

Leesvaardigheid in het vierde leerjaar in Vlaanderen

Verdiepend rapport bij PIRLS 2021
2: PIRLS in relatie tot de eindtermen

Katrijn Denies
Jonas Dockx
Nele Bleukx
Hilde Van Keer
Koen Aesaert

Januari 2024



Geadviseerde referentie voor dit rapport:

Denies, K., Dockx, J., Bleukx, N., Van Keer, H. & Aesaert, K. (2024). Leesvaardigheid in het Vierde Leerjaar in Vlaanderen. Verdiepend Rapport bij PIRLS 2021, 2: PIRLS in Relatie tot de Eindtermen. Leuven: Centrum voor Onderwijseffectiviteit en -evaluatie.

Inhoud

1. INLEIDING	4
1.1. Wat is PIRLS?	4
1.2. Het doel van dit rapport	4
1.3. Hoe PIRLS zich tot de eindtermen verhoudt	5
1.4. Toegevoegde waarde van een PIRLS-blik op de eindtermen	6
2. METHODE	7
2.1. Steekproef en representativiteit	7
2.2. Van toets tot leesvaardigheidsscore	8
2.3. Van PIRLS-score naar peilingsscore	10
2.4. Analyses	10
3. RESULTATEN	12
3.1. Proportie leerlingen die de eindtermen behalen in het vierde leerjaar	12
3.2. Evolutie in het behalen van de eindtermen in het vierde leerjaar	12
3.3. Internationale vergelijking van proporties leerlingen die de eindtermen behalen	13
4. BESLUIT	15

1. Inleiding

1.1. Wat is PIRLS?

De ‘Progress in International Reading Literacy Study’ (PIRLS) is een internationaal onderzoek van de International Association for the Evaluation of Educational Achievement (IEA). Het laat landen toe om elke vijf jaar het begrijpend-leesniveau van hun leerlingen in het vierde leerjaar van het lager onderwijs te bepalen en te vergelijken. PIRLS heeft als indirect doel om de kwaliteit van het onderwijs wereldwijd te verbeteren. Door aan beleidsmakers referentiecriteriën en feedback te bezorgen over hoe de leerlingen in hun onderwijssysteem presteren, biedt het onderzoek immers inspiratie voor curricula en beleidskeuzes.

Vlaanderen nam al deel aan PIRLS in 2006, 2016 en 2021. De eerste twee edities werden uitgevoerd door de KU Leuven. Voor de editie van 2021 sloegen de KU Leuven en de UGent de handen in elkaar. In 2018 vond bovendien - enkel in Vlaanderen - een bijkomend onderzoek plaats dat de naam PIRLS Repeat kreeg (Dockx et al., 2019). De leerlingen die in 2016 hadden deelgenomen aan PIRLS, kregen in het kader van dat onderzoek twee jaar later - dus in het zesde leerjaar - opnieuw leestoetsen aangeboden. Die bevatten deels PIRLS-toetsblokken, en deels vragen uit de periodieke peiling voor Nederlands aan het einde van het lager onderwijs. De bedoeling van dit onderzoek was in de eerste plaats om na te gaan welke PIRLS-score leerlingen op het einde van het lager onderwijs behaalden. De onderliggende vraag luidde: maken de leerlingen in de derde graad van het lager onderwijs een inhaalbeweging? In tweede instantie moest het onderzoek nagaan of PIRLS de begrijpend-leesvaardigheid van de leerlingen meet op een manier die relevant is voor

Vlaanderen. Daartoe werd onderzocht in welke mate de toetsvragen van PIRLS en deze van de peilingen - die georiënteerd zijn op de Vlaamse eindtermen - op dezelfde meetschaal zijn te situeren.

1.2. Het doel van dit rapport

Op 16 mei 2023 werden de eerste resultaten van PIRLS 2021 bekend gemaakt. De prestaties van de leerlingen worden omschreven in zowel een internationaal rapport (Mullis et al., 2023) als een Vlaams rapport (Denies et al., 2023). Na een scherpe daling tussen 2006 en 2016, bleek de begrijpend-leesvaardigheid in het Vlaamse vierde leerjaar opnieuw sterk achteruitgegaan te zijn. De gemiddelde PIRLS-score van de Vlaamse leerlingen was 511, waar die in 2016 nog 525 was en in 2006 zelfs 547. Vlaanderen zakte zo voor het eerst onder de PIRLS-mediaan.

Er zijn verschillende manieren om deze score van 511 betekenis te geven. Een vergelijking met de scores uit eerdere PIRLS-edities, met andere landen en met de mediaan ligt voor de hand. Ook een blik op de verdeling van de leerlingen over de verschillende vaardigheidsniveaus die het PIRLS-onderzoek onderscheidt (van ‘laag’ tot ‘gevorderd’) is informatief.

Eveneens interessant voor de Vlaamse context is een vergelijking van de PIRLS-resultaten met de eindtermen. Dat kan door te concentreren op de score die behaald wordt door een leerling die net de eindtermen voor lezen in het lager onderwijs behaalt (de zogenaamde ‘cesuur’). Die cesuur is door een grote groep experts (leerkrachten, pedagogisch begeleiders, lerarenopleiders, onderzoekers, e.a.) gelegd voor de peilingstoets Nederlands begrijpend lezen (zie verder). Dankzij het PIRLS Repeat-onderzoek kunnen we die cesuur vertalen

naar een precieze score op de PIRLS-schaal. We kunnen daardoor achterhalen welke proportie van de leerlingen in Vlaanderen op het einde van het vierde leerjaar de minimumdoelen voor het volledige lager onderwijs al behaalt. In 2016 bleek dit te gelden voor 53,33% van de leerlingen. In 2021 ligt dit percentage – gezien de lagere gemiddelde score die we behaalden – naar alle waarschijnlijkheid lager. We weten echter niet hoeveel lager. Het doel van dit rapport is om een antwoord te geven op die vraag.

1.3. Hoe PIRLS zich tot de eindtermen verhoudt

De eindtermen voor het Vlaamse basis-onderwijs omschrijven de kennis en vaardigheden die alle leerlingen minimaal moeten beheersen op het einde van het zesde leerjaar. Gezien het grote belang van leesvaardigheid in de studieloopbaan en in het maatschappelijk functioneren van jongeren, zijn er uiteraard ook eindtermen voor het domein lezen binnen het leergebied Nederlands. Ze richten zich naar drie verwerkingsniveaus: beschrijven, structureren en beoordelen. Ze worden weergegeven in Tabel 1.

Tot 2021 werden er op regelmatige basis periodieke peilingen gehouden die in een steekproef van scholen nagingen in welke mate deze eindtermen beheerst werden op het einde van het zesde leerjaar. De laatste peiling Nederlands in het lager onderwijs vond plaats in 2018 (Denis et al., 2019). Daaruit bleek dat 84% van de leerlingen de eindtermen voor lezen beheerste. In 2013 was dat nog 92% (Ameel et al., 2014). Vanaf 2024 worden de periodieke peilingen vervangen door Centrale Toetsen die niet langer in een steekproef van scholen, maar in alle Vlaamse scholen plaatsvinden (Steunpunt Centrale Toetsen in Onderwijs, 2022).

Tabel 1: De Vlaamse eindtermen voor lezen

Nr.	Eindterm
De leerlingen kunnen (verwerkingsniveau = beschrijven) de informatie achterhalen in:	
3.1	voor hen bestemde instructies voor handelingen van gevarieerde aard;
3.2	de gegevens in schema's en tabellen ten dienste van het publiek;
3.3	voor hen bestemde teksten in tijdschriften.
De leerlingen kunnen (verwerkingsniveau = structureren) de informatie ordenen die voorkomt in:	
3.4	voor hen bestemde school- en studieteksten en instructies bij schoolopdrachten;
3.5	voor hen bestemde verhalen, kinderromans, dialogen, gedichten, kindertijdschriften en jeugencyclopédieën.
De leerlingen kunnen (verwerkingsniveau = beoordelen) op basis van, hetzij de eigen mening, hetzij informatie uit andere bronnen, informatie beoordelen die voorkomt in:	
3.6	verschillende voor hen bestemde brieven of uitnodigingen;
3.7	reclame teksten die rechtstreeks verband houden met hun leefwereld.

PIRLS staat los van de Vlaamse eindtermen. De toetsinhoud wordt bepaald op basis van een internationaal bepaalde definitie van begrijpend lezen. Ze is dus expliciet niet gerelateerd aan eender welk nationaal curriculum. Het onderzoek meet twee verschillende leesdoelen en vier leesprocessen. De twee leesdoelen zijn 'lezen gericht op een *persoonlijke ervaring*' en 'lezen gericht op *informatieverwerving*'. In de praktijk komt dit overeen met respectievelijk het lezen van verhalen en het lezen van informatieve teksten. Los van het leesdoel wordt er binnen PIRLS een onderscheid gemaakt tussen vier leesprocessen die aangeven op welke manier leerlingen een tekst proberen te begrijpen:

- ① expliciet vermelde informatie vinden,
- ② eenvoudige conclusies trekken,
- ③ tekstinterpretatie en
- ④ tekstevaluatie.

Er is dus duidelijk sprake van inhoudelijke gelijkenissen tussen wat PIRLS meet onder de noemer 'begrijpend lezen' en wat er in de eindtermen voor lezen staat. Er zijn echter ook verschillen. Hoewel de eindtermen pas bereikt moeten worden op het einde van het zesde leerjaar, omvatten zij bijvoorbeeld amper vereisten omtrent tekstevaluatie. Een Vlaamse peilingstoets zal hooguit een beoordeling vragen van een brief, uitnodiging of reclametekst die aansluit bij de leefwereld van de kinderen, terwijl PIRLS vrij diepgaande evaluaties vraagt van verhalen en informatieve teksten. Ook wezen 10 sleutelinformanten (pedagogisch begeleiders, lerarenopleiders, onderwijsinspecteurs en onderzoekers) in kwalitatief onderzoek naar aanleiding van PIRLS 2016 op enkele andere contrasten (Hoebrechts, 2019). De peilingstoetsen bevatten bijvoorbeeld veelal kortere teksten dan PIRLS, en teksten van meer diverse aard: naast verhalen en informatieve teksten staan er ook schema's, brieven, reclameteksten en tabellen centraal. Waar PIRLS veel open vragen bevat, komen die amper voor in de peilingstoetsen voor lezen.

Toch blijkt uit het PIRLS Repeat-onderzoek dat de PIRLS-toetsen en de peilingstoetsen (die volledig gebaseerd zijn op de Vlaamse eindtermen) hoofdzakelijk een gelijkaardig concept meten. Ze zijn te herleiden naar éénzelfde onderliggende vaardigheid. Er is dus sprake van unidimensionaliteit. We zien uiteraard dat zowel de peilingsvragen (of 'items') onderling als de PIRLS-vragen onderling sterk samenhangen. Samen bekeken hangen de peilingsvragen echter ook sterk samen met de PIRLS-vragen. De samenhang tussen items gaat dus eigenlijk over de items van beide projecten heen. En: wat voorspeld wordt op basis van de PIRLS-vragen lijkt sterk op wat voorspeld wordt op basis van peilingsvragen. PIRLS en de peiling Nederlands meten lezen kortom in sterke mate hetzelfde.

1.4. Toegevoegde waarde van een PIRLS-blik op de eindtermen

Uit de voorgaande paragraaf blijkt dat de manier waarop PIRLS begrijpend lezen operationaliseert relevant is voor de Vlaamse context. De PIRLS-toetsen meten niet een helemaal ander concept dan wat we in Vlaanderen beschouwen en aanleren als 'begrijpend lezen'. PIRLS is dus ook geschikt om uitspraken te doen over de mate waarin leerlingen succesvol leren begrijpend lezen in die Vlaamse context.

Een belangrijke kanttekening is dat PIRLS plaatsvindt op het einde van het vierde leerjaar, terwijl de eindtermen bepalen wat leerlingen pas twee jaar later moeten kunnen. Echter: de lat die gelegd wordt door de eindtermen, is het absolute minimum. Ze bepaalt wat haalbaar en noodzakelijk is voor elke leerling aan het einde van het lager onderwijs. Het gaat dus niet om doelen die bijzonder uitdagend zouden mogen zijn voor de gemiddelde twaalfjarige. Daarom is het alsnog zinvol om al twee jaar eerder, op het einde van de tweede in de plaats van de derde graad, na te gaan hoeveel leerlingen deze doelen op dat moment al bereikt hebben.

Deze controle verschaft enerzijds waardevolle informatie over de kwaliteit van ons begrijpend-leesonderwijs. Als er in het vierde leerjaar nog maar amper leerlingen in de buurt komen van de absolute minimumdoelen, is dat een belangrijk signaal over de weg die we met onze leerlingen aan het afleggen zijn richting een goede begrijpend-leesvaardigheid.

Naast informatie over de mate waarin de leerprocessen in en buiten de klas succesvol verlopen, biedt een PIRLS-blik op de eindtermen ook waardevolle inzichten in de eindtermen zelf. Immers, door de lat die gelegd wordt door de eindtermen te

vergelijken met de PIRLS-schaal, krijgen we een extra indicator van hoe hoog (of laag) onze verwachtingen eigenlijk zijn. Zeker in het licht van de uitwerking van nieuwe minimumdoelen voor het Vlaams lager onderwijs, die op dit moment volop plaatsvindt, is het interessant om hier oog voor te blijven hebben. Dit inzicht kunnen we zowel via een internationale vergelijking verkrijgen als door te kijken naar de verschillende vaardigheidsniveaus die PIRLS definieert. Dus: enerzijds kunnen we achterhalen welke proportie leerlingen in andere landen wellicht onze Vlaamse eindtermen al zou halen in het vierde leerjaar. En anderzijds komen we te weten welk PIRLS-vaardigheidsniveau minimaal nodig is om te kunnen spreken van behaalde eindtermen. Op die laatste vraag bood PIRLS Repeat al een antwoord. Dat onderzoek bepaalde de peilingscesuur (en dus: de score die een leerling moet halen om voldoende beheersing van de Vlaamse eindtermen aan te tonen) naar een score op de schaal van PIRLS 2016. Die bleek te liggen tussen de ondergrensscore die PIRLS hanteert voor het middelmatige niveau en die voor het hoge niveau. Met andere woorden: de Vlaamse eindtermen voor het zesde leerjaar liggen op een leesniveau dat volgens de internationale PIRLS-onderzoekers ‘middelmatig’ is voor een leerling van het vierde leerjaar (Dockx et al., 2019).

2. Methode

2.1. Steekproef en representativiteit

Dit rapport bespreekt gegevens van de steekproef die getrokken werd voor PIRLS 2021. Om tot uitspraken over die gegevens te kunnen komen, maken we echter ook gebruik van analyses op gegevens van de steekproef voor PIRLS 2016 en van het PIRLS Repeat-onderzoek van 2018.

De steekproeven voor PIRLS 2016 en PIRLS 2021 zijn toevalssteekproeven uit alle scholen voor gewoon lager onderwijs in Vlaanderen en een deel van de scholen voor buitengewoon basisonderwijs (BuBaO). De scholen werden geselecteerd rekening houdend met het onderwijsnet en de mate waarin er sprake is van onderwijskansarmoede bij de leerlingen. Na afloop van de dataverzameling berekende de IEA voor elke leerling en elke leerkracht gewichten die gebruikt moeten worden in de statistische analyses, om eventuele scheefftrekkingen door nonrespons terug te corrigeren. Dat maakt dat de uitspraken die we doen representatief zijn: de verschillende onderwijsnetten en profielen van onderwijskansarmoede zijn in de analyses even sterk vertegenwoordigd als in het geheel van Vlaamse klassen van het vierde leerjaar.

De steekproef voor PIRLS Repeat 2018 vloeit voort uit de steekproef voor PIRLS 2016. Alle scholen die toen deelnamen, kregen de vraag om opnieuw deel te nemen met dezelfde leerlingen als zij die in 2016 hadden deelgenomen. In de plaats van in het vierde leerjaar, vond het onderzoek dus plaats in het zesde leerjaar. Er namen aan PIRLS Repeat wat minder scholen deel dan aan PIRLS 2016, maar de steekproef voor PIRLS Repeat bleek daardoor niet vertekend te zijn ten opzichte van die van 2016 (Dockx et al., 2018).

In verband met de steekproef voor PIRLS 2021 merken we op dat we in dit rapport enkel gebruikmaken van de data van de steekproef voor digitale deelname. Omdat Vlaanderen in 2021 de overstap maakte naar digitaal toetsen, waren er twee steekproeven nodig: één voor de digitale toetsing en een ruime extra steekproef voor toetsing met pen-en-papier. Die extra steekproef hielp ons om de 'brug' te maken tussen de toetsing met pen-en-papier in 2016 en de volledig digitale toetsing in 2026. De data van de 'brug'-steekproef werden niet gebruikt voor de inhoudelijke analyses in het internationale rapport en bijgevolg ook niet voor het initiële Vlaamse rapport. Ook in dit verdiepende rapport hanteren we deze aanpak.

In onze analyses betrekken we de gegevens van ongeveer 5 000 leerlingen per PIRLS-editie: 5 198 in 2016, 4 615 in 2018 en 5 114 in 2021.

2.2. Van toets tot leesvaardigheidsscore

Leesvaardigheidstoetsen. PIRLS doet een beroep op leestoetsen: om na te gaan hoe leesvaardig leerlingen zijn. Het gaat om gestandaardiseerde toetsen. Ze werden digitaal (in 2021) of met pen-en-papier (in 2016 en 2018) afgenomen bij voltallige klassen. De leerlingen werden in 2021 ondersteund door een externe, centraal opgeleide begeleider, vooral om de ICT-component in goede banen te leiden. In 2016 en 2018 hield de klasleerkracht toezicht.

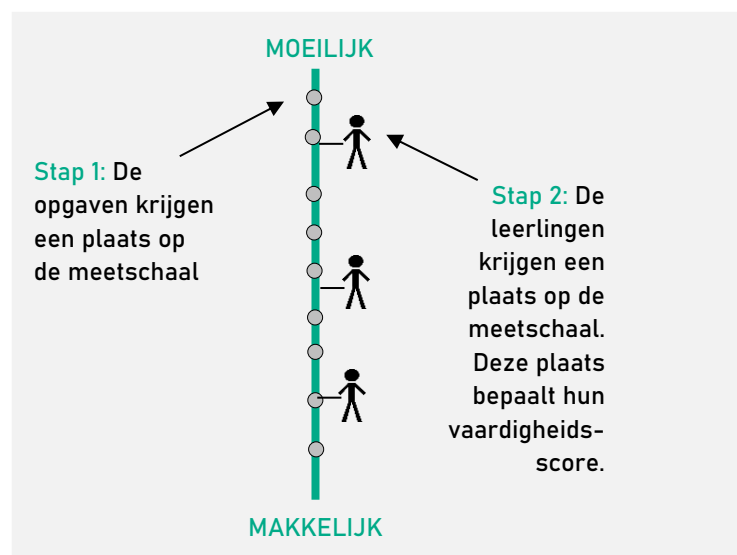
Voor reguliere PIRLS-cycli bestaat elke toets uit twee toetsblokken. Een toetsblok is een geheel van één verhalende of informatieve tekst met daarbij een reeks open vragen en meerkeuzevragen. Voor PIRLS Repeat bestonden de toetsen zowel uit PIRLS-toetsblokken als uit toetsvragen uit de peiling Nederlands begrijpend lezen van 2018. Door toetsvragen van beide studies op te nemen

konden we waardevolle inzichten verwerven in hoe PIRLS en de peiling zich ten opzichte van elkaar verhouden.

Scoring van de toetsen. De antwoorden van de leerlingen werden gescoord aan de hand van een sleutel (voor de meerkeuzevragen) of een gestandaardiseerde scoringsgids (voor de open vragen). Dat resulteerde in databestanden met score 0 voor een foutief antwoord en een hogere score voor een correct antwoord. Doorgaans was deze score 1, maar voor sommige open vragen werd een onderscheid gemaakt tussen gedeeltelijk correcte antwoorden en volledig correcte antwoorden waardoor ook score 2 en soms 3 toegekend konden worden.

Bepalen van een vaardigheidsscore. Volgens werden op de resulterende databestanden analyses toegepast volgens de principes van de item-responstheorie (IRT). Centraal in dit proces staat de meetschaal. Dat is een rangschikking van alle opgaven naargelang hun moeilijkheidsgraad. Die moeilijkheidsgraad wordt statistisch bepaald door te kijken naar de gegeven antwoorden. Items die minder vaak goed beantwoord zijn, zijn moeilijker en staan dus hoger op de meetschaal.

Figuur 1: Het IRT-principe om leesvaardigheidsscores te bepalen



Vervolgens analyseren we voor elke leerling de antwoorden. Elke leerling krijgt daarbij een bepaald vaardigheidsniveau toegekend. Dat vaardigheidsniveau reflecteert de statistische kans om een bepaalde opgave juist op te lossen. We kunnen de leerlingen dus weergeven op dezelfde meetschaal als de opgaven, in toenemende mate van vaardigheid. De positie van elke leerling op de meetschaal vertaalt zich in een leesvaardigheidsscore. Hoe groter de leesvaardigheidsscore van de leerling, hoe sterker hij of zij is in begrijpend lezen. Figuur 1 illustreert deze werkwijze.

Het werken met IRT-analyses laat ons toe om uitspraken te doen over hoe de scores op één PIRLS-toets (bijvoorbeeld uit 2016) zich verhouden tot de scores op een andere PIRLS-toets (bijvoorbeeld uit 2021). Dat kan dankzij overlap in beide toetsen. Doordat we in elke PIRLS-editie niet enkel nieuwe toetsblokken, maar ook toetsblokken die al eens eerder gebruikt zijn inzetten, kunnen we de meetschaal voor PIRLS 2021 immers ‘verankeren’ in de meetschaal voor eerdere PIRLS-edities. In het kader van PIRLS Repeat werden, zoals eerder aangegeven, zelfs PIRLS-toetsblokken gemengd met toetsblokken uit de peiling begrijpend lezen. Die mengeling vormt de sleutel die toegang biedt tot uitspraken over hoe PIRLS-resultaten zich verhouden tot de eindtermen.

Rekening houden met de overstap naar digitaal toetsen. PIRLS 2021 maakte voor het eerst gebruik van digitale toetsen. In dit rapport maken we echter ook gebruik van onderzoeksresultaten die gebaseerd zijn op PIRLS 2016 en PIRLS 2018, die nog steunden op toetsen met pen-en-papier. Dat verschil in toetsmodus maakt een rechtstreekse vergelijking tussen de scores minder vanzelfsprekend. De IEA zorgde voor een equivalering aan de hand van Randomly Equivalent Sample Linking (RESL).

In de RESL-benadering is een sleutelrol weggelegd voor de deelname met pen-en-papier van een grote, representatieve groep leerlingen terwijl een gelijkaardige (en nog grotere) groep leerlingen digitaal deelnam. In elk land dat voor het eerst deelnam aan de digitale toetsen, werden papieren toetsen voorgelegd aan een steekproef van leerlingen die op dezelfde manier (volgens dezelfde stratificatie) uit precies dezelfde doelgroep getrokken werden. Aan beide toetsmodi nam dus een *equivalente* groep leerlingen deel. In het kader van PIRLS 2021 werd eerst een meetschaal opgesteld met de toetsvragen die met pen-en-papier werden afgenomen. Die meetschaal vormt de basis: de moeilijkheidsgraad van de items wordt vastgeklekt. Vervolgens worden op deze meetschaal de digitale versies van dezelfde items toegevoegd, waarbij we het IRT-model opleggen om rekening te houden met de oorsprong van de scores uit een equivalente steekproef. Kort samengevat is het principe als volgt: een digitaal item dat door 65% van de leerlingen goed beantwoord werd, krijgt dezelfde moeilijkheidsgraad toegekend als een ‘papieren’ item dat door 65% van de leerlingen goed beantwoord werd. We veronderstellen immers dat de beide deelnemersgroepen even vaardig zijn. De moeilijkheidsgraad van de pen-en-papier-items en hun digitale equivalent worden vervolgens opnieuw vastgeklekt. Door de leerlingen vervolgens op de totale, gemeenschappelijke meetschaal te plaatsen, bekommen we vaardigheidsscores voor de digitale deelnemers die op betrouwbare wijze vergelijkbaar zijn met de vaardigheidsscores voor leerlingen die deelnamen met pen-en-papier. Bijgevolg zijn dit ook vaardigheidsscores die vergelijkbaar zijn met 2016, ondanks de introductie van digitale toetsen.

Betekenis van de vaardigheidsscore. Het IRT-proces leidde tot een set van vijf ‘plausibele’ leesvaardigheidsscores per leerling. In de analyses wordt veelal rekening gehouden met elk van deze vijf scores. Hoe groter de leesvaardigheidsscore van de leerling, hoe sterker hij of zij is in begrijpend lezen.

2.3. Van PIRLS-score naar peilingsscore

De analyses die we hier bespreken steunen in sterke mate op een cruciale bevinding die voortkwam uit PIRLS Repeat.

Zoals eerder aangegeven werd, bleek dat de toetsvragen van PIRLS en de toetsvragen van de peiling op één onderliggende meetschaal passen. Dat beide onderzoeken dezelfde vaardigheid meten en dat we één meetschaal konden maken dankzij PIRLS Repeat, maakt het mogelijk om te berekenen waar de cesuur van de peiling voor Nederlands lezen zich op de meetschaal van PIRLS bevindt. Op basis van een reeks complexe analyses bleek dat de cesuur voor leesvaardigheid in de peiling Nederlands van 2018 (0.10 op de peilingsschaal), overeenstemt met een waarde van 520.43 op de meetschaal van PIRLS 2016 (zie Dockx et al., 2018 voor meer technische informatie).

2.4. Analyses

Vergelijkingen tussen meerdere PIRLS-cycli mogelijk maken, is een expliciet doel van het PIRLS-onderzoek. De PIRLS-schalen voor verschillende edities zijn dus zo veel mogelijk geëquivalet. Toch is het altijd mogelijk dat er minieme verschuivingen plaatsvinden. Zo werd er voor PIRLS 2021 een grote operatie opgezet om de deelnames met pen-en-papier te equivaleren met de digitale deelnames, én werkte de IEA een nieuwe aanpak uit om in 2021 om te gaan met opengelaten vragen (‘missing values’). Het is niet onwaarschijnlijk dat dit ervoor zorgt dat

de score die op de PIRLS 2016-schaal de peilingsscesuur weergeeft, niet exact dezelfde is als de score op de PIRLS 2021-schaal die wijst op beheersing van de eindtermen.

Om na te gaan of we de cesuurscore voor de PIRLS 2016-schaal exact kunnen overzetten op de PIRLS 2021-schaal, berekenden we eerst of de PIRLS 2016-leerlingen met de antwoorden die ze in 2016 gaven, exact dezelfde vaardigheidsscores als in 2016 zouden behalen als ze hadden deelgenomen in 2021. We stelden een correlatie van .93 vast tussen hun werkelijke scores uit 2016 en de gesimuleerde scores voor 2021. Dat is zeer hoog, maar niet perfect. Dat hoeft ook niet te verbazen: het feit dat er in internationaal vergelijkend onderzoek met vijf willekeurig geselecteerde ‘plausible values’ gewerkt wordt in de plaats van met één puntscore per leerling, kan bijvoorbeeld al bijdragen tot kleine verschillen. Maar mogelijk speelt de nieuwe behandeling van open gelaten vragen in 2021 toch ook een rol. Daarom passen we in functie van dit rapport eerst nog een transformatie toe, om vanuit de PIRLS 2016-score die overeenkomt met de cesuur, uit te komen op de PIRLS 2021-score die overeenkomt met de cesuur voor het beheersen van de Vlaamse eindtermen voor Nederlands begrijpend lezen. We draaien zo het minieme ruis terug dat de IEA ook al beschreef in het technisch rapport bij PIRLS 2021 (von Davier et al., 2023).

We gaan vervolgens na welke proportie van de leerlingen deze cesuurscore bereikt. Zo bekomen we een degelijke benadering van het aandeel leerlingen dat op het einde van het vierde leerjaar al de eindtermen behaalt. We gebruiken hiervoor de *sampling weights* en de vijf *plausible values* die door de IEA berekend werden. Vervolgens herhalen we deze berekening op de datasets van de overige EU-15-landen die deelnamen en van

de drie hoogst presterende landen. Dat doen we om onze bevinding voor Vlaanderen in een internationaal perspectief te plaatsen. Uiteraard gaat het slechts om een projectie: de bepaling van de peilingscesuur op de peilingsschaal gebeurde uitsluitend op Vlaamse data, waardoor we in principe enkel optimaal betrouwbare uitspraken over de Vlaamse context kunnen doen. Toch kan deze werkwijze een goede indicatie opleveren van het aandeel leerlingen in andere landen dat de Vlaamse eindtermen wellicht al behaalt na vier jaar formeel leesonderwijs.

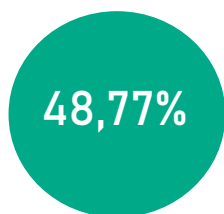
3. Resultaten

Dit hoofdstuk bespreekt onze bevindingen omtrent de beheersing van de eindtermen in het vierde leerjaar, op basis van de data van PIRLS 2021.

Dit zijn de vragen die we in dit hoofdstuk beantwoorden:

- > Welke proportie van de leerlingen die deelnamen aan PIRLS 2021 in Vlaanderen beheerst de eindtermen voor lezen in het lager onderwijs?
- > Hoe verhoudt dit zich tot de resultaten uit 2016?
- > Hoe verhoudt dit zich tot de resultaten van andere deelnemende landen?

3.1. Proportie leerlingen die de eindtermen behalen in het vierde leerjaar

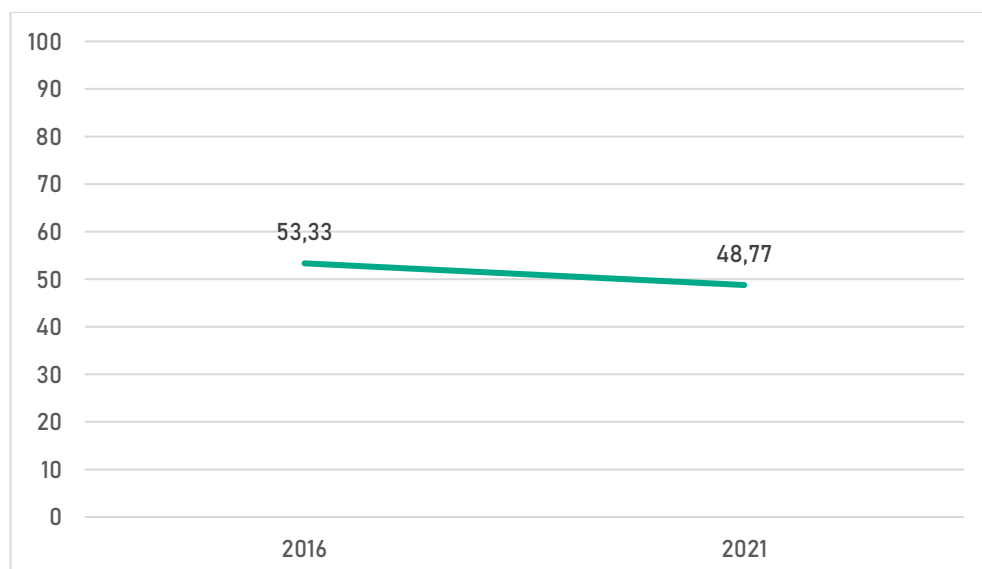


Uit onze analyses blijkt 516,51 de score te zijn die leerlingen moeten behalen op de PIRLS 2021-schaal om aan te tonen dat ze de eindtermen voor lezen in het lager onderwijs behalen. Rekening houdend met de nodige gewichten, behaalt in het Vlaamse vierde leerjaar 48,77% van de leerlingen deze score. We kunnen dus stellen dat in 2021 wellicht een kleine helft van de kinderen op het einde van het vierde leerjaar de eindtermen voor lezen in het lager onderwijs al beheerste.

3.2. Evolutie in het behalen van de eindtermen in het vierde leerjaar

Zoals we al vermoedden op basis van de gedaalde gemiddelde PIRLS-score, blijkt in 2021 een kleinere proportie van de leerlingen over de lat van de eindtermen te springen dan in 2016. In 2016 ging het namelijk nog om 53,33% van de leerlingen in het vierde leerjaar.

Figuur 2: Proportie leerlingen in het vierde leerjaar die de eindtermen voor lezen in het basisonderwijs halen: vergelijking tussen PIRLS 2016 en PIRLS 2021



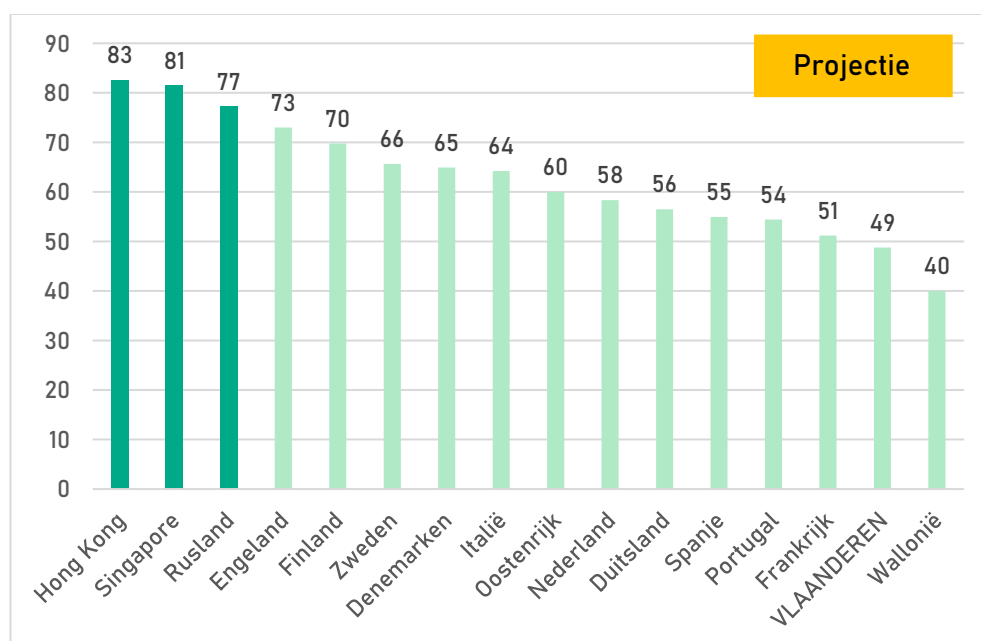
Het verschil met 2016 lijkt misschien eerder klein. Echter, in 2021 zaten er ongeveer 78 220 leerlingen in ons ‘sampling frame’ voor PIRLS (d.i. het aantal leerlingen in het vierde leerjaar van het gewoon, Nederlandstalig lager onderwijs, plus een schatting van het aantal leerlingen dat in het buitengewoon basisonderwijs leesles krijgt op een niveau dat ongeveer dat van het vierde leerjaar is). Uitgaande van dat aantal gaat het toch om meer dan 3 500 kinderen die in 2016 wél al op het einde van de tweede graad de eindtermen voor het lager onderwijs behaalden, maar in 2021 niet meer. In absolute cijfers wordt dus een behoorlijk grote groep kinderen getroffen.

3.3. Internationale vergelijking van proporties leerlingen die de eindtermen behalen

Uit het basisrapport van PIRLS 2021 bleek al dat de gemiddelde leesvaardigheid in het Belgische vierde leerjaar minder goed is dan in de andere EU-15-landen die deelnamen. Enkel in Wallonië liggen de gemiddelde scores nog lager dan in Vlaanderen. De kloof met de drie hoogst presterende landen (Hong Kong, Singapore en Rusland) is groot. We kunnen ons dan afvragen in welke mate het aandeel leerlingen dat daar over de ‘peilingslat’ zou springen, verschilt van onze 48,77%.

In principe kunnen we enkel uitspraken kunnen doen over de verhouding tussen de PIRLS-schaal en de peilingschaal voor lezen bij leerlingen *in het Vlaamse onderwijs*. Toch kunnen we ook schatten wat dit zou betekenen voor andere landen. We veronderstellen dan dat de verhouding tussen de peilingschaal en de PIRLS-schaal ook opgaat voor de populaties in deze andere landen. Dit doen we nadrukkelijk met het nodige voorbehoud: we kunnen er zeker niet van uitgaan dat onze veronderstelling klopt omdat de leerlingen in deze andere landen nooit peilingstoetsen hebben gemaakt. We maken de projectie enkel om didactische redenen, om meer inzicht te geven in hoe de prestaties van leerlingen over verschillende landen binnen PIRLS 2021 zich tot elkaar verhouden.

Figuur 3: Schatting onder voorbehoud van de proportie leerlingen in het vierde leerjaar die de eindtermen voor lezen in het Vlaamse basisonderwijs halen, voor de EU-15-landen waarvoor data van het einde van Grade 4 beschikbaar zijn en de drie hoogst scorende landen in PIRLS 2021



Figuur 3 geeft onze projectie weer van de proportie leerlingen in de vergelijkingslanden en landen uit de top-3, die de Vlaamse eindtermen voor lezen in het basisonderwijs zouden behalen. Wallonië wordt daarbij als aparte vergelijkingsgroep beschouwd omdat Wallonië ook apart van Vlaanderen deelnam aan PIRLS. In Wallonië ligt de proportie leerlingen die de Vlaamse eindtermen voor lezen in het basisonderwijs wellicht zou behalen met vier op de tien leerlingen beduidend lager dan in Vlaanderen. De resultaten suggereren echter dat in elk van de andere onderzochte landen een grotere groep leerlingen de eindtermen zou beheersen. Waar dit in Vlaanderen net geen vijf op de tien leerlingen is, gaat het in Oostenrijk om zes op tien, in Finland om zeven op tien en in Singapore en Hong Kong zelfs om meer dan acht op tien. Nogmaals: dit zijn slechts schattingen, maar ze illustreren wel de grote verschillen die er bestaan tussen Vlaanderen en verschillende andere landen.

4. Besluit

In 2019 leverde het rapport van het PIRLS Repeat-onderzoek voor het eerst een indicatie op van de mate waarin de Vlaamse leerlingen al op het einde van de tweede graad van het lager onderwijs de eindtermen voor lezen aan het einde van het basis-onderwijs beheersen. In dit rapport herhalen we deze berekening op basis van de PIRLS-data van 2021. Mogelijk zullen de Centrale Toetsen over enkele jaren gelijkaardig onderzoek mogelijk maken over de nieuwe minimumdoelen die op dit moment uitgewerkt worden.

Over de huidige eindtermen is dit rapport alvast duidelijk: in 2021 beheerst ongeveer 49% van de leerlingen in het Vlaamse vierde leerjaar deze, en dat is minder dan in 2016 (53%). Het verschil in procentpunten lijkt niet bijzonder groot, maar het gaat alsnog om een behoorlijk groot aantal leerlingen. We kunnen het extra aantal leerlingen dat de eindtermen voor begrijpend lezen niet beheerst in 2021 schatten op ongeveer 3 500.

Over de mogelijke oorzaken van deze daling zullen we het in dit rapport niet hebben. Daarvoor verwijzen we naar ons basisrapport en andere verdiepende rapporten die de gemiddelde PIRLS-scores onder de loep nemen.

Wel dringt er zich een reflectie op over wat de cijfers in dit rapport willen zeggen over de huidige Vlaamse eindtermen voor Nederlands begrijpend lezen in het basisonderwijs. Het lijkt er namelijk op dat die de lat vrij laag leggen. De drempelscore om te kunnen spreken van beheersing van deze eindtermen, ligt op een niveau dat volgens de internationale PIRLS-onderzoekers ‘middelmatic’ is voor leerlingen in het *vierde* leerjaar, terwijl de eindtermen bedoeld zijn voor leerlingen in het *zesde* leerjaar. Bovendien

suggereert onze projectie naar het buitenland dat deze drempelscore door tot wel 8 op de 10 leerlingen al behaald kan worden na vier jaar formeel leesonderwijs. We hoeven ons voor alle duidelijkheid niet volledig te spiegelen aan bijvoorbeeld Hong Kong. De onderwijscontext in het buitenland is de Vlaamse niet en we moeten niet noodzakelijk ambiëren om die volledig over te nemen. Maar het toont wel aan dat het in principe mogelijk is om op vier jaar tijd met een grote meerderheid van de leerlingen al verder te vorderen dan wat pas twee jaar later minimaal vereist wordt in Vlaanderen.

Dat meer dan de helft van de leerlingen in Vlaanderen de eindtermen nog niet beheersen op het einde van het vierde leerjaar, is op papier geen probleem. Het Vlaamse onderwijsbeleid verwacht expliciet pas op het einde van het zesde leerjaar dat de eindtermen beheerst worden. Er is dus nog tijd. Dat is zelfs één van de redenen waarom PIRLS Repeat er is gekomen. De hypothese was dat de achterstand die onze leerlingen in het vierde leerjaar hebben ten opzichte van andere landen ingelopen wordt in derde graad. De focus in de tweede graad, zo werd gedacht, ligt meer op andere domeinen, maar in de derde graad zou er veel meer ingezet worden op begrijpend lezen. Dat bleek niet zo te zijn. De leerlingen boekten in de derde graad net minder leerwinst dan we zouden kunnen verwachten op basis van verschillende bronnen (Dockx et al., 2019). Mede daarom kunnen we toch stellen dat de daling in de tussenstand in het vierde leerjaar ons zorgen moet baren.

PIRLS 2021 bracht geen goed nieuws over het gemiddelde begrijpend-leesniveaue in het Vlaamse vierde leerjaar, en – zo blijkt nu – ook geen goed nieuws over de proportie leerlingen die de eindtermen in dat leerjaar alvast beheerst. Naar onze mening moet dit er vooral voor zorgen dat de ambities weer

stijgen. We weten dat het vroeger beter kon, en we weten dat het in andere landen beter kan. Eerder dan de lat lager te leggen op basis van de huidige ervaringen, moeten we kijken naar het grote potentieel dat aanwezig is. De ambities van de nieuwe minimumdoelen moeten minstens even hoog liggen als die van de huidige eindtermen. Niet alleen wijst het voorliggende onderzoek uit dat dat allicht een haalbare kaart is. Het is ook eenvoudigweg nodig. Begrijpend lezen is een cruciale vaardigheid voor leerlingen: ze leren door te lezen en ze nemen deel aan een maatschappij waar veel informatie, instructies en discussies een schriftelijke vorm aannemen. We zijn het daarom aan onze leerlingen verschuldigd om onze vergrote aandacht voor evidence-informed begrijpend-leesonderwijs vast te houden, en om voor elk van hen een zo goed mogelijk leesbegrip na te streven.

Referenties

Ameel, E., Van Nijlen, D., Denis, J., Crynen, M., Van Gorp, K., Janssen, R. (2014). *Peiling Nederlands (Lezen - Luisteren) in het basisonderwijs - Eindrapport*. Leuven: KU Leuven, Steunpunt Toetsontwikkeling en Peilingen.

Denies, K., Bleukx, N., Pelgrims, L., Laga, J., Van Steertegem, K., Dockx, J., Vanbuel, M., Van Keer, H. & Aesaert, K. (2023). *Leesvaardigheid in het vierde leerjaar in Vlaanderen: Resultaten van PIRLS 2021 in internationaal vergelijkend perspectief*. Leuven: KU Leuven, Centrum voor Onderwijseffectiviteit en -evaluatie.

Denis, J., Janssen, R., Aesaert, K. (2019). *Peiling Nederlands lezen, luisteren en schrijven in het basisonderwijs 2018 - Brochure*. Leuven: KU Leuven, Steunpunt voor Toetsontwikkeling en Peilingen.

Dockx, J., Van Landeghem, G., Aesaert, K., Van Damme, J., & De Fraine, B. (2019). *Begrijpend lezen van het vierde naar het zesde leerjaar. Herhalingsmeting van PIRLS in 2018 vergeleken met PIRLS 2016*. Leuven: KU Leuven, Centrum voor Onderwijseffectiviteit en -evaluatie.

Hoebrechts, H. (2019). *Eindtermen lezen of internationale leesstandaarden?* [Ongepubliceerde masterproef]. Leuven: KU Leuven.

Mullis, I. V. S., von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021 International Results in Reading*. Boston: Boston College, TIMSS & PIRLS International Study Center.

Steunpunt Centrale Toetsen in Onderwijs (2022). *Toetsen voor onderwijsontwikkeling. Krachtlijnen voor de centrale toetsen in Vlaanderen*. Gent: Steunpunt Centrale Toetsen in Onderwijs.

von Davier, M., Mullis, I. V. S., Fishbein, B., & Foy, P. (Eds.). (2023). *Methods and Procedures: PIRLS 2021 Technical Report*. Boston: Boston College, TIMSS & PIRLS International Study Center.

Medewerkers

PIRLS 2021 in Vlaanderen werd uitgevoerd onder supervisie van:

Prof. dr. Koen Aesaert (KU Leuven)

Prof. dr. Hilde Van Keer (UGent)

Dr. Katrijn Denies (KU Leuven)

Het onderzoeksteam bestond uit:

Nele Bleukx

Catharina Custers

Katrijn Denies

Antonis Dervenis

Jonas Dockx

Ilka Fidlers

Jana Laga

Lore Pelgrims

Marieke Vanbuel

Kim Van Steertegem

De internationale projectaansturing was in handen van:

International Association for the Evaluation of Educational Achievement (IEA)

TIMSS & PIRLS International Study Center, Boston College

Dank:

Dit rapport steunt volledig op de waardevolle inbreng van honderden leerlingen. Onze grote dank gaat uit naar hen en hun scholen en leerkrachten, voor hun tijd en inzet.

