NATIONAAL KENNISINSTITUUT ONDERWIJS NCO

Van M- tot E-toets: inzicht in de leergroei van het afgelopen halfjaar

Rapportage op basis van gegevens uit het leerlingvolgsysteem

door het Nationaal Cohortonderzoek Onderwijs

Technische Toelichting regulier basisonderwijs (BO)

Rapportage november 2026

Inhoud

1. Algemeen 2
1.1 Privacy 2
1.2 Data 2
1.3 Procedure 2
2. Contact 4
3. Uitleg type figuren 5
4. Extra toelichting Hoofdstuk 1 – Absolute vaardigheidsscore per jaargroep 6
5. Extra toelichting Hoofdstuk 2 – Absolute leergroei per jaargroep 8
6. Extra toelichting Hoofdstuk 3 – De leergroei per domein 11
7. Extra toelichting Hoofdstuk 4 – Leergroei uitgesplitst naar geslacht 13
8. De voorspelde referentiewaarde 15
9. Berekening infoboxen 18

1. Algemeen

Deze technische toelichting hoort bij de schoolrapportages “Van M- tot E-toets: inzicht in de leergroei van het afgelopen halfjaar”, die het Nationaal Cohortonderzoek Onderwijs (NCO) heeft uitgebracht in november 2026 over de leergroei in het reguliere basisonderwijs (BO). In deze rapportage is gekeken naar de leergroei tussen de M-toets van schooljaar 2025-2026 en de E-toets van 2025-2026 van leerlingen op drie domeinen: begrijpend lezen, spelling (niet-werkwoorden) en rekenen-wiskunde. In deze technische toelichting wordt dieper ingegaan op de gebruikte definities en de berekeningswijzen.

1.1 Privacy

De leerling- en schoolgegevens uit het NCO en de daarop gebaseerde schoolrapportages zijn vertrouwelijk. Ze worden niet gedeeld met en zijn niet toegankelijk voor het Ministerie van Onderwijs, Cultuur en Wetenschap, de Inspectie van het Onderwijs of anderen. Alle informatie wordt door het Centraal Bureau voor de Statistiek (CBS) en onderzoekers van het Nationaal Cohortonderzoek Onderwijs (NCO) in een beveiligde omgeving opgeslagen en bewerkt.

1.2 Data

Alle schoolbesturen in het primair onderwijs zijn benaderd met het verzoek om gegevens uit het leerlingvolgsysteem (LVS) beschikbaar te stellen aan het NCO. Het gaat om LVS-toetsgegevens van groep 3 t/m 8 voor de domeinen begrijpend lezen, spelling en rekenen-wiskunde, van de toetsleveranciers Boom, Bureau ICE (IEP), Cito en Diataal. Het doel is om meer inzicht te krijgen in de leergroei van uw leerlingen over een periode van het afgelopen halfjaar. In totaal hebben zich ruim 4800 reguliere scholen aangemeld waarvan er ongeveer 4250 een rapportage ontvangen[1].

Het databestand achter deze leergroei-rapportages bestaat allereerst uit de toetsgegevens van leerlingen afkomstig uit de LVS-systemen die uw school gebruikt. Voor de berekening van de voorspelde referentiewaarde worden de gegevens van het CBS gebruikt: het Stelsel van Sociaal-statistische Bestanden (SSB) en gegevens uit het Register Onderwijs Deelnemers (ROD) van Dienst Uitvoering Onderwijs (DUO). De gegevens bevatten bijvoorbeeld informatie over: burgerlijke status ouders, geboorteland kind, geboorteland ouders, samenstelling huishouden, inkomen ouders, ouders werkend/zelfstandig/uitkering etc.

De gebruikte data voor de leergroei-rapportages bevatten gegevens op het niveau van leerlingen en scholen, maar geen informatie over klassen en docenten. Het kan echter zijn dat in deze data administratieve fouten voorkomen, omdat ze bijvoorbeeld door de school per ongeluk verkeerd zijn ingevoerd. Dit zou voor een school tot verkeerde uitkomsten kunnen leiden. De onderzoekers van het NCO trachten de overduidelijke fouten uit de data te halen en eventueel te corrigeren indien mogelijk. Echter, niet alle gemaakte fouten zullen even duidelijk zijn. Mocht u twijfelen over de gegevens in uw rapportage, dan staan we u graag te woord.

1.3 Procedure

Allereerst hebben de besturen een contract getekend met het Nationaal Kennisinstituut Onderwijs (NKO) (voorheen NRO). De besturen zijn immers juridisch verantwoordelijk en ‘eigenaar’ van de data. Omdat het om niet-bijzondere persoonsgegeven gaat, is het niet nodig dat ouders hiervoor

[1] Niet alle aangemelde scholen ontvangen een rapportage omdat deze scholen niet tijdig hun data hebben geëxporteerd of omdat ze niet de juiste toetsen hebben afgenomen om de leergroei te berekenen.

toestemming geven. Wel hebben de deelnemende scholen alle ouders geïnformeerd over het voornemen om de toetsgegevens te leveren aan het NKO zodat de leergroei-rapportages vertrouwelijk aan het bestuur kunnen worden geleverd. En aan CBS met het doel te koppelen aan de NCO-dataset[2].Hiervoor zijn voorbeeldbrieven beschikbaar gesteld. Ouders zijn in de gelegenheid gesteld om daar bezwaar tegen te maken gedurende een bepaalde periode. De scholen registreren deze bezwaren in het LeerlingAdministratieSysteem (LAS). De softwareleveranciers van de LAS-en in het (s)bo (ParnasSys en ESIS) hebben hiervoor een aparte 'bezwaarknop' ingebouwd. Voordat de feitelijke upload naar het CBS en NKO plaatsvindt, moeten de scholen op een 'verzendknop' klikken. Zo verklaart de school dat alle procedures doorlopen zijn en de softwareleveranciers tot levering over kunnen gaan. De softwareleveranciers leveren vervolgens de gegevens via een beveiligd kanaal aan het NKO en CBS. Het CBS 'verrint' deze gegevens (proces van pseudonimisering) en stelt de data beschikbaar aan het NCO-team. Aan het NKO wordt geen BSN-nummer van de leerling geleverd. Wel ontvangt het NKO een leerling-id: een uniek ID voor elke leerling. Het NCO-team bewerkt deze data en baseert hier de leergroei-rapportages op.

In de losse datasets van ParnasSys en ESIS zitten gegevens die de school in hun softwaresysteem over de leerling heeft geregistreerd zoals:

- BSN nummer (CBS pseudonimiseert deze, dit wordt niet aan het NKO geleverd)
- Brinnummer
- Vestigingsnummer
- Postcode school
- LeerlingID (ParnasSys en ESIS pseudonimiseren deze)
- Inschrijfdatum op school
- Eventuele uitschrijfdatum van school
- Jaargroep
- Klasnaam
- Vaardigheidsscore
- OSOtoetscode
- Jaargroep van afname toets
- Afnamedatum toets
- Geslacht
- Geboortemaand en geboortejaar

[2] Voor meer informatie over de NCO-dataset, zie: Haelermans, C., Huijgen, T., Jacobs, M., Levels, M., van der Velden, R., van Vugt, L., van Wetten, S., (2020). Using Data to Advance Educational Research, Policy, and Practice: Design, Content, and Research Potential of the Netherlands Cohort Study on Education. *European Sociological Review 36*(4), p. 643–662, https://doi.org/10.1093/esr/jcaa027

2. Contact

Voor vragen over de inhoud van uw rapportage kunt u terecht bij het NCO, klik hier om naar de website te gaan. Ook kunt u contact met ons opnemen via info@nationaalcohortonderzoek.nl.

- De technische toelichting voor BO en SBO is hier te vinden op de website van het NCO.
- De voorbeeldrapportages voor de leergroei-rapportages BO en SBO kunt u vinden op de website van het ROA. Bekijk hier een voorbeeldrapportage.
- Op deze pagina kunt u uitleg vinden over de figuren en uitlegvideo's.
- Meer uitleg over begrippen die gebruikt worden in de rapportages is ook te vinden op de website van het NCO. Deze pagina vindt u hier.

3. Uitleg type figuren

Hieronder ziet u een overzicht van alle figuren die opgenomen zijn in de leergroei-rapportage.

Tabel 1 Overzicht van figuren

Hoofdstuk	Nr.	Soort figuur	Inhoud
H1	F.1.1	Staafdiagram	Gemiddelde absolute vaardigheidsscore begrijpend lezen per jaargroep van meest recente halfjaar
	F1.2	Staafdiagram	Gemiddelde absolute vaardigheidsscore spelling per jaargroep van meest recente halfjaar
	F1.3	Staafdiagram	Gemiddelde absolute vaardigheidsscore rekenen-wiskunde per jaargroep van meest recente halfjaar
H2	F2.1	Staafdiagram	Gemiddelde absolute leergroei begrijpend lezen per jaargroep over meest recente halfjaar
	F2.2	Staafdiagram	Gemiddelde absolute leergroei spelling per jaargroep over meest recente halfjaar
	F2.3	Staafdiagram	Gemiddelde absolute leergroei rekenen-wiskunde per jaargroep over meest recente halfjaar
H3	F3.1	Gestandaardiseerde staafdiagram	Gestandaardiseerde leergroei van M-toets 2025-2026 tot E-toets 2025-2026
H4	F4.1	Gestandaardiseerde staafdiagram	Verschillen in gestandaardiseerde leergroei voor meisjes en jongens

Staafdiagram

Een staafdiagram geeft de gemiddelde absolute vaardigheidsscores (Hoofdstuk 1) of de leergroei in vaardigheidsscores weer op een bepaald domein (Hoofdstuk 2). Deze wordt uitgesplitst naar jaargroep (bijv. Figuur 2.1). In de figuurtitel wordt aangegeven over welk domein het gaat. In de staafdiagrammen kunt u uw school vergelijken met het landelijk gemiddelde en met de voorspelde referentiewaarde. Voor Boom, Diataal en Bureau ICE (IEP) geldt dat de referentiewaarde alleen in Hoofdstuk 3 en 4 getoond kan worden, gebaseerd op de gestandaardiseerde leergroei.

Gestandaardiseerde staafdiagram

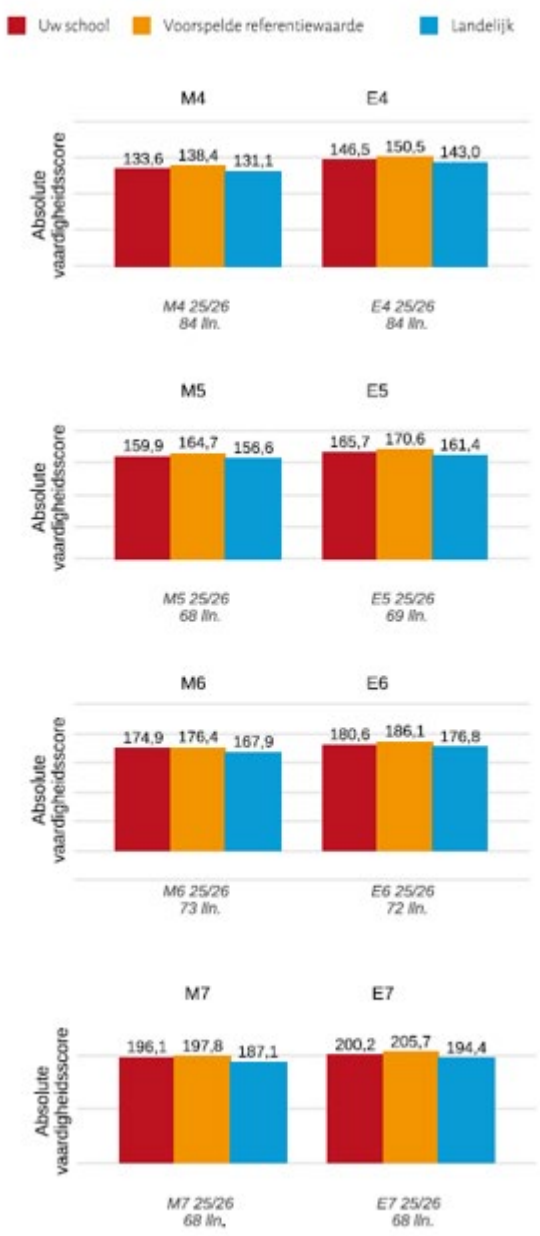
Gestandaardiseerde staafdiagrammen (bijv. Figuur 3.1) tonen de gestandaardiseerde leergroei in vaardigheidsscores. De drie domeinen (begrijpend lezen, spelling en rekenen-wiskunde) hebben alle drie een eigen schaal met andere minimale en maximale scores. Dit maakt de vergelijking tussen de domeinen moeilijker. Door te standaardiseren worden de schalen van de verschillende domeinen gelijk getrokken, waardoor een vergelijking tussen de domeinen wel mogelijk wordt.

4. Extra toelichting Hoofdstuk 1 – Absolute vaardigheidsscore per jaargroep

In dit hoofdstuk hebben we de gemiddelde absolute vaardigheidsscore per toets weergegeven. We kijken hierbij naar de absolute vaardigheidsscores ten tijden van de M-toets in schooljaar 2025-2026 en de E-toets in schooljaar 2025-2026. De vaardigheidsscores worden in hoofdstuk 2 gebruikt om de leergroei te berekenen. In dit hoofdstuk kunt u de absolute vaardigheidsscores vergelijken met het landelijk gemiddelde, en waar mogelijk met de voorspelde referentiewaarde.

Voor CITO-scholen in de G4-steden is er daarnaast een extra vergelijking mogelijk met de gemiddelde absolute vaardigheidsscore binnen de stad. Voor deze rapportage tonen we dit alleen voor scholen in Amsterdam, Utrecht en Den Haag, omdat alleen in deze steden voldoende (minstens ongeveer 67% (2/3) van het totale aantal) scholen deelnemen aan het project om een betrouwbare vergelijking te kunnen maken.

Voorbeeld: In figuur 1 laten we de absolute vaardigheidsscores voor begrijpend lezen voor de M-toets en de E-toets van de huidige groep 4 tot en met groep 7 leerlingen zien. We zien dat leerlingen in groep 6 een absolute vaardigheidsscore van 180,6 hebben behaald op de E-toets in schooljaar 2025-2026. Dit is hoger dan het landelijk gemiddelde van 176,8. Een halfjaar eerder, voor de M-toets groep 5 in schooljaar 2025-2026, hadden deze leerlingen een absolute vaardigheidsscore van 174,9. De leergroei voor deze leerlingen is dus de vaardigheidsscore op de E-toets (180,6) min de vaardigheidsscore op de M-toets een halfjaar eerder (174,9). Deze leergroei is terug te vinden in figuur 2 in het volgende hoofdstuk.



Figuur 1 Gemiddelde absolute vaardigheidsscore begrijpend lezen per jaargroep over een halfjaar (M-toets en E-toets)

5. Extra toelichting Hoofdstuk 2 – Absolute leergroei per jaargroep

In dit hoofdstuk hebben we gekeken naar de gemiddelde leergroei van de leerlingen op uw school. We vergelijken de leergroei tussen de M-toets van schooljaar 2025-2026 en de E-toets van schooljaar 2025-2026 met de leergroei tussen de M-toets van schooljaar 2024-2025 en de E-toets van schooljaar 2024-2025.

De absolute leergroei van de M-toets tot de E-toets is berekend door per schooljaar, per leerjaar en per leerling de vaardigheidsscore op de E-toets te verminderen met de vaardigheidsscore op de M-toets van een halfjaar daarvoor. Vervolgens hebben we een schoolgemiddelde berekend. Dit schoolgemiddelde kan door afronding iets afwijken van het daadwerkelijke verschil tussen de M-toets en de E-toets. Daarnaast geven we het landelijk gemiddelde en, wanneer mogelijk, ook de voorspelde referentiewaarde (zie Hoofdstuk 8 voor meer uitleg over de voorspelde referentiewaarde). Voor CITO geldt dat we het landelijk gemiddelde hebben berekend voor alle CITO-scholen die LVS-gegevens hebben geleverd. Van de overige toetsaanbieders (Boom, Diataal en Bureau ICE (IEP)) hebben we de gemiddelde leergroei gebaseerd op landelijke data, verkregen van de toetsaanbieder zelf.

Voor CITO-scholen in de G4-steden is er daarnaast een extra vergelijking mogelijk met de gemiddelde leergroei binnen de stad. Voor deze rapportage tonen we dit alleen voor scholen in Amsterdam, Utrecht en Den Haag, omdat alleen in deze steden voldoende (minstens ongeveer 67% (2/3) van het totale aantal) scholen deelnemen aan het project om een betrouwbare vergelijking te kunnen maken.

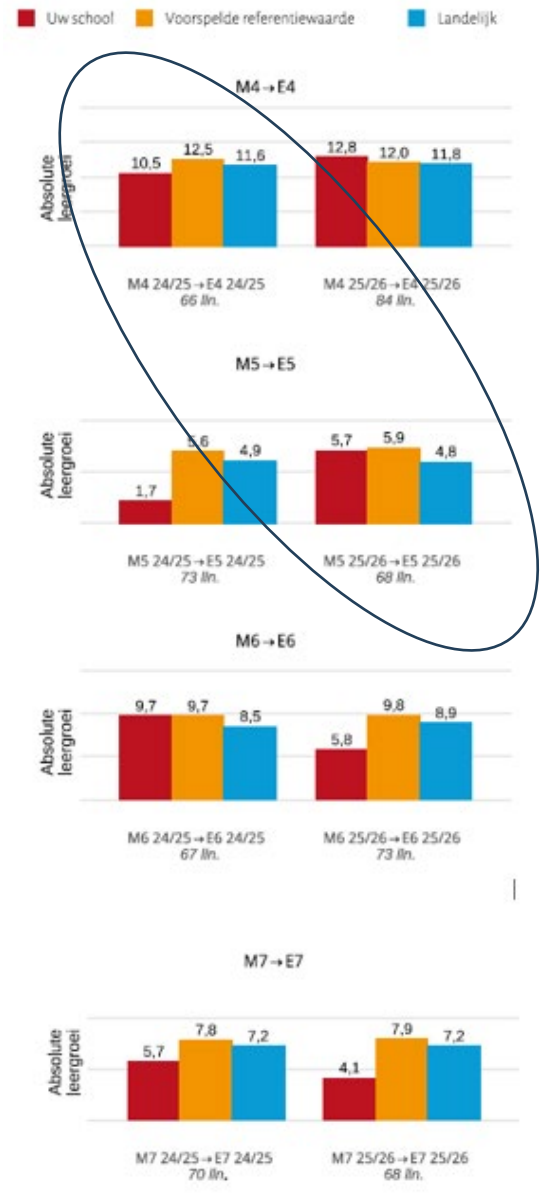
Let op!

- Voor Boom geldt dat we de landelijke gemiddeldes laten zien die gebaseerd zijn op de nieuwe LVS Begrijpend Lezen toets die in schooljaar 2022-2023 verschenen is. Mocht uw school nog de oude begrijpend lezen toets van 2012 afnemen, dan hebben wij geen leergroei voor uw school kunnen berekenen.

- Voor IEP geldt over het algemeen dat een leerling een toets krijgt die past bij het niveau van de leerling en de aangeboden lesstof. Toetsen worden dus niet altijd in één bepaalde jaargroep afgenomen. Om een leergroei-berekening te maken is het voor ons van belang dat wij weten in welke jaargroep en op welk moment een toets wordt afgenomen. Omdat dit bij veel IEP toetsen varieert, hebben wij het afnamemoment van een aantal IEP-toetsen toegewezen aan de jaargroep waarin deze het meest is afgenomen in onze data.

Voorbeeld: In Figuur 2 zien we de absolute leergroei van de M-toets naar de E-toets voor groep 4 t/m 7 voor begrijpend lezen. De score boven de balken geeft het verschil in vaardigheidsscore weer tussen de M-toets en de E-toets een halfjaar later. In dit geval zien we dat de leerlingen in groep 6 op deze school tussen de M-toets 2025/26 en de E-toets 2025/26 een leergroei doormaakten van 5,8. Dat is lager dan het landelijk gemiddelde. Een jaar eerder, tussen de M-toets 2024-2025 en de E-toets 2024-2025, was de leergroei 9,7. Dit was hoger dan het landelijk gemiddelde. De leergroei op deze school is dit schooljaar (2025-2026) afgenomen in vergelijking met de leergroei van een jaar eerder (2024-2025).

Als we het figuur horizontaal per jaargroep lezen, dan kijken we naar verschillende leerlingen. De leerlingen die vorig jaar in groep 4 zaten, zitten op dit moment niet nogmaals in groep 4. Aanvullend kunnen we de figuren ook diagonaal lezen (zie de omcirkeling in de figuren). De ovalen in het figuur geven dezelfde groep leerlingen weer: in 2024-2025 zaten deze leerlingen in groep 4, maar één jaar later (in 2025-2026) zitten zij in groep 5. Wanneer er diagonaal gekeken wordt, kan er dus binnen het cohort gekeken worden. In onderstaand figuur zien we bijvoorbeeld dat de groep 4-leerlingen in 2024-2025 een lagere absolute leergroei doormaakte (10,5) in vergelijking met het landelijk gemiddelde (11,6). Een jaar later, wanneer deze leerlingen in groep 5 zitten, zien we dat de leerlingen wel een hogere leergroei (5,7) hebben dan het landelijk gemiddelde (4,8).



Figuur 2 Gemiddelde absolute leergroei begrijpend lezen per jaargroep over een halfjaar (van M-toets naar E-toets)

6. Extra toelichting Hoofdstuk 3 – De leergroei per domein

Om te kijken in welk domein de vertraging in leergroei het grootst is geweest, maken we gebruik van de gestandaardiseerde leergroei. In de domeinen begrijpend lezen, spelling en rekenen-wiskunde worden verschillende type vaardigheidsscores gebruikt, waardoor de vaardigheidsscores van deze domeinen niet direct met elkaar vergeleken kunnen worden. In dit hoofdstuk hebben we de scores daarom omgezet zodat deze wél onderling vergeleken kunnen worden. Dit noemen we standaardiseren. We hebben de standaardisatie per domein en jaargroep uitgevoerd omdat de leergroei binnen een domein over de groepen heen niet hetzelfde is. Leerlingen in bepaalde jaargroepen hebben namelijk op bepaalde domeinen een grotere absolute leergroei dan leerlingen uit andere jaargroepen. Dit is met name belangrijk voor de kleinere basisscholen in de dataset en voor basisscholen met een scheve verdeling van jaargroepen (bijv. wanneer er op een basisschool meerdere groepen 3 zijn en één groep 7). Bij het standaardiseren nemen we alle reguliere en speciale basisscholen samen en standaardiseren we de waarde van de school op de waarde van het landelijk gemiddelde. Hierdoor zal het landelijk gemiddelde dus altijd een waarde van 0 krijgen. We drukken de gestandaardiseerde leergroei uit in standaarddeviaties, afgekort is dat ‘sd’.

Voor CITO-scholen in de G4-steden is er daarnaast een extra vergelijking mogelijk met de gemiddelde gestandaardiseerde leergroei binnen de stad. Voor deze rapportage tonen we dit alleen voor scholen in Amsterdam, Utrecht en Den Haag, omdat alleen in deze steden voldoende (minstens ongeveer 67% (2/3) van het totale aantal) scholen deelnemen aan het project om een betrouwbare vergelijking te kunnen maken.

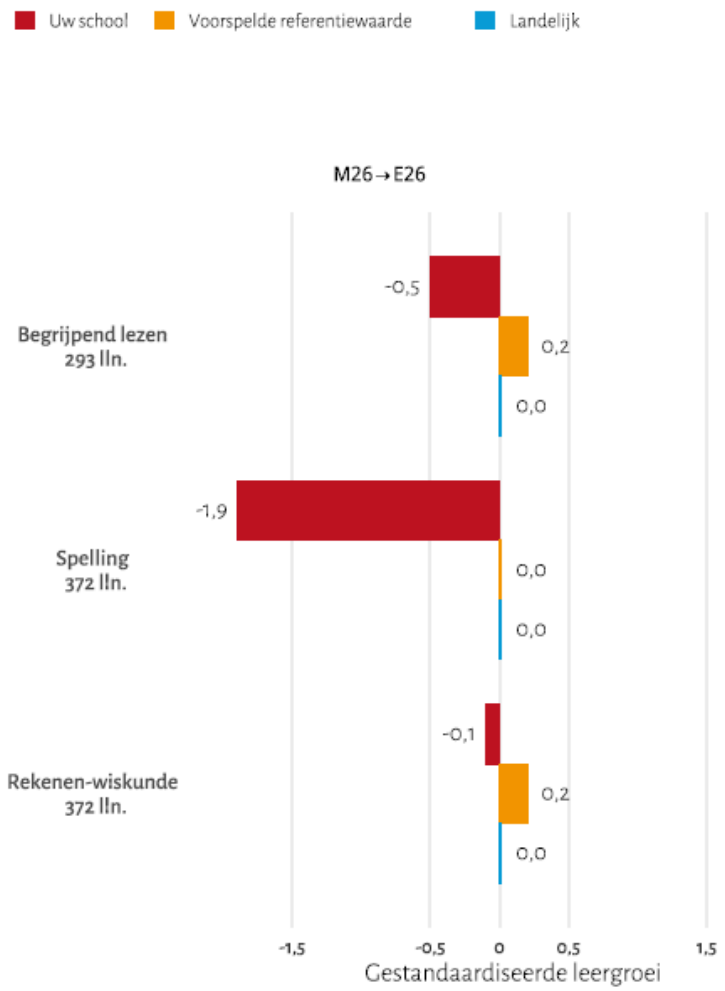
Let op!

Sinds de rapportage van november 2026 wordt de gemiddelde G4-leergroei van Hoofdstuk 3 en Hoofdstuk 4 op een andere manier berekend die nauwkeuriger is dan voorheen en beter aansluit bij de berekening voor Hoofdstuk 1 en Hoofdstuk 2.

Wanneer de gestandaardiseerde leergroei 0,5 sd of minder afwijkt van het landelijk gemiddelde, beschouwen wij dit als ‘ongeveer gelijk aan het landelijk gemiddelde’. Een grotere afwijking betekent dat scholen lager/hoger scoren dan het landelijk gemiddelde. In enkele extreme gevallen bleek de sd groter te zijn dan -2 sd of +2 sd voor een school. Wanneer dit het geval is dan hebben wij deze waarde afgekapt naar -2 sd of +2 sd.

Voorbeeld: In Figuur 3 zien we de gestandaardiseerde leergroei van alle drie de domeinen. Allereerst kan de leergroei voor de verschillende domeinen vergeleken worden met het landelijk gemiddelde en de voorspelde referentiewaarde. Zo zien we bijvoorbeeld bij spelling dat de leerlingen een lagere leergroei hebben in vergelijking met het landelijk gemiddelde en de voorspelde referentiewaarde. Bij rekenen-wiskunde behalen de leerlingen van deze school een ongeveer gelijke leergroei vergeleken met het landelijk gemiddelde en de voorspelde referentiewaarde. Daarnaast kunnen we ook tussen de domeinen vergelijken. We kunnen bijvoorbeeld de leergroei voor begrijpend lezen vergelijken met spelling. We zien

dat de leergroei voor spelling lager is vergeleken met begrijpend lezen. Door deze vergelijking te maken kunt u bepalen of een bepaald domein eventueel extra aandacht nodig heeft.



Figuur 3 Gestandaardiseerde leergroei van M-toets 2025-2026 tot E-toets 2025-2026

7. Extra toelichting Hoofdstuk 4 – Leergroei uitgesplitst naar geslacht

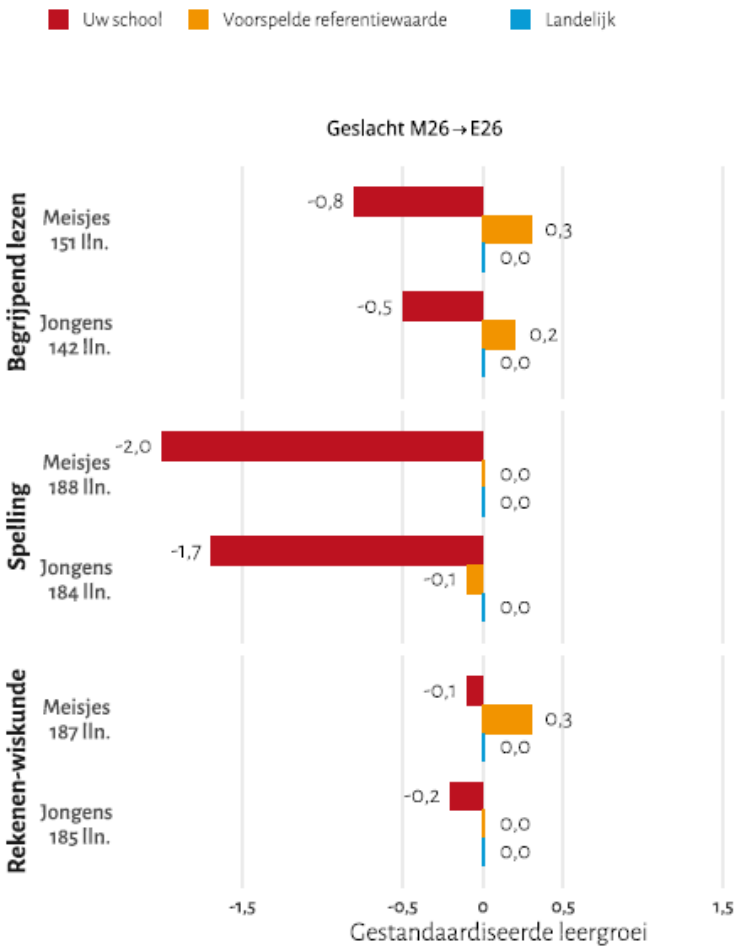
In dit hoofdstuk presenteren we een figuur die de verschillen in leergroei laten zien tussen meisjes en jongens. Het gaat hier – net als in Hoofdstuk 3 – om de gestandaardiseerde leergroei, zodat de leergroei tussen de domeinen vergeleken kan worden. Ook hier kijken we naar de leergroei van uw school, de voorspelde referentiewaarde en het landelijk gemiddelde. Het landelijk gemiddelde wordt hier berekend per subgroep.

Voor CITO-scholen in de G4-steden is er daarnaast een extra vergelijking mogelijk met de gemiddelde gestandaardiseerde leergroei binnen de stad. Voor deze rapportage tonen we dit alleen voor scholen in Amsterdam, Utrecht en Den Haag, omdat alleen in deze steden voldoende (minstens ongeveer 67% (2/3) van het totale aantal) scholen deelnemen aan het project om een betrouwbare vergelijking te kunnen maken.

Let op!

Sinds de rapportage van november 2026 wordt de gemiddelde G4-leergroei van Hoofdstuk 3 en Hoofdstuk 4 op een andere manier berekend die nauwkeuriger is dan voorheen en beter aansluit bij de berekening voor Hoofdstuk 1 en Hoofdstuk 2.

Voorbeeld: In Figuur 4 zien we dat meisjes een lagere leergroei hebben behaald op begrijpend lezen en spelling in vergelijking met het landelijk gemiddelde en de voorspelde referentiewaarde. Jongens hebben op rekenen-wiskunde een leergroei die ongeveer gelijk is aan het landelijk gemiddelde en de voorspelde referentiewaarde. Een onderlinge vergelijking laat zien dat meisjes en jongens op rekenen-wiskunde een ongeveer gelijke leergroei hebben behaald. Op begrijpend lezen behaalde meisjes een ongeveer gelijke leergroei dan jongens.



Figuur 4 Verschillen in leergroei voor jongens en meisjes (M-toets 2025-2026 tot E-toets 2025-2026)

8. De voorspelde referentiewaarde

De voorspelde referentiewaarde is een score die specifiek voor uw school is berekend, door rekening te houden met een aantal kenmerken van uw leerlingenpopulatie en kenmerken van de school zelf. De scores zijn berekend op basis van de gegevens van alle deelnemende Nederlandse basisscholen (brin-vestiging) die geregistreerd staan bij DUO, waarbij we berekenen hoe een school zou scoren die wat betreft bepaalde leerlingenpopulatie- en schoolkenmerken vergelijkbaar is met uw school. De scores van de voorspelde referentiewaarde op uitkomstmaten zijn berekend door de gegevens van alle deelnemende Nederlandse basisscholen (brin-vestiging) te vergelijken via een lineaire regressieanalyse.[3] In een regressieanalyse wordt gekeken of er een correlatie is tussen de verschillende leerlingenpopulatie-kenmerken en schoolkenmerken op een bepaalde uitkomstmaat.[4] Door middel van een regressievergelijking op alle scholen wordt berekend wat de meest waarschijnlijke score is van een school met dezelfde kenmerken als uw school.

Er wordt voor iedere berekening rekening gehouden met de volgende leerlingenpopulatiekenmerken:

- Het percentage leerlingen op de school met een niet-Europese herkomst
- Het percentage leerlingen op de school uit huishoudens met een laag gestandaardiseerd inkomen
- Het percentage leerlingen op de school waarvan minstens één ouder een hoog opleidingsniveau heeft
- Het percentage leerlingen op de school waarvan de vader werkt
- Het percentage leerlingen op de school waarvan de moeder werkt
- Het percentage meisjes op de school (we beschikken hierbij enkel over de historische gegevens die werden geregistreerd bij de geboorte van de leerling)
- Het percentage leerlingen op de school uit eenpersoonshuishoudens
- De gemiddelde gezinsgrootte van leerlingen op de school

[3] Voor meer informatie over de berekening van een lineaire regressieanalyse zie: Howitt & Cramer (2007).

[4] De bijbehorende formule is als volgt:

$Y' = \alpha + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 \ldots + \beta_{12}X_{12} + \varepsilon$

Met de bijbehorende variabelen:

Y' = voorspelde score op de uitkomstmaat

X_1= percentage leerlingen op de school met een niet-Europese herkomst

X_2= percentage leerlingen op de school uit huishoudens met een laag gestandaardiseerd inkomen

X_3= percentage leerlingen op de school waarvan minstens een ouder een hoog opleidingsniveau heeft

X_4= percentage leerlingen waarvan de vader werkt

X_5= percentage leerlingen waarvan de moeder werkt

X_6= percentage meisjes op de school (we beschikken hierbij enkel over de historische gegevens die werden geregistreerd bij de geboorte van de leerling)

X_7= percentage leerlingen op de school uit eenpersoonshuishoudens

X_8= de gemiddelde gezinsgrootte van leerlingen op de schoolX_9= totaal aantal leerlingen op de school (groep 1 t/m groep 8)

X_{10}= grondslag: openbaar (=Referentie), algemeen bijzonder, confessioneel bijzonder, gemengd

X_{11}= stedelijkheidsgraad: minder dan 500 adressen/km2 (=Referentie), tussen 500 en 1000 adressen/km2, tussen 1000-1500 adressen/km2, tussen 1500-2500 adressen/km2, meer dan 2500 adressen/km2

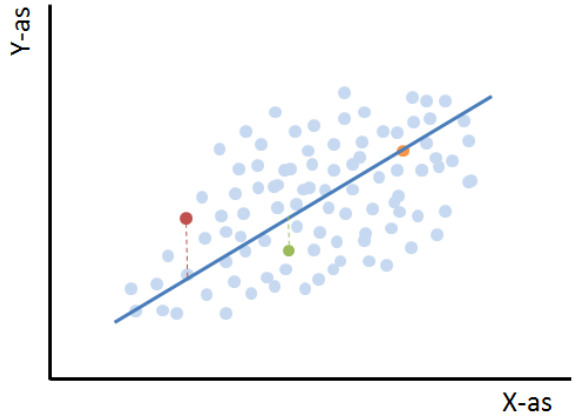
X_{12}= leerling-leraar ratio

We houden ook rekening met de volgende schoolkenmerken:

- Totaal aantal leerlingen op de school van groep 1 t/m groep 8 (schoolgrootte)
- Grondslag: openbaar, algemeen bijzonder, confessioneel bijzonder, gemengd
- Stedelijkheid van de vestigingsplaats
- Verhoudingsgetal tussen het aantal leerlingen en het aantal leraren: minder als 15 leerlingen per leerkracht, 15-20 leerlingen per leerkracht en meer dan 20 leerlingen per leerkracht (leerling-leraar ratio)
    - Leerling-leerkracht ratio is op basis van de vier-cijferige brin. Dit betekent dat hoofdvestigingen en nevenvestigingen dezelfde leerling-leraar ratio hebben.

Indien een voorspelde referentiewaarde door de regressieanalyse onder de 0% en boven de 100% is geschat, zijn deze waarden afgerond op respectievelijk 0% en 100%.

Een fictief en vereenvoudigd voorbeeld van een regressieanalyse is te zien in onderstaande grafiek. Gebaseerd op de leerlingenpopulatie- en schoolkenmerken (x-as) en de uitkomstmaat (y-as) krijgt iedere school een puntje in deze grafiek. Wanneer dit gedaan wordt voor alle deelnemende scholen in Nederland, ontstaat er een puntenwolk en kan er een regressielijn getrokken worden door alle verschillende scholen. Deze regressielijn geeft aan wat, gebaseerd op alle deelnemende scholen in Nederland, de verwachte score zou zijn op de uitkomstmaat onder controle van alle leerlingenpopulatie- en schoolkenmerken. Deze regressielijn is de score van de voorspelde referentiewaarde.



Figuur 5 Voorbeeld van een regressieanalyse

Voorbeeld: De uitkomstmaat op de y-as is de leergroei van leerlingen. Voor alle deelnemende basisscholen in Nederland weten we de leergroei van leerlingen en iedere school heeft met zijn specifieke achtergrondkenmerken hiervoor een punt gekregen in de grafiek. De bijbehorende regressielijn is de voorspelde referentiewaarde en laat zien wat de verwachte leergroei zou zijn op basis van deze specifieke combinatie van leerlingenpopulatie- en schoolkenmerken. We zien dan dat de rode school een hogere leergroei heeft dan verwacht, en dat de groene school een lagere leergroei heeft dan op basis van zijn

leerlingenpopulatie- en schoolkenmerken verwacht zou worden. De oranje school heeft precies evenveel leergroei als verwacht zou worden, want deze school ligt op de regressielijn en verschilt dus niet van de voorspelde referentiewaarde.

9. Berekening infoboxen

In de rapportage vindt u kaders met kijktips en onderzoeksresultaten. Dit laatste noemen we infoboxen. In de infoboxen met landelijke bevindingen lichten wij de resultaten toe uit ons eigen onderzoek. De infoboxen in de rapportage zijn gebaseerd op de resultaten van de factsheets over de basisvaardigheden van leerlingen in het Nederlands basisonderwijs die in 2026 door het NCO zijn uitgebracht. Dit betreft de factsheets “Basisvaardigheden factsheet 1 – Taalvaardigheid”, “Basisvaardigheden factsheet 2 – Rekenvaardigheid”, “Basisvaardigheden factsheet 3 – Leerlingkenmerken” en “Basisvaardigheden factsheet 4 – Schoolkenmerken”. In deze factsheets worden de absolute vaardigheidsscores van de Cito eindtoets (E-toets) voor begrijpend lezen, spelling, en rekenen-wiskunde bekeken in de periode tussen schooljaar 2021-2022 tot en met 2025-2026. Ook wordt er gekeken naar het percentage leerlingen dat een bepaald referentieniveau behaald heeft. De bevindingen van deze factsheets zijn samengevat in twee infoboxen in deze rapportage. De oorspronkelijke factsheets kunt u terugvinden op deze pagina op de website van het NCO.

De gegevens in de infoboxen zijn gebaseerd op de volgende leerlingenpopulatie- en schoolkenmerken:

Jaargroep: Dit betreft de groep waar de leerling in zat op het moment dat de toets werd afgenomen, zoals geregistreerd door DUO.

Opleidingsniveau ouders: Indien minstens één van de juridische ouders hoogopgeleid was (minimaal een hbo-opleiding afgerond), dan valt de leerling onder de categorie ‘hoogopgeleide ouders’ (2). Indien de hoogst behaalde opleiding van minimaal één van de ouders mbo 2-4, havo of vwo was, dan valt de leerling onder de categorie ‘gemiddeld opgeleide ouders’ (1) en indien beide ouders laag opgeleid zijn (maximaal vmbo-gt, havo/vwo onderbouw) dan valt de leerling onder de categorie ‘laag opgeleide ouders’ (0).

Gezinssamenstelling/ouderlijke structuur: Er is een opsplitsing gemaakt tussen leerlingen die in tweeoudergezinnen wonen en leerlingen die in eenoudergezinnen wonen. Onder tweeoudergezinnen verstaan wij leerlingen die wonen met beide juridische ouders of één van de juridische ouders met een partner (0). Onder eenoudergezinnen vallen leerlingen die staan ingeschreven in een huishouden met één juridische ouder, zonder partner (1). Merk op dat het hier gaat om de geregistreerde inwoners bij de Basisregistratie Persoonsgegevens (BRP) (Gemeentelijke Basis Administratie (GBA)-bestanden CBS) waardoor het kan voorkomen dat als bijvoorbeeld de partner van de moeder (nog) niet staat ingeschreven in dit huishouden, de leerling onterecht als wonend in een eenoudergezin wordt beschouwd. Leerlingen die zonder ouders wonen, bijvoorbeeld omdat ze begeleid wonen, zijn buiten beschouwing gelaten.

Geslacht: Geslacht zoals vastgelegd in de BRP. In geval van geslachtwijziging is alleen het laatst geldende geslacht vastgelegd. Personen waarvan het geslacht onbekend is worden als vrouw bestempeld in lijn de CBS werkwijze. meisje (1) of jongen (0).

Schoolweging: De schoolweging wordt berekend aan de hand van vijf omgevingskenmerken: het opleidingsniveau van de ouders, het gemiddeld opleidingsniveau van alle moeders op school, het land van herkomst van de ouders, de verblijfsduur van de moeder in Nederland en of de ouders in de

schuldsanering zitten. De schoolweging heeft een schaal tussen 20 en 40, en is een maat voor de complexiteit van de leerlingenpopulatie op een school. Een hoge weging betekent een hoge mate van complexiteit. We hebben zeven categorieën onderverdeeld; drie waarden van de schoolweging in één categorie, waarbij de hoogste vier en laagste vier waarden samengenomen zijn in (elk) een eigen categorie.

Stedelijkheid: Dit is gebaseerd op de plaats waar de school staat en het aantal adressen per km^2. We hebben dit onderverdeeld in vijf categorieën: 0) zeer sterk stedelijk (>= 2500 adressen/km^2), 1) sterk stedelijk (1500 – 2500 adressen/km^2), 2) matig stedelijk (1000 – 1500 adressen/km^2), 3) weinig stedelijk (500 – 1000 adressen/km^2) en 4) niet stedelijk (<500 adressen/km^2).

Schoolgrootte: Dit is gebaseerd op het aantal leerlingen op een school en onderverdeeld in drie categorieën: 0) kleine scholen (maximaal 140 leerlingen), 1) gemiddelde scholen (tussen de 141 en 220 leerlingen) en 2) grote scholen (minimaal 221 leerlingen). De categorieën zijn, op basis van NCO-data, in drie ongeveer gelijke groepen verdeeld.

Omdat vaardigheidsscores van verschillende toetsleveranciers vooralsnog niet eenvoudig met elkaar kunnen worden vergeleken, baseren we ons enkel op Cito-vaardigheidsscores (welke het meest voorkomen) van begrijpend lezen, spelling en rekenen-wiskunde. De oorspronkelijke factsheets kunt hier terugvinden.