



Technische toelichting

Samenvatting

Het College voor Toetsen en Examens (CvTE) past vanaf 2027 de manier aan waarop de grensscores voor referentieniveaus en toetsadviezen worden vastgesteld. Deze wijziging volgt uit onderzoek naar aanleiding van een onafhankelijke audit op het normeringsproces van de doorstroomtoetsen.

De gewijzigde normeringsmethodiek maakt gebruik van alle beschikbare afnamegegevens sinds 2018 en zorgt voor een stabielere normering van de doorstroomtoetsen. De effecten van deze wijziging zijn voor het eerst zichtbaar in de resultaten van de doorstroomtoets van 2027.

De gevolgen voor toetsadviezen zijn naar verwachting zeer beperkt. Bij de referentieniveaus kunnen ten opzichte van voorgaande jaren verschuivingen zichtbaar zijn. Deze verschillen zijn naar verwachting het meest zichtbaar bij Lezen 2F (zie tabellen 1 en 2 voor de verwachte effecten op basis van de afname van 2026).

De wijziging betekent niet dat leerlingen minder goed zijn gaan presteren of dat de inhoud van de referentieniveaus verandert. Wel verandert de manier waarop referentieniveaus aan prestaties worden gekoppeld.

Aanleiding

Het CvTE en Stichting Cito evalueren de normering van de doorstroomtoetsen voortdurend. Daarbij wordt regelmatig onderzocht of de gebruikte methoden nog aansluiten bij de meest actuele psychometrische inzichten en of verbeteringen mogelijk zijn.

In 2024 heeft ETS Global een onafhankelijke audit uitgevoerd op het normeringsproces van de doorstroomtoetsen. In de aanbevelingen bij deze audit wees ETS Global op de mogelijkheid om gebruik te maken van gemeenschappelijke kalibratie. Naar aanleiding hiervan heeft Stichting Cito onderzocht of deze methode toepasbaar was binnen het Nederlandse toetsstelsel en welke gevolgen invoering ervan zou hebben.

Uit dit onderzoek bleek dat een methode kan worden toegepast die gebruikmaakt van alle beschikbare afnamegegevens sinds 2018. Het jaar 2018 vormt daarbij een logisch startpunt, omdat in dat jaar voor het eerst het gezamenlijke anker werd ingezet dat nog steeds wordt gebruikt voor de koppeling tussen toetsjaren. Hierdoor kan gebruik worden gemaakt van alle beschikbare gegevens die op deze wijze met elkaar verbonden zijn.

Door alle beschikbare gegevens sinds 2018 gezamenlijk te gebruiken, wordt meer informatie benut dan bij de tot dan toe gebruikte werkwijze. Dit maakt de normering minder gevoelig voor kleine onzekerheden die kunnen optreden bij het jaarlijks overbrengen van de schaal tussen afnamejaren.

Op basis van deze bevindingen heeft het CvTE besloten de normeringsmethodiek aan te passen. De invoering vindt gefaseerd plaats. In 2026 is de nieuwe kalibratiemethode ingevoerd. In 2027 volgt daarnaast een herijking van de grensscores voor referentieniveaus en toetsadviezen.

Kalibratie en normering

*Na iedere afname worden de opgaven van de doorstroomtoets geanalyseerd. Daarbij wordt op basis van de antwoorden van leerlingen geschat hoe moeilijk opgaven zijn en hoe goed zij verschillen tussen leerlingen met verschillende vaardigheidsniveaus zichtbaar maken. Ook wordt bepaald hoe de vaardigheid van leerlingen zich verhoudt tot de moeilijkheid van de opgaven. Dit proces noemen we **kalibratie**.*

De uitkomsten van de kalibratie worden geplaatst op een gezamenlijke vaardigheidsschaal. Op deze schaal kan worden bepaald hoe de vaardigheid van leerlingen zich verhoudt tot vooraf

*vastgestelde grensscores voor referentieniveaus en toetsadviezen. Het koppelen van deze grensscores aan de vaardigheidsschaal noemen we de **normering** van de toets.*

De normering tot en met 2025

Tot en met 2025 werd de schaal stap voor stap van jaar op jaar overgebracht. Daarbij vormden de kalibratieresultaten van eerdere afnamejaren steeds het uitgangspunt voor de normering van een volgend afnamejaar.

Deze methode maakte het mogelijk om de schaal voor referentieniveaus en toetsadviezen door de jaren heen stabiel te houden. Daardoor konden grensscores op de vaardigheidsschalen voorafgaand aan de afname worden vastgesteld en was de interpretatie van resultaten voor toetsaanbieders relatief eenvoudig.

Parameterfixatie

*Tot en met 2025 werd bij de normering gebruikgemaakt van **parameterfixatie**. Daarbij werden kenmerken van ankeropgaven, zoals de geschatte moeilijkheid van een opgave, overgenomen uit een eerder afnamejaar. Deze kenmerken vormden vervolgens het uitgangspunt voor de kalibratie en normering van een nieuw afnamejaar.*

Met deze methode werd de vaardigheidsschaal stapsgewijs van jaar op jaar overgebracht. Deze werkwijze vormde de basis voor de normering binnen het huidige doorstroomtoetsstelsel. Hierdoor konden resultaten uit verschillende afnamejaren met elkaar worden vergeleken en bleef de koppeling met de bestaande referentieniveaus en toetsadviezen behouden.

Het jaar 2018 vormt hierbij een belangrijk ijkpunt. In dat jaar werd voor het eerst het gezamenlijk anker ingezet dat nog steeds wordt gebruikt voor de koppeling tussen toetsjaren. Daarnaast was de vaardigheidsschaal in dat jaar voor het laatst rechtstreeks gekoppeld aan de oorspronkelijke normstelling van de referentieniveaus. Daarmee is vastgelegd welke toetsprestaties horen bij referentieniveaus zoals 1F, 2F en 1S.

De methode maakte het jarenlang mogelijk om resultaten uit verschillende afnamejaren met elkaar te vergelijken. Een beperking van deze aanpak is echter dat kleine onzekerheden uit eerdere schattingen kunnen doorwerken in latere afnamejaren. Wanneer dit gedurende meerdere jaren gebeurt, kan de schaal geleidelijk afwijken van de oorspronkelijke positionering.

De invoering van gemeenschappelijke kalibratie in 2026

In 2026 is de eerste stap gezet in de invoering van de nieuwe normeringsmethode. Daarbij is de wijze waarop de kalibratie wordt uitgevoerd aangepast.

Vanaf 2026 wordt gebruikgemaakt van **gemeenschappelijke kalibratie**. Hierbij worden alle beschikbare afnamegegevens sinds 2018 gezamenlijk gebruikt. Hierdoor ontstaat een stabielere basis voor de vaardigheidsschaal waarop de normering van de doorstroomtoetsen is gebaseerd.

Omdat de vaardigheidsschaal met deze methode opnieuw wordt vastgesteld, kunnen de bestaande grensscores voor referentieniveaus en toetsadviezen niet zonder meer worden overgenomen. Het CvTE heeft daarom gekozen voor een gefaseerde invoering.

In 2026 is uitsluitend de nieuwe kalibratiemethode ingevoerd. De herijking van de grensscores voor referentieniveaus en toetsadviezen volgt in 2027. Om de overgang zorgvuldig te laten verlopen zijn de grensscores in 2026 zo vastgesteld dat de koppeling tussen toetsprestaties en referentieniveaus en toetsadviezen aansluit bij die van 2025.

Herijking van de grensscores in 2027

Bij de normering van de doorstroomtoets van 2027 volgt de tweede stap in de invoering van de gewijzigde normeringsmethode: de herijking van de grensscores voor referentieniveaus en toetsadviezen.

Met de invoering van gemeenschappelijke kalibratie beschikt het CvTE over een stabielere vaardigheidsschaal waarop de resultaten van de eind- en doorstroomtoetsen sinds 2018 zijn geplaatst. Hierdoor kan worden beoordeeld in hoeverre de huidige positie van de grensscores nog aansluit bij de oorspronkelijke normstelling van de referentieniveaus en toetsadviezen. Bij de doorstroomtoets van 2027 worden de grensscores herijkt, zodat deze opnieuw aansluiten bij die oorspronkelijke normstelling. Daarna worden de grensscores jaarlijks volgens deze methodiek vastgesteld.

De herijking betekent niet dat de inhoud van de referentieniveaus verandert of dat leerlingen iets anders moeten kennen of kunnen dan voorheen. Wel kan de koppeling tussen toetsprestaties en referentieniveaus of toetsadviezen veranderen.

Herijking van de grensscores

Voor de herijking wordt teruggegrepen op de oorspronkelijke normstelling van de referentieniveaus en toetsadviezen.

Voor de referentieniveaus wordt daarbij gebruikgemaakt van de wijze waarop de oorspronkelijke normstelling in 2018 was gekoppeld aan de Centrale Eindtoets. Voor de toetsadviezen wordt aangesloten bij de meest recente normstelling waarop de geldende adviesgrenzen zijn gebaseerd. Deze normstelling wordt mede onderbouwd door toelatings- en doorstroomonderzoek en kan op basis van nieuw onderzoek worden geactualiseerd.

*De grensscores worden vervolgens opnieuw gepositioneerd op de vaardigheidsschaal die voortkomt uit de gemeenschappelijke kalibratie. Hiervoor wordt gebruikgemaakt van een **equipercentiel-equivalering**. Daarbij wordt voor ieder referentieniveau en toetsadvies bepaald welk punt op de nieuwe vaardigheidsschaal overeenkomt met de geldende normstelling in het gehanteerde referentiejaar. Concreet wordt een grensscore gezocht die in het referentiejaar leidt tot hetzelfde percentage leerlingen dat het betreffende referentieniveau of toetsadvies behaalt als onder de oorspronkelijke normering.*

De herijking leidt daarmee niet tot een nieuwe inhoudelijke normstelling. De oorspronkelijke normstelling blijft het uitgangspunt; alleen de positie van de grensscores op de opnieuw vastgestelde vaardigheidsschaal wordt aangepast.

Verwachte effecten van de herijking

Om inzicht te krijgen in de mogelijke gevolgen van de herijking zijn analyses uitgevoerd op de resultaten van de doorstroomtoetsafname van 2026. Daarbij zijn de daadwerkelijk toegepaste grensscores vergeleken met de grensscores die zouden zijn toegepast na herijking.

De hieronder weergegeven verschillen geven daarmee een indicatie van de verwachte orde van grootte van de effecten. De daadwerkelijke effecten in 2027 kunnen hiervan afwijken, omdat deze mede afhankelijk zijn van de resultaten van de leerlingen die de doorstroomtoets in dat jaar maken.

Indicatieve effecten op toetsadviezen

De verwachte effecten op de landelijke verdeling van toetsadviezen zijn beperkt. De verschillen blijven voor alle toetsadviezen binnen ongeveer één procentpunt.

Tabel 1. Verdeling van toetsadviezen bij toepassing van de grensscores van 2026 en de herijkte grensscores

Toetsadvies	2026 (toegepast)	2026 (na herijking)	Vershil
pro/vmbo bb	2,6%	3,0%	0,4%
vmbo bb/kb	14,3%	15,0%	0,7%
vmbo kb/gl-tl	22,5%	22,3%	-0,2%
vmbo gl-tl/havo	29,4%	28,4%	-1,0%
havo/vwo	20,2%	19,7%	-0,5%
vwo	11,0%	11,5%	0,5%

Indicatieve effecten op referentieniveaus

Voor de referentieniveaus laten de analyses grotere verschillen zien dan voor de toetsadviezen. De omvang van deze verschillen varieert per onderdeel en referentieniveau. Op basis van de doorrekening van de afname van 2026 wordt het grootste verschil zichtbaar bij Lezen 2F.

Tabel 2. Verschillen in het aandeel leerlingen dat een referentieniveau behaalt tussen de grensscores van 2026 en de herijkte grensscores

Onderdeel	Niveau	2026 (toegepast)	2026 (na herijking)	Vershil
Lezen	1F	97,9%	96,0%	-1,9%
Lezen	2F	72,0%	65,3%	-6,7%
Rekenen	1F	89,7%	88,3%	-1,4%
Rekenen	1S	40,6%	40,3%	-0,3%
Taalverzorging	1F	93,8%	94,6%	0,8%
Taalverzorging	2F	52,6%	55,2%	2,6%

De verschillen ontstaan doordat de grensscores na de herijking op een andere plek op de vaardigheidsschaal komen te liggen. Hierdoor kan dezelfde toetsprestatie tot een ander referentieniveau of toetsadvies leiden dan onder de grensscores die in 2026 zijn gebruikt. Dit betekent niet dat leerlingen minder goed zijn gaan presteren of dat de inhoud van de referentieniveaus verandert.

Vergelijkbaarheid van resultaten

De herijking van de grensscores kan leiden tot verschillen in de resultaten ten opzichte van eerdere afnamejaren. Hierdoor moeten ontwikkelingen tussen 2026 en 2027 met voorzichtigheid worden geïnterpreteerd.

Veranderingen in percentages behaalde referentieniveaus kunnen vanaf 2027 deels het gevolg zijn van de nieuwe positionering van de grensscores en hoeven daarom niet uitsluitend samen te hangen met veranderingen in leerlingprestaties.

Vanaf de doorstroomtoets van 2027 worden de resultaten vastgesteld volgens dezelfde normeringsmethodiek. Hierdoor blijven resultaten van opeenvolgende afnamejaren vanaf 2027 onderling vergelijkbaar.



**Toetsen
primair onderwijs**

cvtetoetsenpo.nl

**Centrale examens
middelbaar
beroepsonderwijs**

examenbladmbo.nl

**Centrale- en
staatsexamens
voortgezet onderwijs**

examenblad.nl

**Staatsexamen
Nederlands als
tweede taal (Nt2)**

staatsexamensnt2.nl