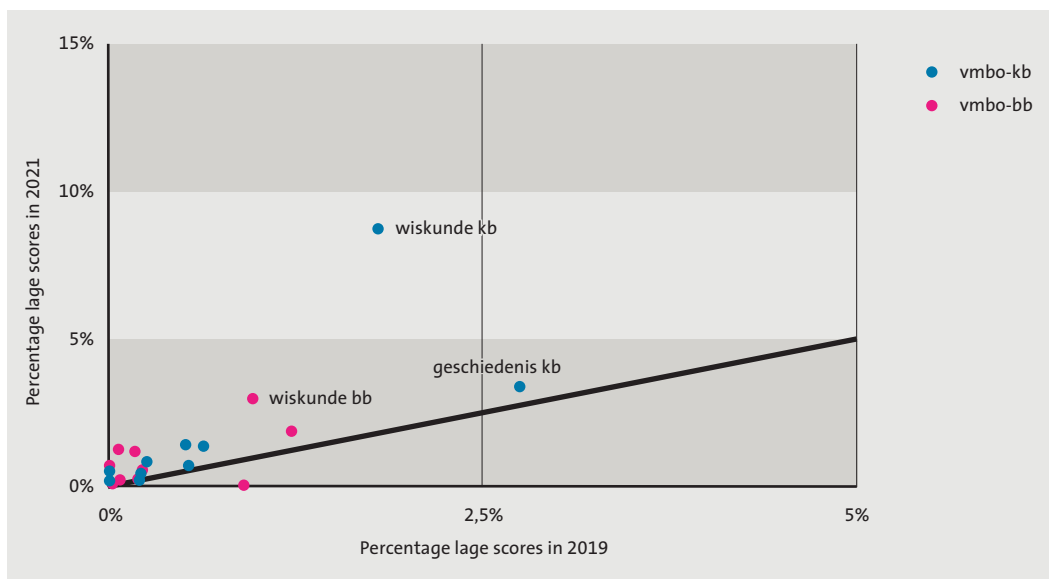
In Figuur 4 en Figuur 5 staan per centraal examen de percentages eerstekansers met een lage score, dat wil zeggen een score lager dan 25% van de maximale score. Op de horizontale as staan de percentages kandidaten in 2019 en op de verticale as de percentages in 2021. Bijvoorbeeld het examen wiskunde B in havo kende in 2019 een percentage lage scores van 2% en in 2021 van 12%. De schuine lijn geeft aan wanneer het percentage in 2021 gelijk was aan dat in 2019. De punten in Figuur 4 en Figuur 5 liggen grotendeels aan de linkerkant van de grafiek. Dit betekent dat lage scores voor de corona-pandemie maar beperkt voorkwamen bij de centrale examens, gemiddeld bij minder dan 1% van de kandidaten, en hooguit 3% bij een enkel examen.

Bij vrijwel alle centrale examens, in alle onderwijsniveaus, is het percentage kandidaten met een lage score (minder dan 25% van de maximale score) groter dan in 2019 (Figuur 4 en Figuur 5). In vwo toont Latijnse taal en cultuur een veel hoger percentage examenkandidaten met een lage score (10%) dan in 2019 (3%). Bij havo springt, net zoals bij de herkansingen, wiskunde B eruit met 12% kandidaten met een lage score in 2021 (tegenover 2% in 2019), maar ook bij filosofie en geschiedenis was het percentage ruim boven de 5%. In vmbo-gl/tl lag bij drama het percentage kandidaten met een lage score net boven de 5% in 2021.



Figuur 4 Percentage eerstekansers met een lage score per CSE in 2019 en 2021

In vmbo-kb en vmbo-bb vertonen de wiskunde-examens de grootste toename aan kandidaten met een lage score. Wiskunde op kb-niveau is overigens het enige vak dat meer dan 5% lage scores kende bij de digitale examens, namelijk 9%.



Figuur 5 Percentage eerstekansers met een lage score per CBT-examen in 2019 en 2021

Opvallend is dat wiskunde er bij meerdere onderwijsniveaus uitspringt met relatief veel lage scores. De daling in vaardigheid bij het cluster 'wiskunde' zou dus deels verklaard kunnen worden door de hoeveelheid lage scores. Enkele andere genoemde centrale examens, zoals Latijn bij vwo, filosofie en geschiedenis bij havo, en geschiedenis bij vmbo-kb, zijn niet opgenomen in de berekeningen van de vaardigheidsontwikkeling. De vaardigheidsdaling bij deze centrale examens zou dus groter kunnen zijn dan de schattingen in Figuur 1 en Figuur 2 van het betreffende cluster van examens. Deze bevinding sluit aan bij de derde reden waarom er een aangepaste normeringswijze van de centrale examens nodig was. De centrale examens in het normhandhavingsonderzoek waren niet altijd representatief voor alle andere centrale examens. Echter, als er slechts een enkel examen met veel lage scores per cluster zou worden

toegevoegd, dan zou de gemiddelde vaardigheidsontwikkeling van dat cluster niet veel wijzigen; de diversiteit van de vaardigheidsontwikkeling binnen het cluster zou wel toenemen, wat zich zou uiten in een groter betrouwbaarheidsinterval van het cluster.

Als een laagscorende kandidaat het examen in pre-corona-omstandigheden had gemaakt, zou zijn/haar score rond het landelijk gemiddelde kunnen zijn geweest. Dit is al snel een verschil van 2 cijferpunten. Bij enkele centrale examens gaat het soms om 5% meer leerlingen met lage scores ten opzichte van 2019. Ofwel om $(5\% * 2 \text{ cijferpunten} =)$ gemiddeld 0,1 cijferpunt daling die verklaard zou kunnen worden door de lage scores. Als dit gemiddeld bij alle centrale examens in een cluster het geval is (zoals bij wiskunde), dan zou dit effect voor ongeveer 0,1 cijferpunt hebben bijgedragen aan de geschatte vaardigheid van dat cluster.

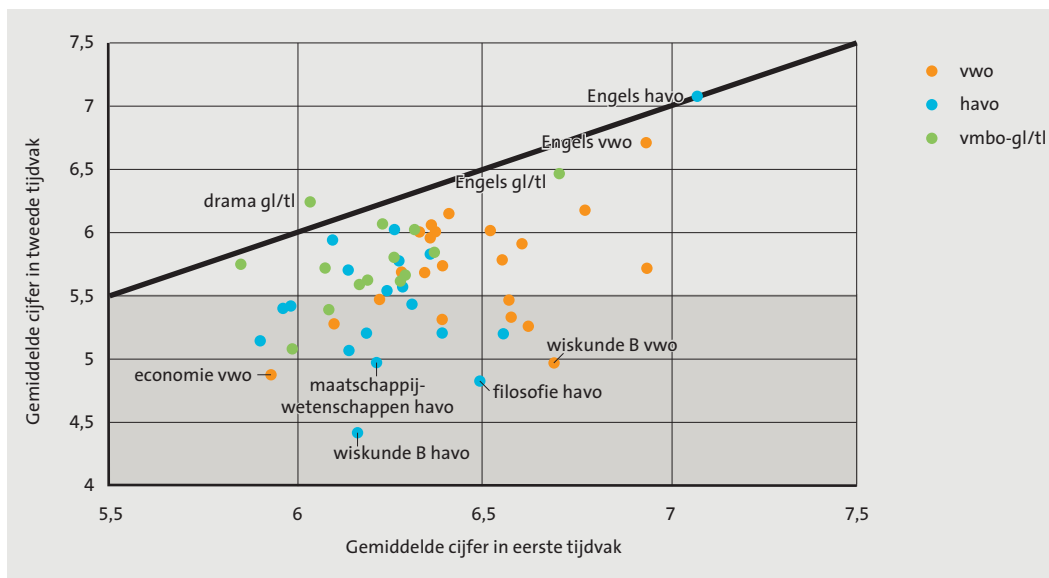
3.2.3 Eerstekansers in tweede tijdvak en hun gemiddelde cijfer

In 2021 hebben de kandidaten van de centraal schriftelijke examens de mogelijkheid gekregen om het examen te spreiden over twee tijdvakken bestaande uit een eerste tijdvak voor eerste afnames en een tweede tijdvak voor eerste afnames en herkansingen. De derde deelvraag was of de kandidaten in het eerste tijdvak wel representatief waren voor alle eerstekansers, aangezien de vaardigheidsontwikkeling geschat is op kandidaten uit het eerste tijdvak. Deze vraag had geen betrekking op de digitale centrale examens in vmbo-bb en vmbo-kb, omdat daarbij niet in tijdvakken afgenomen wordt.

In 2021 deden per centraal examen gemiddeld 5% van de kandidaten voor het eerst examen in het tweede tijdvak. In vmbo-gl/tl hebben examenkandidaten minder gebruik gemaakt van de mogelijkheid om het centraal examen te spreiden. Gemiddeld deed 3% van de vmbo-gl/tl-kandidaten per centraal examen de eerste afname in het tweede tijdvak, terwijl dit in havo gemiddeld 6% per centraal examen was en in vwo 5% per centraal examen. In elk niveau zijn er enkele centrale examens waarbij veel meer leerlingen ervoor hebben gekozen om de eerste afname pas in het tweede tijdvak te doen (zie Tabel 2).

Tabel 2 Percentage eerstekeerders dat het vak voor de eerste keer in het tweede tijdvak aflegt

Vak	% eerstekeerders in het tweede tijdvak 2021		
	VWO	HAVO	VMBO-GL/TL
Nederlands	1,9%	2,7%	1,3%
Engels	2,3%	2,7%	1,4%
Duits	2,2%	2,4%	2,1%
Frans	4,0%	3,1%	2,8%
Spaans	1,7%	3,3%	1,6%
Griekse taal en cultuur	6,8%		
Latijnse taal en cultuur	11,0%		
Aardrijkskunde	4,0%	5,9%	4,0%
Bedrijfseconomie	3,7%	5,7%	
Economie	3,7%	5,4%	4,0%
Filosofie	12,8%	7,6%	
Geschiedenis	7,7%	6,7%	3,8%
Maatschappijwetenschappen	7,9%	4,2%	4,9%
Kunst (algemeen)	3,2%	9,3%	
Beeldende vakken			2,1%
Drama			11,0%
Muziek	7,1%	3,6%	1,6%
Tek., hand., text.	7,3%	6,4%	
Biologie	6,0%	8,2%	4,1%
Natuurkunde/ nask1	6,0%	6,9%	3,9%
Scheikunde/ nask 2	5,3%	4,3%	4,2%
Wiskunde A	3,7%	5,3%	
Wiskunde (B)	3,4%	7,3%	3,8%
Wiskunde C	4,8%		



Figuur 6 Gemiddelde cijfer eerstekansers in het eerste tijdvak en eerstekansers in het tweede tijdvak

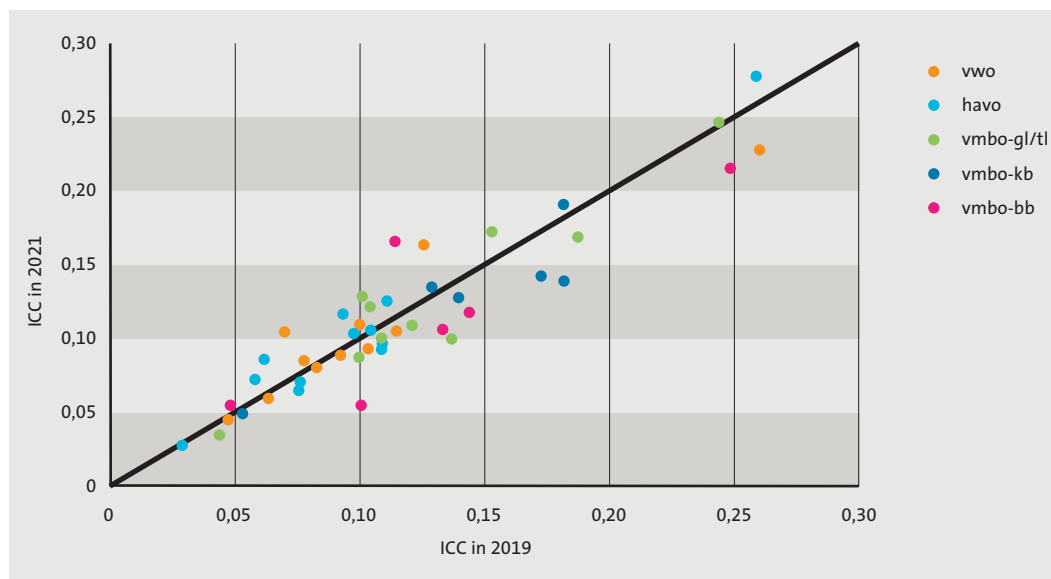
In Figuur 6 is gekeken wat de gemiddelde cijfers van de eerstekansers in het eerste tijdvak (horizontale as) en in het tweede tijdvak (verticale as) waren, per centraal examen met meer dan 150 kandidaten in het eerste tijdvak en minstens 25 eerstekansers in het tweede tijdvak. Alleen bij drama in vmbo-gl/tl behaalden de eerstekansers in het tweede tijdvak gemiddeld een hoger cijfer (6,2) dan de eerstekansers in het eerste tijdvak (6,0). Dit zou kunnen komen doordat één hele school (van de 47 scholen die het vak afnamen), met een groot aantal kandidaten, de deelname in het eerste tijdvak ongeldig heeft moeten verklaren en in het tweede tijdvak weer deel moest nemen, wat de cijfers bij dit vak vertekent. Bij Engels havo behaalden de eerstekansers van het tweede tijdvak gemiddeld een even hoog cijfer als de eerstekansers in het eerste tijdvak. Bij *alle* overige centrale examens op alle drie onderwijsniveaus behaalden de eerstekansers in het tweede tijdvak gemiddeld een lager cijfer dan de eerstekansers in het eerste tijdvak. Figuur 6 toont ook dat in havo en vwo bij een flink aantal centrale examens het gemiddelde cijfer onder de 5,5 (onvoldoende) was. In het tweede tijdvak zijn in vwo 8 van de 22 afgebeelde examens in totaal gemiddeld onvoldoende (36% van de examens), in havo 11 van de 19 (58% van de examens) en in vmbo-gl/tl 2 van de 14 examens (14%).

Er is gestreefd om de examens van 2021 gemiddeld even streng te normeren in het tweede tijdvak als in het eerste tijdvak. Gemiddeld lagen de cijfers van de eerstekansers in het tweede tijdvak 0,7 cijferpunt lager dan van de kandidaten in het eerste tijdvak. Een dergelijke gemiddelde daling in cijfers duidt op een lagere vaardigheid van de eerstekansers in het tweede tijdvak, dan die van de kandidaten in het eerste tijdvak. De schattingen van de vaardigheidsontwikkeling (Figuur 1 en Figuur 2) zijn gebaseerd op de kandidaten in het eerste tijdvak. Gezien de geringe percentages kandidaten die het per centraal examen betreft, is de vaardigheid van de totale populatie echter niet veel lager dan weergegeven in Figuur 1 en Figuur 2. Een rekenvoorbeeld: als 5% van de populatie 0,7 cijferpunt lager scoort, dan valt de gemiddelde vaardigheid $0,05 \cdot 0,7 = 0,035$ cijferpunt lager uit. Dus als de eerstekansers in het tweede tijdvak wel zouden zijn meegenomen in de berekeningen van Figuur 1 en Figuur 2, dan zouden de gemiddelden maar beperkt, veel minder dan 0,1 cijferpunt, lager uitgevallen zijn dan nu het geval is. Bij een enkel vak zou het effect groter kunnen zijn, maar dan moet zowel het percentage eerstekansers in het tweede tijdvak hoog zijn, als het verschil in gemiddeld cijfer tussen deze kandidaten en de kandidaten in het eerste tijdvak. Vanwege de onzekerheid bij een enkel vak, is de vaardigheidsschatting juist gedaan op clusters van centrale examens.

3.2.4 Spreiding tussen scholen

De vierde deelvraag spitste zich toe op de vraag of de vaardigheidsontwikkeling gelijk was op alle scholen. Daartoe werd zowel voor de cijfers van 2021 als voor die van 2019 een maat (ICC) berekend die de relatieve invloed van een school op de cijfers van haar kandidaten aangeeft.

Per centraal examen is de ICC op de cijfers van 2021 en 2019 uitgerekend (Figuur 7). Opvallend is dat er wel verschillen in ICC zijn tussen centrale examens, maar dat gemiddeld genomen er geen verschil is tussen ICC's in 2021 ten opzichte van die in 2019. De vaardigheidsontwikkeling lijkt dus in min of meer gelijke mate bij alle scholen plaats te hebben gevonden, want de relatieve verschillen tussen scholen zijn niet structureel groter of kleiner geworden.



Figuur 7 ICC (intraklasse correlatie) per centraal examen in 2021 en 2019

4 Conclusie

4 Conclusie

De hoofdvraag van het onderzoek was hoe de vaardigheid van de kandidaten in 2021 zich verhoudt tot de vaardigheid van de examenpopulaties voor de corona-pandemie. De vaardigheid in drie van de vijf onderwijsniveaus is gedaald, bij vmbo-kb is deze min of meer constant gebleven en vwo zit op de rand tussen daling en geen verschil. De gemiddelde vaardigheid in Engels is daarentegen gestegen. De vaardigheid voor de clusters exact, mens en maatschappij (inclusief kunst), overige moderne vreemde talen, Nederlands en wiskunde is gedaald. Daarbij vond de grootste gemiddelde daling bij wiskunde plaats (van 0,4 à 0,5 cijferpunt).

De eerste deelvraag betrof de invloed van herkansingen op de gemiddelde vaardigheid van kandidaten. Omdat in 2021 kandidaten twee centrale examens mochten herkansen in plaats van een, was de verwachting dat het aantal herkansers in 2021 hoger zou zijn dan in 2019. Er is inderdaad gemiddeld een kleine toename in het percentage herkansers, met enkele uitschieters. Vooral bij wiskunde in havo en vwo was er een grote toename in het percentage herkansers. Desondanks zou het totale gemiddelde cijfer van een centraal examen door deze herkansingen nauwelijks stijgen. De herkansingen hebben dan ook geen noemenswaardig effect op de geschatte vaardigheid van de populatie.

Een tweede deelvraag om de daling in vaardigheid te duiden bekeek een ander aspect, namelijk het percentage kandidaten met een lage score. De geschatte vaardigheid zou beïnvloed kunnen zijn doordat kandidaten de mogelijkheid kregen om één vak weg te strepen bij de uitslagbepaling. Er was geen informatie beschikbaar over welke kandidaten bij welk examen doelbewust van deze mogelijkheid gebruik wilden maken. Wel kwamen kandidaten met lage scores in 2021 vaker voor dan in 2019, wat de gemiddelde geschatte vaardigheid naar beneden trok. De oorzaak van de toename in lage scores is niet af te leiden uit de scores zelf. Onduidelijk is welk deel van deze daling veroorzaakt is door kandidaten die niet of minder goed hun best hebben gedaan op het centraal examen, en welk deel een gevolg was van minder goede voorbereiding vanwege de corona-pandemie. Bij de kernvakken, zoals wiskunde in havo en vwo, was de wegstreepregeling niet van toepassing en daarom kan een toename in lage scores bij deze centrale examens niet veroorzaakt zijn door deze regeling.

Bij de percentages lage scores waren er enkele uitschieterende centrale examens, zoals wiskunde B, filosofie en geschiedenis in havo, Latijn in vwo en wiskunde in vmbo-kb. Filosofie havo, geschiedenis havo en Latijn zijn niet meegenomen in de schatting van de vaardigheid. De centrale examens van het normhandhavingsonderzoek (zie Tabel 1) zijn daarmee wellicht niet representatief voor alle vakken. Dit betekent dat de geschatte vaardigheidsontwikkeling van een cluster van centrale examens niet zonder meer van toepassing is op de centrale examens die wel onder het cluster zouden vallen, maar niet in het normhandhavingsonderzoek zijn opgenomen.

Per centraal examen kunnen extra lage scores al snel leiden tot een extra daling van de vaardigheid van 0,1 cijferpunt. Als dit bij veel examens in een cluster het geval is (zoals bij wiskunde), dan is het denkbaar dat een substantieel deel van de vaardigheidsdaling veroorzaakt is door lage scores.

In 2021 kregen kandidaten de gelegenheid om het centraal examen te spreiden. Kandidaten konden ervoor kiezen om een centraal examen in het tweede tijdvak te doen in plaats van in het eerste tijdvak. De derde deelvraag haakte hierop aan door te kijken of de kandidaten waarop de vaardigheidsschatting berekend was, wel representatief waren voor alle kandidaten.

Per centraal examen heeft ongeveer 5% van de kandidaten in 2021 gebruik gemaakt van de mogelijkheid om een afname in het tweede tijdvak te doen in plaats van het eerste tijdvak. Dit varieerde van ongeveer 2% bij Engels en Duits (alle drie niveaus) tot 12,8% voor filosofie in vwo. In vmbo-gl/tl hebben examenkandidaten minder gebruik gemaakt van de mogelijkheid om het centraal examen te spreiden dan in havo en vwo. Bij vrijwel alle centrale examens haalden deze kandidaten, die hun eerste afname hadden uitgesteld tot het tweede tijdvak, gemiddeld lagere cijfers dan de kandidaten die al in het eerste tijdvak het centraal examen hadden gedaan. Dit verschil liep bij een aantal examens op tot meer dan 1,5 cijferpunt (wiskunde B in havo en vwo, filosofie in havo). Bij een aantal centrale examens, zoals wiskunde B in havo en vwo, haalden de eerstekansers in het tweede tijdvak gemiddeld zelfs een onvoldoende. Het lijkt erop dat de kandidaten de mogelijkheid tot examenspreiding hebben benut om een vak waarin ze minder goed zijn in het tweede tijdvak te maken. Gezien de meestal beperkte percentages eerstekansers in het tweede tijdvak zijn de effecten daarvan op de gemiddelde vaardigheidsschattingen per onderwijsniveau of inhoudelijk cluster van centrale examens echter ook beperkt (minder dan 0,1 cijferpunt).

De laatste deelvraag betrof de spreiding tussen scholen. De relatieve verschillen tussen de scholen in cijfers zijn gemiddeld niet toegenomen. De scholen hebben dus bij de centrale examens de onderlinge verschillen in schoolgemiddelden per examen niet laten toenemen, wat erop duidt dat de vaardigheidsontwikkeling bij alle scholen in min of meer gelijke mate plaats vond.

De factoren die bekeken zijn, lijken geen noemenswaardige invloed te hebben op de vaardigheidsschattingen in Figuur 1 en Figuur 2. Gemiddeld laat de vaardigheid op havo, vmbo-gl/tl en vmbo-bb een daling laat zien van rond de 0,2 à 0,3 cijferpunt. Bij vmbo-kb is de daling gemiddeld ongeveer 0,15 cijferpunt, maar daarbij is de spreiding tussen de centrale examens binnen dit onderwijsniveau zo groot, dat de gemiddelde daling niet significant is. Bij vwo is het gemiddelde verschil in vaardigheid tussen 2021 en pre-corona ook negatief (-0,1), maar ook hierbij is de spreiding tussen de centrale examens zo groot dat dit gemiddelde net niet significant is. Bij de inhoudelijke clusters van examens is er overall een vaardigheidsdaling tussen de 0,25 tot 0,45 cijferpunt, met uitzondering van Engels waar een vaardigheidsstijging van ongeveer 0,1 cijferpunt te zien is. De invloed van herkansende kandidaten en van eerstekansers in het tweede tijdvak op deze schattingen is beperkt. Doordat enkele examens met veel kandidaten met lage scores niet zijn meegenomen, zou de gemiddelde vaardigheidsdaling bij sommige clusters groter kunnen zijn dan geschat. Scholen zijn niet meer van elkaar gaan verschillen in gemiddelde vaardigheid op het centraal examen dan voor de corona-pandemie.

5 Context bij de resultaten

5 Context bij de resultaten

Met uitzondering van het cluster Engels, zijn bij de andere clusters van centrale examens dalingen in vaardigheid van 0,2 tot 0,45 cijferpunt gevonden. De vraag die rijst, is hoe groot een vaardigheidsdaling van 0,2 tot 0,45 cijferpunt nu eigenlijk is. Het eerdere onderzoek van Stichting Cito naar de vaardigheidsontwikkelingen volgens PISA en centrale examens¹¹ laat zien dat de resultaten op de centrale examens sinds 2015 vrij stabiel zijn, maar wel jaarlijks rond de 0,1 tot 0,2 cijferpunt fluctueren. Het valt op in 2021 dat voor vrijwel alle centrale examens een *daling* te zien is. Dus de fluctuaties zijn in 2021 bijna alle in negatieve richting, en deels groter dan verwacht kon worden op basis van voorgaande jaren.

Een alternatieve manier om de omvang van een daling te beoordelen, is door een vergelijking te maken met de spreiding in cijfers tussen kandidaten per centraal examen voor de coronapandemie. Dit geeft een idee over het aantal leerlingen dat in 2019 per centraal examen lager scoorde dan het gemiddelde in 2021. In 2019 was de gemiddelde standaarddeviatie in cijfers gemiddeld ongeveer 1,1. Daarmee zou een daling van 0,25 tot 0,55 cijferpunt bij een enkel centraal examen een klein effect genoemd kunnen worden, volgens de interpretatie van effectmaten van Cohen¹².

Een lichte daling van de vaardigheid of prestaties is in lijn met bevindingen in ander nationaal en internationaal onderzoek dat leerlingen door de schoolsluitingen een leerachterstand hebben opgelopen. In het najaar 2020 werden de resultaten van verschillende onderzoeken gepubliceerd, waaruit bleek dat in het primair onderwijs leerlingen een leerachterstand hadden opgelopen^{13,14,15,16}. In het onderzoek van Oxford en Leeruniek werd na de eerste lockdown de opgelopen leerachterstand in het primair onderwijs geschat op 7-8 lesweken^{17,18}.

11 Van der Molen et al. (2019). Vaardigheidsontwikkelingen volgens PISA en examens. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.

12 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, Routledge.

13 Lek, K. Feskens, R. en Keuning, J. (2020). Het effect van afstandsonderwijs op leerresultaten in het PO. Arnhem: Cito.

14 Engzell, P., Frey, A. en Verhagen, M. D., (2020). Learning Inequality During the Covid-19 Pandemic, No ve4z7, SocArXiv, Center for Open Science.

15 Engzell, P., Frey, A. en Verhagen, M. D., (2020). Learning loss due to school closures during the COVID-19 pandemic, *Proceedings of the National Academy of Sciences* Apr 2021, 118 (17) e2022376118; DOI: 10.1073/pnas.2022376118.

16 Meshcheriakova, O., Borghans, L., Haenbeukers, R. en Schils, T (2020). Limburgse leerlingen op de basisschool ten tijde van corona.

17 Leeruniek (2021). Wat was de impact van 1 jaar corona op de kernvakken in het primair onderwijs?

18 Frey, A., van der Leeuw, R., Verhagen, M. D. (2021). Additional analyses and background information for “Het basisonderwijs 1 jaar na het begin van Corona, wat was de impact op toetsresultaten van de kernvakken?”.

Uit onderzoek een jaar later bleek dat de leerachterstanden inmiddels deels zijn ingelopen^{19,20}. Ook in het voortgezet onderwijs is in 2021 gebleken dat brugklassers lagere schooladviezen^{21,22,23} hebben gekregen en een lagere startvaardigheid hebben. Vooral bij rekenen leverden leerlingen lagere prestaties²⁴. Tegelijkertijd is er ook internationaal onderzoek waaruit blijkt dat de opgelopen leerachterstanden in het secundair onderwijs minder zijn dan in het primair onderwijs²⁵.

Er zijn aanwijzingen dat kandidaten de ruimere herkansing en spreidingsmogelijkheden strategisch, doch beperkt hebben ingezet. In dit onderzoek is het percentage herkansers per centraal examen doorgaans gestegen ten opzichte van 2019. Uit de Examenmonitor²⁶ blijkt dat het percentage kandidaten dat een of twee centrale examens herkanst, licht is afgenomen. Deze schijnbare tegenstrijdigheid komt door een andere berekening van de percentages. In dit onderzoek wordt bij ieder examen afzonderlijk geteld hoeveel eerstekansers er waren, en hoeveel daarvan herkansten. In de examenmonitor wordt er per onderwijsniveau gekeken naar het aantal kandidaten dat één of twee examens herkanst, ten opzichte van het totaal aantal kandidaten in het betreffende onderwijsniveau. Het merendeel van de kandidaten herkanst helemaal geen centrale examens. Maar doordat de meerderheid van de kandidaten die wel herkanst, in 2021 juist twee centrale examens herkanst, terwijl dat voorheen altijd maar één centraal examen betrof, zijn de percentages herkansers per centraal examen gestegen. De Examenmonitor laat tevens zien dat het wegstrepen van de resultaten van één vak veruit de meest gebruikte maatregel is. Tien procent van de eindexamenkandidaten is geslaagd dankzij de mogelijkheid om een vak buiten beschouwing te houden bij de uitslagbepaling. Dit zou het grotere aantal kandidaten met een lage score bij de niet-kernvakken mogelijk kunnen verklaren. In dit onderzoek bleek dat van de onderzochte factoren het percentage lage scores de meeste invloed heeft op de vaardigheidsontwikkeling.

19 Leerunie (2021). Wat was de impact van 1 jaar corona op de kernvakken in het primair onderwijs?

20 Frey, A., van der Leeuw, R., Verhagen, M. D. (2021). Additional analyses and background information for “Het basisonderwijs 1 jaar na het begin van Corona, wat was de impact op toetsresultaten van de kernvakken?”.

21 Emons, W., Frissen, S., & Straat, H. (2021). Effecten van de Schoolsluitingen op de Leerresultaten voor Nederlands en Engels Leesvaardigheid & Woordenschat, en Rekenen-Wiskunde, in het Voortgezet Onderwijs. Arnhem: Centraal Instituut voor Toetsontwikkeling (CITO).

22 (Z.D). Impact van corona. Lagere startvaardigheden brugklassers: onderzoek voor de Staat van het Onderwijs 2021.

23 Hartgers, M., Traag, T., Wielenga, L. (2021). De schooladviezen in groep 8: verschillen tussen groepen leerlingen. CBS.

24 Emons, W., Frissen, S., en Straat, H. (2021). Effecten van de Schoolsluitingen op de Leerresultaten voor Nederlands en Engels Leesvaardigheid & Woordenschat, en Rekenen-Wiskunde, in het Voortgezet Onderwijs. Arnhem: Centraal Instituut voor Toetsontwikkeling (CITO).

25 Tomasik, M. J. , Helbling, L. A. en Moser, U. (2020). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *Int. J. Psychol.*, doi:10.1002/ijop.12728 (2020).

26 DUO (2021). Examenmonitor VO 2021. Examenmonitor Voortgezet Onderwijs 2021 | Kamerstuk | Rijksoverheid.nl

6 Literatuur

6 Literatuur

CBS (2021). Mentale gezondheid in eerste helft 2021 op dieptepunt.

<https://www.cbs.nl/nl-nl/nieuws/2021/35/mentale-gezondheid-in-eerste-helft-2021-op-dieptepunt>

Cito (Z.D.). Impact van corona. Lagere startvaardigheden brugklassers: onderzoek voor de Staat van het Onderwijs 2021. <https://www.cito.nl/kennis-en-innovatie/actuele-themas/impact-van-corona/effecten-afstandsonderwijs-brugklas>

Cito (Z.D.). Tools en informatie voor docenten: Wolf. <https://www.cito.nl/onderwijs/voortgezet-onderwijs/centrale-examens-voortgezet-onderwijs/tools-en-informatie-voor-docenten/wolf>.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, Routledge.

College voor Toetsen en Examens (Z.D.). Examenblad.nl: Normering centrale examens. <https://www.examenblad.nl/onderwerp/normering-centrale-examens>

College voor Toetsen en Examens (30 november 2016). Examenblad.nl: De normeringssystematiek van de Centrale Examens VO. https://www.examenblad.nl/document/de-normeringssystematiek-van-de-ce/2017/f=/De_normeringssystematiek_van_de_CEs_vo_30_november_2016.pdf

College voor Toetsen en Examens (2017). Examenblad.nl: Wat is Facet? <https://www.examenblad.nl/veel-gevraagd/wat-is-facet/2017/vwo?horizon=vg41h1h4i9qe>

College voor Toetsen en Examens (11-06-2019). De N-term maken we samen. <https://www.cvte.nl/actueel/nieuws/2019/6/11/de-n-term-maken-we-samen>

College voor Toetsen en Examens (19-02-2021). Examenblad.nl: Normering centrale examens 2021. <https://www.examenblad.nl/nieuws/20210219/normering-centrale-examens-2021/2021>

College voor Toetsen en Examens (18 maart 2021). Examenblad: Maatregelen coronavirus en de centrale examens VO 2021. <https://www.examenblad.nl/nieuws/20210318/maatregelen-coronavirus-en-de/2021>

College voor Toetsen en Examens (2021). Persmap centraal examen 2021. <https://www.cvte.nl/documenten/publicaties/2021/04/02/persmap-centraal-examen-2021>

DUO (2021). Examenmonitor VO 2021. Examenmonitor Voortgezet Onderwijs 2021 | Kamerstuk | Rijksoverheid.nl

Emons, W., Frissen, S., & Straat, H. (2021). Effecten van de Schoolsluitingen op de Leerresultaten voor Nederlands en Engels Leesvaardigheid & Woordenschat, en Rekenen-Wiskunde, in het Voortgezet Onderwijs. Arnhem: Centraal Instituut voor Toetsontwikkeling (CITO). Effecten afstandsonderwijs brugklas (cito.nl)

Engzell, P., Frey, A. en Verhagen, M. D., (2020). Learning loss due to school closures during the COVID-19 pandemic, *Proceedings of the National Academy of Sciences* Apr 2021, 118 (17) <https://www.pnas.org/content/118/17/e2022376118>

Engzell, P., Frey, A. en Verhagen, M. D., (2020). Learning Inequality During the Covid-19 Pandemic, No ve4z7, *SocArXiv*, Center for Open Science. <https://osf.io/preprints/socarxiv/ve4z7/>

Frey, A., van der Leeuw, R., Verhagen, M. D. (2021). Additional analyses and background information for “Het basisonderwijs 1 jaar na het begin van Corona, wat was de impact op toetsresultaten van de kernvakken?”. <https://www.nponderwijs.nl/documenten/rapporten/2021/06/15/technical-appendix-corona-onderzoek>

Hartgers, M., Traag, T., Wielenga, L. (2021). De schooladviezen in groep 8: verschillen tussen groepen leerlingen. CBS. <https://www.cbs.nl/nl-nl/longread/statistische-trends/2021/de-schooladviezen-in-groep-8-verschillen-tussen-groepen-leerlingen?onepage=true>

ImpactEd (2001). Pupil learning and wellbeing during the Covid-19 pandemic. <https://impacted.org.uk/covid-19>

Inspectie van het Onderwijs (24-11-2020). Nieuwsbericht: Onderwijs tijdens COVID-19: scholen en instellingen hebben zorgen over de continuïteit van het onderwijs en over leerachterstanden van leerlingen en studenten. <https://www.onderwijsinspectie.nl/actueel/nieuws/2020/11/24/onderwijs-tijdens-covid-19-scholen-en-instellingen-hebben-zorgen-over-de-continuïteit-van-het-onderwijs-en-over-leerachterstanden-van-leerlingen-en-studenten>.

Inspectie van Onderwijs (13-5-2020). Nieuwsbericht: Onderwijs op afstand tijdens COVID-19: grote inspanningen en de nodige aandachtspunten. <https://www.onderwijsinspectie.nl/actueel/nieuws/2020/05/13/onderwijs-op-afstand-tijdens-covid-19-grote-inspanningen-en-de-nodige-aandachtspunten>

Inspectie van Onderwijs (2021). De Staat van het Onderwijs. <https://www.onderwijsinspectie.nl/documenten/rapporten/2021/04/14/de-staat-van-het-onderwijs-2021>

Leeruniek (2021). Wat was de impact van 1 jaar corona op de kernvakken in het primair onderwijs? <https://www.nponderwijs.nl/documenten/rapporten/2021/06/15/impact-corona-op-kernvakken-po>

Lek, K. Feskens, R. en Keuning, J. (2020). Het effect van afstandsonderwijs op leerresultaten in het PO. Arnhem: Cito. https://www.cito.nl/-/media/files/kennis-en-innovatie-onderzoek/201104_cito_onderzoeksrapport-effect-afstandsonderwijs.pdf?la=nl-nl

Meshcheriakova, O., Borghans, L., Haenbeukers, R. en Schils, T (2020). Limburgse leerlingen op de basisschool ten tijde van corona. <https://www.educatieveagendalimburg.nl/bijdragen/2020>

Oberon (2020). Periodieke peiling COVID-19 in het funderend onderwijs Meting week 9 t/m 13 november 2020. <https://www.rijksoverheid.nl/documenten/kamerstukken/2020/12/03/periodieke-peiling-covid19-november-20201130b>.

Onderwijsraad (29 juni 2020). Advies vooruitzien voor jonge generaties: tweede advies over de gevolgen van de coronacrisis voor het onderwijs. p. 15. <https://www.rijksoverheid.nl/documenten/rapporten/2020/06/29/advies-or-vooruitzien-voor-jonge-generaties>

Paechter, M., Luttenberger, S., Macher, D., Berding, F., Papousek, I., en Weiss E.M. et al. (2015). The effects of nine-week summer vacation: losses in mathematics and gains in reading. EURASIA J Math Sci Technol Educ (2015) 11(6):1339–413. doi:10.12973/eurasia.2015.1397a. <https://www.ejmste.com/download/the-effects-of-nine-week-summer-vacation-losses-in-mathematics-and-gains-in-reading-4444.pdf>

Rijksoverheid (Z.D.). Coronavirus tijdlijn: Maart 2020: Maatregelen tegen verspreiding coronavirus, intelligente lockdown. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/maart-2020-maatregelen-tegen-verspreiding-coronavirus>

Rijksoverheid (Z.D.). Eindexamens 2021. <https://www.rijksoverheid.nl/onderwerpen/eindexamens/eindexamens-2021>

Rijksoverheid (Z.D.). Eindexamens 2021: Wanneer zijn de examens voortgezet onderwijs in 2021? <https://www.rijksoverheid.nl/onderwerpen/eindexamens/eindexamens-2021/wanneer-zijn-de-examens-voortgezet-onderwijs-in-2021>.

Rijksoverheid (15-03-2020). Nieuwsbericht: Aanvullende maatregelen onderwijs, horeca, sport. <https://www.rijksoverheid.nl/actueel/nieuws/2020/03/15/aanvullende-maatregelen-onderwijs-horeca-sport>

Slob, A. (08-04-2020). Kamerbrief tijdelijke slaag-zakregeling examens vo. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/documenten/kamerstukken/2020/04/08/kamerbrief-tijdelijke-slaag-zakregeling-examens-vo>

Slob, A. (12-02-2021). Kamerbrief over aanvulling besluit eindexamen voortgezet onderwijs 2021. <https://www.rijksoverheid.nl/documenten/kamerstukken/2021/02/12/aanvulling-besluit-eindexamen-voortgezet-onderwijs-2021>

Slob, A. (31-8-2020). Kamerbrief over voorbereiding nieuwe schooljaar. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/documenten/kamerstukken/2020/08/31/start-nieuwe-schooljaar>

Slob., A. (20-11-2020). Kamerbrief over stand van zaken november 2020 eindexamen VO 2021. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/documenten/kamerstukken/2020/11/20/eindexamen-vo-2021-stand-van-zaken-november-2020>

Slob., A. (24-11-2020). Kamerstuk Continuïteit van onderwijs in schooljaar 2020-2021. <https://www.rijksoverheid.nl/documenten/kamerstukken/2020/11/24/continuïteit-van-onderwijs-in-schooljaar-2020-21-stand-van-zaken>

Slob, A. (24-11-20). Kamerbrief over continuïteit van onderwijs in schooljaar 2020 - 21 stand van zaken. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/documenten/kamerstukken/2020/11/24/continuïteit-van-onderwijs-in-schooljaar-2020-21-stand-van-zaken>.

Slob, A. (16-12-2020). Kamerbrief over Besluit eindexamen voortgezet onderwijs 2021. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/documenten/kamerstukken/2020/12/16/besluit-eindexamen-voortgezet-onderwijs-2021>.

Slob, A. (22-01-2021). Kamerstuk 31293, nr. 576. Brief van de minister voor basis- en voortgezet onderwijs en media aan de Voorzitter van de Tweede Kamer der Staten-Generaal. <https://zoek.officielebekendmakingen.nl/kst-31293-576.html>

Snijders, T.A.B. & Bosker, R.J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling* (sec. ed.). London, Sage.

Tomasik, M. J. , Helbling, L. A. en Moser, U. (2020). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *Int. J. Psychol.*, doi:10.1002/ijop.12728 (2020). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7753520/>

Universiteit van Amsterdam (uva.nl). Welke impact hebben de coronamaatregelen op onderwijs(on)gelijkheid? - <https://www.uva.nl/shared-content/faculteiten/nl/faculteit-der-maatschappij-en-gedragwetenschappen/nieuws/2020/03/welke-impact-hebben-de-coronamaatregelen-op-onderwijsongelijkheid.html>

Van der Molen, P., Schouwstra, S., Feskens, R., & Van Onna, M. (2019). *Vaardigheidsontwikkelingen volgens PISA en examens*. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling. <https://www.cito.nl/kennis-en-innovatie/kennisbank/cte-rapport-vaardigheidsontwikkelingen-volgens-pisa-en-examens>

<https://www.rijksoverheid.nl/documenten/rapporten/2019/04/18/vaardigheidsontwikkelingen-volgens-pisa-en-examens>

Van Engelshoven, I. & Slob, A. (17-02-2021). Persconferentie Nationaal Programma Onderwijs. <https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/documenten/toespraken/2021/02/17/persconferentie-nationaal-programma-onderwijs>

VO-raad (23-03-2021). Nieuws: Leerlingen gemiddeld 47% van de tijd op school. <https://www.vo-raad.nl/nieuws/leerlingen-gemiddeld-47-van-de-tijd-op-school>

VO-raad (28-05-2021). Nieuws: Kamer stemt in met loslaten 1,5-meterregel voor leerlingen. <https://www.vo-raad.nl/nieuws/kamer-stemt-in-met-loslaten-1-5-meterregel-voor-leerlingen>

wetten.nl - Regeling - Regeling omzetting scores in cijfers VO - BWBR0037590 (overheid.nl)

Appendix

Appendix I | Context van het Centraal Examen 2021

Maatregelen in VO tijdens pandemie

Schooljaar 2019-2020

Nadat in januari 2020 de eerste berichten over het corona-virus in China kwamen, zijn de ontwikkelingen snel gegaan. Begin februari ontdekte men de eerste corona-gevallen in Italië en 27 februari werd de eerste coronabesmetting in Nederland geconstateerd. Hierna komt het onderwijs al snel in een ongekende situatie, waarin scholen gesloten worden en op afstand onderwijs werd gegeven.

Op 9 maart worden in een persconferentie met premier Rutte de eerste instructies voor algemene hygiënemaatregelen gegeven (handen wassen, in de elleboog niezen, papieren zakdoekjes gebruiken en geen handen meer schudden). Nog geen week later, op 15 maart, worden de Nederlandse scholen in het basisonderwijs, voortgezet onderwijs (vo) en middelbaar beroepsonderwijs (mbo) gesloten. Scholen gaan het onderwijs op afstand organiseren met prioriteit voor eindexamenkandidaten 2020 in het voortgezet onderwijs^{27,28}. Nog in maart 2020 kondigt Minister Slob aan dat de eindtoets niet zal worden afgenomen bij leerlingen in groep 8 en dat het centraal examen in het voortgezet onderwijs komt te vervallen. Leerlingen krijgen een diploma op basis van de resultaten op hun schoolexamens, maar de slaag-zakregeling blijft zo veel mogelijk in stand²⁹.

Pas 11 mei gaat het basisonderwijs weer gedeeltelijk open en per 2 juni volgt het voortgezet onderwijs. Veel scholen in het voortgezet onderwijs kiezen voor een hybride vorm van lesgeven (gedeeltelijk digitaal en gedeeltelijk op school), omdat leerlingen in het voortgezet onderwijs 1,5 meter afstand moeten houden. Deze regel vervalt op 1 juli en de deuren van alle scholen in het voortgezet onderwijs gaan weer open. In de periode dat de scholen gesloten waren worden verschillende maatregelen genomen om scholen te ondersteunen bij het onderwijs op afstand³⁰. Zo is een aanvullende investering gedaan om scholen te ondersteunen met devices en internetverbindingen voor leerlingen die daar thuis niet over beschikten. Er zijn miljoenen vrijgemaakt voor inhaal- en ondersteuningsprogramma's om achterstanden in te lopen die zijn ontstaan door de getroffen maatregelen en om extra hulp in de klas in te schakelen. Veel informatie die nodig is voor het organiseren van onderwijs op afstand en hybride onderwijs wordt gebundeld op de website www.lesopafstand.nl.

27 Rijksoverheid (Z.D.). Coronavirus tijdlijn: Maart 2020: Maatregelen tegen verspreiding coronavirus, intelligente lockdown.

28 Rijksoverheid (15-03-2020). Nieuwsbericht: Aanvullende maatregelen onderwijs, horeca, sport.

29 Slob, A. (08-04-2020). Kamerbrief tijdelijke slaag-zakregeling examens vo.

30 Slob, A. (22-01-2021). Kamerstuk 31293, nr. 576. Brief van de minister voor basis- en voortgezet onderwijs en media aan de Voorzitter van de Tweede Kamer der Staten-Generaal.

Schooljaar 2020-2021

In de zomer 2020 treffen scholen voorbereidingen om op een veilige wijze volledig open te gaan, volgens de richtlijnen van het RIVM³¹. Vanaf maandag 31 augustus zijn alle scholen in het primair en voortgezet onderwijs weer volledig open. Tegelijkertijd loopt het aantal besmettingen langzaam op vanaf de zomer en stapsgewijs worden aangescherpte maatregelen aangekondigd.

In oktober is sprake van de tweede golf en 13 oktober wordt een gedeeltelijk lockdown ingesteld. De scholen blijven open, maar er komt een mondkapjes-maatregel: In het voortgezet onderwijs (VO), MBO en hoger onderwijs (HO) draagt iedereen een mondkapje buiten de les. Het kabinet stelt daarnaast geld beschikbaar om de ventilatie in scholen te verbeteren en om leerlingen te helpen achterstanden als gevolg van corona in te lopen (bijvoorbeeld voor bijlessen)³². Voor scholen komt er een meldpunt voor schoolsluiting vanwege corona. Dat meldpunt valt onder de verantwoordelijkheid van de Inspectie van het Onderwijs. Daarnaast start de periodieke peiling COVID-19 in het funderend onderwijs waarin onder meer de afwezigheid van leerlingen en lesuitval vanwege corona-gerelateerde aspecten wordt gemeten³³.

In november en december wordt de gedeeltelijk lockdown verder verzaamd, bijvoorbeeld met de mondkapjesplicht. Scholen en instellingen hebben steeds meer zorgen over de continuïteit van het onderwijs in schooljaar 2020–21 en over leerachterstanden van leerlingen en studenten³⁴³⁵. Het kabinet trekt opnieuw extra geld uit om leerlingen van een laptop of tablet te voorzien en voor scholen om extra hulp in te schakelen. Vlak voor de Kerstvakantie, op 14 december, gaat Nederland weer in lockdown en sluiten alle scholen. Vanwege het grote belang van de eindexamens heeft het kabinet besloten dat tijdens de lockdown examenkandidaten les blijven volgen op school³⁶. Het centraal examen gaat in aangepaste vorm door, zo mogen kandidaten de examens spreiden over twee tijdvakken en komt er een derde tijdvak voor herkansing. Ook krijgen alle kandidaten een extra herkansing voor een centraal examen. Dat betekent dat het centraal examen van twee vakken kan worden herkanst.

Het jaar begint met de aanvang van de vaccinatie. De scholen blijven echter gesloten. Alleen leerlingen in de examenklas, kwetsbare leerlingen en leerlingen in alle leerjaren van het vmbo, vso en praktijkonderwijs (pro) met beroepsgerichte vakken mogen wel op school lessen volgen. Vanaf 13 januari moeten deze leerlingen ook op school onderling 1,5 m afstand houden. Eind januari informeert minister Slob (OCW) over een nationaal programma voor het funderend onderwijs om de effecten van COVID-19 op het onderwijs ook op de lange termijn op te vangen. In die kamerbrief waarschuwt minister Slob dat het niet realistisch is “om te verwachten dat alle

31 Slob, A. (31-8-2020). Kamerbrief over voorbereiding nieuwe schooljaar.

32 Slob, A. (24-11-2020). Kamerstuk Continuïteit van onderwijs in schooljaar 2020-2021.

33 Oberon (2020). Periodieke peiling COVID-19 in het funderend onderwijs Meting week 9 t/m 13 november 2020.

34 Inspectie van het Onderwijs (24-11-2020). Nieuwsbericht: Onderwijs tijdens COVID-19: scholen en instellingen hebben zorgen over de continuïteit van het onderwijs en over leerachterstanden van leerlingen en studenten.

35 Slob, A. (24-11-20). Kamerbrief over continuïteit van onderwijs in schooljaar 2020 - 21 stand van zaken.

36 Slob, A. (16-12-2020). Kamerbrief over Besluit eindexamen voortgezet onderwijs 2021.

kinderen de opgelopen achterstand op de verschillende ontwikkelgebieden dit jaar al volledig inhalen”. Minister Slob wil de achterstanden over de gehele linie in kaart brengen.

Alhoewel op 8 februari de basisscholen weer openen, blijven de scholen in het voortgezet onderwijs in februari gesloten. Ook de avondklok die op 23 januari is ingegaan blijft gelden in februari. Op 12 februari informeert minister Slob de kamer dat ze extra middelen vrijmaken om extra ondersteuning te kunnen bieden aan examenkandidaten, bijvoorbeeld voor de vormgeving van (individuele) ondersteuning van examenkandidaten via de scholen, zoals de inzet van studenten uit het hoger onderwijs³⁷. Ook houdt minister Slob een persconferentie over de investeringen in het onderwijs via het Nationaal Programma Onderwijs³⁸. Met dit programma wordt beoogd leerlingen, docenten en scholen voor een paar schooljaren extra te ondersteunen om opgelopen vertragingen weg te werken.

Uiteindelijk mogen leerlingen in het voortgezet onderwijs per 1 maart weer minstens 1 dag fysiek naar school. Volgens de peiling van de VO-raad in maart onder ruim 500 schoolleiders brengen leerlingen gemiddeld 47% van de onderwijstijd door op school³⁹. Half maart komt er een update van het servicedocument examens 2021 beschikbaar; voorts verschijnt het besluit eindexamens 2021⁴⁰. Helaas lopen tegelijkertijd in maart de besmettingen op en is sprake van de derde golf.

Nadat vanaf 5 april de afnames van de digitale examens in vmbo-bb en vmbo-kb beginnen, begint op 17 mei het eerste tijdvak van het centraal schriftelijk examen (zie Tabel 3). Pas na het eerste tijdvak vervalt de 1,5 meter afstandsregel voor leerlingen binnen school en kunnen scholen in het voortgezet onderwijs per 31 mei weer volledig fysiek onderwijs gaan geven⁴¹. Vanaf 7 juni zijn alle scholen verplicht om alle leerlingen weer voltijds op school te ontvangen.

37 Slob, A. (12-02-2021). Kamerbrief over aanvulling besluit eindexamen voortgezet onderwijs 2021.

38 Van Engelshoven, I. & Slob, A. (17-02-2021). Persconferentie Nationaal Programma Onderwijs.

39 VO-raad (23-03-2021). Nieuws: Leerlingen gemiddeld 47% van de tijd op school.

40 Rijksoverheid (Z.D.). Eindexamens 2021.

41 VO-raad (28-05-2021). Nieuws: Kamer stemt in met loslaten 1,5-meterregel voor leerlingen.

Tabel 3 Afnamedagen van het centraal eindexamen 2021⁴²

Tijdvak	Afnamedagen	Bekendmaking N-termen (uitslag)
Tijdvak 1: Eerste afnames	17 mei t/m 1 juni	10 juni
Tijdvak 2: Eerste afnames en herkansingen	14 juni t/m 25 juni	2 juli
Tijdvak 3: Herkansingen	6 t/m 9 juli*	15 juli
Afname periode BB- en KB-flex	5 april t/m 23 juli	16 juni

Maatregelen centraal examen 2021

Aan het begin van het schooljaar 20/21 informeert minister Slob de kamer over de voorbereiding van het nieuwe schooljaar⁴³. Omdat het voortgezet onderwijs volledig opengaat, is zijn streven dat het centraal examen in de normale vorm zal plaatsvinden. Wel geeft minister Slob scholen enige extra ruimte voor het organiseren van het eindexamen door de aanlever- en afnametermijnen voor examen(s)(resultaten) te verruimen. Vanwege waargenomen onderwijsachterstanden en twijfels over de haalbaarheid van het centraal schriftelijk en praktisch examen (cspe) valt in november⁴⁴ het besluit dat het beroepsgerichte profielvak dit jaar zal worden afgesloten met een schoolexamen in plaats van een centraal examen. Eind 2020, op het moment dat de scholen net weer opnieuw gesloten zijn in verband met het oplopende aantal besmettingen met corona, informeert minister Slob (OCW) de Tweede Kamer over het eindexamen in het voortgezet onderwijs in 2021⁴⁵. In het belang van kandidaten heeft minister Slob besloten het centraal examen door te laten gaan. Tegelijkertijd worden maatregelen genomen om kandidaten “de best mogelijke uitgangspositie geven in termen van tijd, flexibiliteit, compensatie en kansengelijkheid en er vervolgens sprake zal zijn van een volwaardig diploma voor doorstroom naar het vervolgonderwijs”. Eindexamenkandidaten kunnen bijvoorbeeld les blijven volgen op school tijdens de tweede lockdown.

Op 12 februari informeert Minister Slob de Tweede Kamer over de aanvullingen op het besluit eindexamens voortgezet onderwijs 2021⁴⁶. De aanleiding voor de aanvulling zijn de veranderde omstandigheden, met name de herinvoering van de anderhalve meterregel in het onderwijs voor leerlingen onderling. Minister Slob heeft besloten dat eindexamenkandidaten die anders

42 Rijksoverheid (Z.D.). Eindexamens 2021: Wanneer zijn de examens voortgezet onderwijs in 2021?

43 Slob, A. (31-08-20). Kamerbrief over voorbereiding nieuwe schooljaar.

44 Slob, A. (20-11-2020). Kamerbrief over stand van zaken november 2020 eindexamen VO 2021.

45 Slob, A. (16-12-2020). Kamerbrief over Besluit eindexamen voortgezet onderwijs 2021.

46 Slob, A. (12-02-2021). Kamerbrief over aanvulling besluit eindexamen voortgezet onderwijs 2021.

net gezakt zouden zijn dit jaar het eindresultaat van één vak (niet zijnde een kernvak) mogen wegstrepen als zij daarmee kunnen slagen voor het diploma.

Op 2 april 2021 publiceert het College voor Toetsen en Examens de definitieve persmap centraal examens 2021⁴⁷, nadat de minister de tweede kamer en de scholen heeft geïnformeerd. In de persmap worden de ruimere mogelijkheden opgesomd die kandidaten in 2021 krijgen bij het afleggen van het examen in 2021. De volgende maatregelen zijn genomen⁴⁸:

- De afnameperiode van de flexibele digitale centrale examens vmbo BB en KB is verlengd⁴⁹;
- Het beroepsgerichte profielvak zal dit jaar worden afgesloten met alleen een schoolexamen⁵⁰;
- Een extra tijdvak en extra herkansing die op de volgende wijze zijn vormgegeven:⁵¹
 - De mogelijkheid voor de leerling om het centraal examen te spreiden over twee tijdvakken bestaande uit een eerste tijdvak voor eerste afnames en een tweede tijdvak voor eerste afnames en herkansingen;
 - Uitbreiding tweede tijdvak van 4 naar 10 afnamedagen;
 - Een extra herkansing voor het centraal examen: een leerling kan in twee vakken uit het centraal examen herkansen;
- Een 4-daags derde tijdvak (herkansingstijdvak) op de scholen na het tweede tijdvak voor de afronding van de herkansingen.
- De mogelijkheid om een vak weg te strepen om te kunnen slagen⁵². Bij de vaststelling van de uitslag kan, indien dat nodig is om de leerling te laten slagen, het eindcijfer van één vak, niet zijnde een kernvak, buiten de bepaling van de definitieve uitslag worden gehouden. Het volledige eindexamen moet wel zijn afgelegd alvorens gebruik gemaakt kan worden van deze mogelijkheid.

47 College voor Toetsen en Examens (2021). Persmap centraal examen 2021.

48 College voor Toetsen en Examens (18 maart 2021). Examenblad: Maatregelen coronavirus en de centrale examens VO 2021.

49 Slob, A. (31-08-20). Kamerbrief over voorbereiding nieuwe schooljaar.

50 Slob, A. (20-11-2020). Kamerbrief over stand van zaken november 2020 eindexamen VO 2021.

51 Slob, A. (16-12-2020). Kamerbrief over Besluit eindexamen voortgezet onderwijs 2021.

52 Slob, A. (12-02-2021). Kamerbrief over aanvulling besluit eindexamen voortgezet onderwijs 2021.

Appendix II | Berekeningswijze vaardigheidsverschil tussen 2021 en pre-corona

Onderliggend aan Figuur 1 en Figuur 2 zijn berekeningen uitgevoerd. In deze appendix worden de details daarvan besproken. Het doel was om de prestaties op de examens in 2021 te relateren aan de prestaties in de jaren voor de corona-pandemie.

Gebruikte data

Voor de schatting van de vaardigheid van 2021 zijn alleen de kandidaten meegenomen die deelnamen aan het eerste tijdvak, en voor de digitale examens alleen de kandidaten die t/m 9 juni 2021 hebben deelgenomen. De eerstekansers in het tweede tijdvak zijn buiten beschouwing gelaten omdat er geen geschikte equivaleringsgegevens beschikbaar waren voor het tweede tijdvak. Omdat het aantal eerstekansers bij het tweede tijdvak substantieel lager was, is de aanname dat de prestaties van alle eerstekansers niet veel verschillen van de kandidaten die opgenomen zijn in deze analyses. Met één van de deelvragen wordt onderzocht of deze kandidaten van het eerste tijdvak representatief zijn voor alle kandidaten.

Om de pre-corona-vaardigheid te schatten werd gekozen welke afnamejaren de pre-corona-conditie het best weerspiegelen. Een schatting van vaardigheid wordt beter naar mate deze op meer afnamegegevens gebaseerd is. Dan moeten deze afnamegegevens wel een weerspiegeling zijn van eenzelfde vaardigheid. Het laatste pre-corona-jaar was 2019. Om eventuele toevallige fluctuaties in de vaardigheid van 2019 op te vangen, zijn de gegevens van 2015 tot en met 2018 ook meegenomen. Er is daarbij wel eerst gecontroleerd, indien mogelijk, of de vaardigheid in de afnameperiode 2015 tot en met 2019 wel stabiel was.

Bij de AIP, pretest- en posttest-examens heeft CvTE in het voorjaar van 2020 besloten of een centraal examen in de jaren 2015 – 2019 een stabiele vaardigheid had, of dat er sprake was van een trend in de vaardigheid. Bij de stabiele centrale examens is in de berekening de gemiddelde vaardigheid van de populaties 2015 tot en met 2019 gebruikt als pre-corona-vaardigheid. Bij de zogenaamde trend-examens is alleen de vaardigheid van 2019 als pre-corona-vaardigheid gebruikt.

Bij de standaardsettingen is gekozen om altijd 2019 als pre-corona-jaar te gebruiken. Het is bij standaardsetting-examens niet mogelijk om vast te stellen of de populaties in de jaren voor 2019 een stabiele vaardigheid hadden, of niet.

Methoden

In Tabel 1 wordt aangegeven welke centrale examens in de berekening voor de vaardigheidsschattingen per niveau en cluster zijn meegenomen. Tevens wordt daar aangegeven welke methode (Anchor in package, pretest, posttest of standaardbepaling) gebruikt is. Bij Anchor in package, pretest en posttest is een trendschattingwijze gebruikt om de vaardigheid voor en na corona te vergelijken. Bij de standaardbepalingen is op een andere manier de daling in cijfer tussen pre-corona en post-corona berekend. In alle gevallen resulteerde een verschil in gemiddeld cijfer tussen de pre-corona-populatie en de populatie in 2021, met daarbij een schatting van de nauwkeurigheid van dit verschil.

Trendschatting

Bij een trendschatting zijn alle afnamegegevens van 2015 tot en met 2021, plus eventuele pre- of posttesten opgenomen in één IRT-analyse. Daarvoor moest de moeilijkheid van de opgaven in deze afnamejaren wel aan elkaar gerelateerd kunnen worden via direct hergebruik van opgaven (Anchor in package) of via hergebruik van ankeropgaven in pre- of posttestboekjes. De IRT-analyse⁵³ leverde naast parameters die de moeilijkheid van de opgaven weergaven, ook schattingen van vaardigheid op de latente trek. Deze vaardigheden op de latente trek werden vervolgens uitgedrukt als scores op een (grote) set van items. Dus iedere kandidaat in de periode 2015 – 2021 had ‘virtueel’ een score behaald op dezelfde mega-toets; een score tussen 0 en de maximum score op de totale set van items.

Vervolgens werd per centraal examen een representatieve pre-corona-periode aangewezen, zoals hierboven beschreven bij de gebruikte data. Van de kandidaten in de pre-corona-periode werd de gemiddelde score berekend, van de kandidaten in 2021 eveneens. Een daling in gemiddelde scores tussen deze twee gemiddelden reflecteerde dus een daling in vaardigheid. Het verschil in gemiddelde scores van 2021 en de pre-corona-periode werd uitgedrukt in een verschil in gemiddeld cijfer⁵⁴ om de getallen beter te kunnen duiden, en om centrale examens te kunnen aggregeren.

De puntschatting van dit zogenaamde contrast werd voorzien van een nauwkeurigheidsmaat⁵⁵. In de berekening van deze nauwkeurigheidsmaat werd rekening gehouden met de nauwkeurigheid van de schatting van het IRT-model op de data. Bij een centraal examen met relatief veel hergebruik van opgaven en grotere aantallen afnames (Anchor in package), is deze nauwkeurigheid groter dan bij centrale examens waarbij hergebruik beperkt is tot kleine ankerstetjes in pre- of posttest boekjes met relatief weinig afnames.

Vaardigheidsverschil bij standaardsetting

Bij de gebruikte standaardsettingen werd voorafgaand aan de afname aan experts gevraagd om van twee examens aan te geven welke gemiddelde scores zij verwachten op iedere opgave. Het ene examen was een referentie-examen, het andere examen het eerste tijdvak-examen van 2021 (2021-1). Deze inschattingen gaven een verschil in moeilijkheid weer tussen het referentie-examen en het examen van 2021-1. Omdat de afnamegegevens en N-term van het referentie-examen al beschikbaar waren, kon een schatting gemaakt worden van de N-term van het examen 2021-1. Door de spreiding in de antwoorden van de experts te beschouwen, kon de nauwkeurigheid van deze schatting van de N-term berekend worden.

Vervolgens is na afloop van het examen gekeken welk gemiddeld cijfer de kandidaten van 2021-1 behaald zouden hebben als de N-term uit de standaardsetting was toegepast. Dit gemiddelde cijfer weerspiegelt de vaardigheid van de populatie na corona.

53 Vanwege het grote aantal centrale examens zijn er geen uitgebreide analyse op modelfit gedaan. Dit is eveneens een reden waarom de afzonderlijke trendschattingen niet gepresenteerd worden.

54 Uit de gebruikelijke manier waarop scores in cijfers worden omgezet, namelijk met een N-term, volgt dat een verschil in cijfer gelijk is aan een verschil in score vermenigvuldigd met een constante. Deze constante is gelijk aan 9 gedeeld door de maximale score.

55 Standard Error, deze SE is omgekeerd gerelateerd aan de nauwkeurigheid: een grotere SE wijst op minder nauwkeurigheid.

De nauwkeurigheid van dit gemiddelde cijfer was direct af te leiden uit de nauwkeurigheid van de geschatte N-term. Het (constante) gemiddelde cijfer van 2019 werd als pre-corona-vaardigheid gebruikt. Het verschil tussen 2021-vaardigheid en pre-corona-vaardigheid kende dezelfde nauwkeurigheid als de nauwkeurigheid van de geschatte N-term.

Examen-overstijgende berekening

Omdat de nauwkeurigheid bij ieder centraal examen niet bijzonder hoog was, en er bij ieder examen jaarlijks wel schommelingen voorkomen in vaardigheid, zijn de resultaten van meerdere centrale examens geaggregeerd om een vaardigheidsschatting te doen. Dit was mogelijk omdat alle examen-specifieke verschillen op dezelfde cijferpuntschaal waren uitgedrukt. Enerzijds zijn er berekeningen per onderwijsniveau uitgevoerd (Figuur 1), anderzijds per inhoudelijk cluster van centrale examens (Figuur 2). De clustering van de centrale examens voor de examen-overstijgende berekeningen staat weergegeven in Tabel 1.

De examen-overstijgende schatting van het verschil in vaardigheid tussen de pre-corona- en 2021-populaties werd berekend als een gewogen gemiddelde van de examen-specifieke verschillen D_k in vaardigheid,

$$\bar{D}_{overstijgend} = \frac{\sum w_k D_k}{\sum w_k},$$

waarbij de gewichten w_k de nauwkeurigheid van de examen-specifieke schatting weergaven⁵⁶. Omdat de examen-specifieke verschillen onderling kunnen variëren, is de variatie daartussen gebruikt om een betrouwbaarheidsinterval rondom het examen-overstijgende gemiddelde te berekenen. Ook hierbij is rekening gehouden met de gewichten/nauwkeurigheid van de examen-specifieke verschillen.

$$V(\bar{D}_{overstijgend}) = \frac{1}{(\sum w_k)^2} \sum w_k^2 (D_k - \bar{D}_{overstijgend})^2$$

Het 90%-betrouwbaarheidsinterval in Figuur 1 en Figuur 2 loopt van

$$\bar{D}_{overstijgend} - 1,645 * \sqrt{V(\bar{D}_{overstijgend})} \text{ tot } \bar{D}_{overstijgend} + 1,645 * \sqrt{V(\bar{D}_{overstijgend})}.$$

⁵⁶ $w_k = 1/(SE_k)^2$

Vaardigheid Examenkandidaten 2021

Cito

Amsterdamseweg 13
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
klantenservice@cito.nl
www.cito.nl

Fotograaf: Gijs Versteeg
© Cito B.V. Arnhem (2021)