

Toelichting bij de adviesnormeringen Cito Kijk- en luistertoetsen MVT 2024/2025

Onderzoek naar aanleiding van verspreiden van antwoorden via Telegram

Tijdens de afnameweek kregen we signalen binnen dat er antwoorden van verschillende toetsen op Telegram zouden circuleren. Na onderzoek bleek dat hier inderdaad sprake van was, en dat het vooral leek te gaan over de toetsen vwo van Duits, Engels en Frans. Cito heeft klanten toen gemaild en gesuggereerd om, bij twijfel en waar mogelijk, een andere jaargang in te zetten dan de nieuwste. Daarnaast heeft Cito aangifte gedaan bij de politie. Dat onderzoek is nog steeds gaande. We zullen later hierover meer informatie verstrekken. Er was geen sprake van een hack bij Cito.

We willen ervoor zorgen dat leerlingen die op een eerlijke manier de toets hebben gemaakt, de normering krijgen die ze verdienen. We hebben daarom bij het vaststellen van de normering met extra aandacht gekeken naar eventueel afwijkende resultaten die kunnen duiden op voorkennis. Wij hebben bij deze analyse geen afwijkende, hoge scores aangetroffen. Ook gaf de cijferverdeling binnen scholen geen aanleiding om te spreken van veelvoorkomende fraude. Alleen bij Engels vwo zien we een hogere, landelijke vaardigheid, maar die toename ligt in lijn met die van voorgaande jaren. De populatie bij Engels lijkt gewoon elk jaar vaardiger te worden. Er is voor ons geen aanleiding om te spreken van grootschalige fraude, en we vertrouwen erop dat we tot een eerlijke normering zijn gekomen.

Normering Duits vwo aangepast

Over de toets Duits vwo zijn meerdere meldingen binnengekomen. Docenten gaven aan dat veel leerlingen in tijdnood kwamen (bij de digitale variant) en/of de toets te moeilijk vonden. Dat was terug te zien in de data. Bij Duits vwo is daarom de normering versoepeld. Duits vwo is dit jaar de enige toets waarbij we versoepeld hebben.

Landelijke gemiddelden en percentages onvoldoende per toets

Hieronder geven we een overzicht van de verschillende gegevens per toets van de afgelopen drie jaar. Waar mogelijk zijn er aparte gegevens voor de digitale variant en de papieren variant. Voor sommige papieren varianten zijn er te weinig afnames binnengekomen om landelijke gegevens te kunnen rapporteren. Duits bb en Spaans (alle niveaus) worden alleen digitaal afgenomen. Van Frans vmbo kb zijn te weinig afnames te rapporteren. Deze normering wordt altijd vastgelegd op basis van de normering van Frans vmbo gt.

Spaans havo ook in de Caraïben afgenomen

De Kijk- en luistertoetsen Spaans worden ook afgenomen op scholen in de Caraïben. De scores daar zijn beduidend hoger, omdat Spaans voor deze leerlingen meer een moedertaal is. Dit heeft geen gevolgen voor de normering. We rapporteren in dit overzicht alleen de gegevens van de scholen in Nederland.

Voor meer informatie over hoe de Kijk- en luistertoetsen worden gemaakt, en hoe de normering tot stand komt: zie de laatste pagina van dit document. Die tekst staat ook bij de veelgestelde vragen voor leerlingen op cito.nl/luistertoetsen.

Duits

Duits vwo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	460	5.919	382	6.594	222	6.176
gemiddelde	6,4	6,3	6,2	6,2	6,1	6,1
% onvoldoende	28	29	32	33	36	31

Duits vmbo kb	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	–	303	–	253	–	331
gemiddelde	–	6,0	–	6,1	–	6,1
% onvoldoende	–	31	–	30	–	31

Duits havo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	337	4.597	267	2.187	–	3.189
gemiddelde	6,2	6,2	6,0	6,0	–	6,2
% onvoldoende	31	30	37	35	–	28

Duits vmbo bb*	2023	2024	2025
aantal kandidaten	95	61	71
gemiddelde	6,1	6,3	6,0
% onvoldoende	25	25	18

* alleen digitaal af te nemen

Duits vmbo gt	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	–	3.403	–	3.431	–	3.018
gemiddelde	–	6,2	–	6,1	–	6,1
% onvoldoende	–	30	–	29	–	30

Engels

Engels vwo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	911	12.864	827	13.728	706	13.062
gemiddelde	6,7	6,7	6,6	6,5	6,7	6,8
% onvoldoende	26	25	27	26	25	22

Engels havo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	1.286	15.533	966	18.023	581	13.563
gemiddelde	6,7	6,7	6,7	6,7	6,8	6,8
% onvoldoende	23	21	22	23	20	18



Engels vmbo gt	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	277	8.389	–	11.796	384	8.577
gemiddelde	6,6	6,6	–	6,6	6,6	6,6
% onvoldoende	29	26	–	27	28	27

Engels vmbo bb	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	–	1.092	–	1.281	–	1.577
gemiddelde	–	7,2	–	7,2	–	7,2
% onvoldoende	–	15	–	16	–	19

Engels vmbo kb	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	–	2.773	–	2.620	–	3.612
gemiddelde	–	6,7	–	6,6	–	6,9
% onvoldoende	–	27	–	26	–	22

Frans

Frans vwo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	244	5.064	–	5.399	–	5.468
gemiddelde	6,1	6,1	–	6,2	–	6,1
% onvoldoende	36	34	–	34	–	35

Frans vmbo gt*	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	–	1.064	–	1.448	–	1.190
gemiddelde	–	6,0	–	6,0	–	6,0
% onvoldoende	–	34	–	38	–	38

Frans havo	2023p	2023d	2024p	2024d	2025p	2025d
aantal kandidaten	192	3.276	–	3.655	–	2.614
gemiddelde	6,0	6,0	–	6,2	–	6,0
% onvoldoende	31	34	–	28	–	35

* In de correctiesleutels van Frans vmbo gt (in de bijlagen in Woots) stond een foute sleutel. Opgave 12 van het audiodeel vmbo gt in versie 1 moet A zijn, niet C. Dit is gecorrigeerd op maandag 3 februari om 17.00 uur 's middags.



Spaans (alleen digitaal)

Spaans vwo	2023	2024	2025
aantal kandidaten	1.397	1.640	1.420
gemiddelde	6,6	6,4	6,6
% onvoldoende	22	28	25

Spaans havo elementair	2023	2024	2025
aantal kandidaten	313	388	360
gemiddelde	6,3	6,5	6,9
% onvoldoende	24	17	14

Spaans vwo elementair	2023	2024	2025
aantal kandidaten	191	278	391
gemiddelde	6,6	6,2	6,6
% onvoldoende	12	21	25

Spaans vmbo	2023	2024	2025
aantal kandidaten	225	356	395
gemiddelde	6,0	6,3	6,3
% onvoldoende	36	19	22

Spaans havo	2023	2024	2025
aantal kandidaten	849	983	955
gemiddelde	6,8	6,2	6,4
% onvoldoende	33	23	36

Hoe komt een Kijk- en luistertoets tot stand?

Wij maken onze toetsen met een team van toetsconstructeurs. Mensen die bij Cito werken en het hele jaar door bezig zijn om materiaal te zoeken of te maken, en daar opgaven bij te bedenken. Soms halen we het materiaal van omroepen, soms nemen wij zelf interviews af of schrijven we teksten die we laten inspreken.

Al dat materiaal leggen we voor aan een adviescommissie. Daarin zitten docenten die lesgeven in het voortgezet onderwijs. Wij vragen hun dan te kijken of ze het materiaal geschikt vinden voor leerlingen: is het interessant, is het niet te moeilijk, wordt er niet te snel gesproken? Als zij akkoord geven, gaan wij fragmenten selecteren uit het materiaal en er meerkeuze-opgaven bij bedenken. We bespreken de opgaven intensief met elkaar. De fragmenten (audio of video) worden gemonteerd door de media-afdeling van Cito.

Dat gaat weer naar de adviescommissie, die gevraagd wordt om het toetsonderdeel te maken. Zij geven dan allerlei soorten kritiek: het goede antwoord moet duidelijker, het woordgebruik van de opgaven moet eenvoudiger of korter, één van de foute antwoorden moet minder vals gemaakt worden, enzovoort. Wij passen de opgaven dan een laatste keer aan.



Vervolgens wordt het toetsonderdeel geproeftoetst. Dat gebeurt door middel van een steekproef onder leerlingen voor wie de toets bedoeld is: vwo 6, havo 5 of vmbo 4. Honderden leerlingen in heel Nederland maken dan de opgaven. Voor hen is dit een goede manier om te oefenen voor de 'echte' Kijk- en luistertoets. Voor ons is het dé manier om te zien hoe moeilijk de opgaven nu echt zijn en of de opgaven eerlijk meten. We vragen ook feedback aan de docenten die de proeftoetsen afnemen. Na zo'n proeftoets vallen er altijd een aantal opgaven af. Opgaven die te moeilijk waren, of veel te makkelijk. Opgaven die onduidelijk waren, of waarbij door heel veel leerlingen maar wat werd gegokt. Alleen de opgaven die goed functioneerden, worden uiteindelijk opgenomen in een Kijk- en luistertoets.

Doordat wij de Kijk- en luistertoetsen zo samenstellen uit opgaven uit de proeftoetsen, kunnen we goed voorspellen hoe moeilijk de Kijk- en luistertoets is. Deze informatie nemen we mee bij de normering van de Kijk- en luistertoetsen.

Hoe komt de normering tot stand?

Als je een Kijk- en luistertoets maakt, komt er allereerst een score uit. Een leerling heeft bijvoorbeeld 26 van de 35 opgaven goed. Dan weet je alleen nog niet welk cijfer daarbij hoort. Om die score te kunnen vertalen naar een cijfer, is een normering nodig. Om uit te leggen hoe wij dat doen, moeten we eerst wat theorie uitleggen.

De score die je behaalt op een Kijk- en luistertoets is eigenlijk een weergave van hoe goed je bent in kijkluistervaardigheid. Iemand die bij de Kijk- en luistertoets Duits havo 12 goede antwoorden heeft gegeven is minder 'Duits kijkluistervaardig' dan iemand die op die toets 30 goede antwoorden had. Wij vinden het belangrijk dat, ongeacht in welk jaar je de Kijk- en luistertoets maakt, en of je de papieren of digitale toets maakt, je dezelfde kijkluistervaardigheid moet laten zien om een voldoende te halen. Omdat wij vrij goed weten hoe moeilijk de toets is, weten we ook bij welke score je genoeg kijkluistervaardigheid hebt laten zien om een voldoende te halen. En dat is niet altijd bij dezelfde score. Hoe kan dat?

Een voorbeeld: de toets Frans vwo 2022 had 35 opgaven en je had een voldoende bij score 24. De toets Frans vwo 2023 heeft ook 35 opgaven, maar daar heb je pas een voldoende bij score 26. Dat lijkt niet eerlijk, alsof er strenger is genormeerd. Maar er is niet strenger genormeerd. De toets Frans vwo 2023 was dan iets makkelijker dan die van 2022. Als een toets makkelijker is, dan moet je meer scorepunten halen om de voldoende te 'verdienen'. Je hebt dan, in dit voorbeeld, pas met score 26 bewezen dat je de kijkluistervaardigheid hebt die wij altijd eisen om een voldoende te halen. Bij de moeilijker toets uit 2022 had je met score 24 al bewezen dat je genoeg kijkluistervaardigheid had.

Maar stel dat de toetsen Frans vwo uit 2022 en 2023 wel exact even moeilijk zijn. Dan moet de grens tussen onvoldoende en voldoende (=de cesuur) bij dezelfde score liggen. Maar als de toets van 2023 drie opgaven meer bevat, dan zal de grens tussen onvoldoende en voldoende bij de toets uit 2023 ook hoger liggen.

Samengevat: als je bij de ene toets bij score 26 een 7,1 had, betekent dat dus niet altijd dat je bij de toets van een ander jaar bij score 26 ook weer een 7,1 hebt.



We normeren nooit strenger, we versoepelen wanneer nodig

In principe leggen we dus bij elke toets de lat bij dezelfde vaardigheid als vorige jaren. Als we na de afnameweek zien dat dat prima cijfers oplevert, dan is duidelijk dat we die normering één op één over kunnen nemen. Als de toets goed gemaakt is, dan zal het landelijke gemiddelde hoger uitvallen, en/of is het percentage leerlingen dat een onvoldoende haalt lager. Maar: als we zien dat met de normering het landelijke gemiddelde opeens lager dan een 6,0 is, en/of dat meer dan 40% van de leerlingen een onvoldoende gaat scoren, dan grijpen we in. We versoepelen dan de normering: we leggen de lat iets lager, zodat er weer acceptabele, landelijke cijfers verschijnen. Dit zullen we dan wel vermelden bij het publiceren van de normering.

Wat houdt de ERK-indicatie eigenlijk in?

De Kijk- en luistertoetsen zijn geen ERK-toetsen zoals de toetsen van Goethe, Cambridge of DELF. In 2013 zijn de Kijk- en luistertoetsen door buitenlandse experts bij een zogenaamde 'standaardsetting' gekoppeld aan de ERK-niveaus. Hierdoor kunnen we op basis van het cijfer dat je gehaald op de Kijk- en luistertoets een indicatie geven op welk ERK-niveau je zit. Dit geeft dus een idee van je ERK-niveau, maar is geen bewijs ervan.

