

Leesvaardigheid in het vierde leerjaar in Vlaanderen

Verdiepend rapport bij PIRLS 2021
1: Verschillen tussen netten



Jonas Dockx
Katrijn Denies
Lore Pelgrims
Hilde Van Keer
Koen Aesaert

30 mei 2023

Geadviseerde referentie voor dit rapport:

Dockx, J., Denies, K., Pelgrims, L., Van Keer, H. & Aesaert, K. (2023). Leesvaardigheid in het vierde leerjaar in Vlaanderen. Verdiepend rapport bij PIRLS 2021, 1: verschillen tussen netten. Leuven: Centrum voor Onderwijseffectiviteit en -evaluatie.

Inhoud

1. INLEIDING	1
1.1. Wat is PIRLS?	1
1.2. Het doel van dit rapport	1
2. METHODE	2
2.1. Steekproef en representativiteit	2
2.2. Bepalen van een vaardigheidsscore	4
2.3. Bepalen van achtergrondkenmerken	4
2.4. Analyses	4
2.5. Netten vergelijken	5
3. RUWE VERSCHILLEN TUSSEN NETTEN	6
3.1. Situatie in 2021	6
3.2. Evolutie ten opzichte van 2016	7
4. VERSCHILLEN TUSSEN NETTEN NA CORRECTIE VOOR ACHTERGRONDKENMERKEN	9
4.1. Verantwoording	9
4.2. Variabelen in het nettomodel	10
4.3. Verschillen tussen de netten in 2016 en 2021 in het 'nettomodel'	11
5. CONCLUSIE	12

1. Inleiding

1.1. Wat is PIRLS?

PIRLS staat voor 'Progress in International Reading Literacy Study'. Het is een internationale studie naar de leesvaardigheid van leerlingen in het vierde leerjaar van het lager onderwijs. Tientallen landen werken elke vijf jaar samen om het begrip- en leesniveau van hun leerlingen te bepalen en te vergelijken. Het onderzoek wordt internationaal aangestuurd door de International Association for the Evaluation of Educational Achievement (IEA). Doorheen de jaren is PIRLS uitgegroeid tot dé internationale barometer voor begrip- en leesniveau in het lager onderwijs.

Vlaanderen nam in 2021 voor de derde keer deel aan PIRLS. PIRLS 2021 werd uitgevoerd door de KU Leuven en de UGent.

PIRLS heeft als indirect doel om de kwaliteit van het onderwijs wereldwijd te verbeteren. Het onderzoek bezorgt beleidsmakers referentiecriteriën en feedback over hoe de leerlingen in hun onderwijssysteem presteren. Zo kan het onderzoek inspiratie bieden voor curricula en beleidskeuzes.

Ook voor Vlaanderen is de deelname aan PIRLS dus een belangrijke vorm van kwaliteitsmonitoring. De combinatie van verschillende metingen die toelaten om longitudinaal of internationaal te vergelijken, geeft ons een genuanceerd beeld van de aanpak, context en opbrengsten van het Vlaamse leerplichtonderwijs doorheen de tijd. PIRLS is één van de initiatieven die dit doel helpen te bereiken voor begrip- en leesniveau in het basisonderwijs. Mede op basis van deze informatie kan de Vlaamse Regering gefundeerde beleidsbeslissingen nemen.

1.2. Het doel van dit rapport

Op 16 mei 2023 werden de eerste resultaten van PIRLS 2021 in Vlaanderen bekend gemaakt. Aanvullend op het internationale rapport (Mullis et al., 2023) biedt het rapport 'Leesvaardigheid in het vierde leerjaar in Vlaanderen: resultaten van PIRLS 2021 in internationaal vergelijkend perspectief' (Denies et al., 2023) een gedetailleerd overzicht van het onderzoeksdesign en van de prestaties van de Vlaamse leerlingen in verhouding tot die in de andere landen. Het rapport gaat ook in op relevante kenmerken van leerlingen en hun gezin, leerkrachten en scholen en het schetst verbanden tussen deze kenmerken en de prestaties.

Eén schoolkenmerk dat tot dusver niet belicht werd, is het net. Uit het net waartoe de school behoort kunnen we afleiden of ze georganiseerd is door of in opdracht van een overheid (de Vlaamse overheid, een stad, gemeente of provincie), of door een privé-organisatie of -persoon. Binnen elk net bieden één of meerdere instellingen, koepels of organisaties pedagogische ondersteuning. Het is dus niet onrealistisch dat het net waartoe een school behoort samenhangt met de prestaties van haar leerlingen. In dit verdiepend rapport lichten we daarom enkele analyses toe die de gemiddelde prestatieverschillen tussen de netten in kaart brengen en trachten te verklaren.

In Hoofdstuk 2 lichten we kort enkele basiskenmerken van de steekproef toe, samen met de werkwijze om een vaardigheidsscore te bepalen en om achtergrondkenmerken in kaart te brengen. Deze informatie is nodig om de analyses beter te kunnen begrijpen. Hoofdstuk 3 toont de ruwe verschillen tussen de netten en in Hoofdstuk 4 gaan we dieper in op de rol van de achtergrondkenmerken van de leerlingen.

2. Methode

2.1. Steekproef en representativiteit

Figuur 1 beschrijft de manier waarop de steekproef voor PIRLS 2021 in Vlaanderen werd getrokken door de IEA. Stap 4 uit deze figuur is in deze context extra belangrijk omdat deze zorgde voor een goede vertegenwoordiging van verschillende types van scholen wat betreft twee kenmerken:

- > We maakten een onderscheid tussen **vrije scholen en scholen die deel uitmaken van het officieel onderwijs** dat georganiseerd wordt door of in opdracht van de overheid (gemeenschapsonderwijs of gesubsidieerd officieel onderwijs). Er was dus geen apart stratum voor elk individueel net, maar de steekproef beoogde wel representativiteit wat betreft de verdeling tussen 'vrij' en 'officieel' onderwijs.
- > Om de samenstelling van het leerlingenpubliek in kaart te brengen, keken we naar de **gemiddelde score voor onderwijskansarmoede (OKI)** van alle leerlingen in de school. Leerlingen kregen één punt per indicator die op hen van toepassing is: (1) ze verkrijgen een schooltoelage, (2) hun moeder heeft een laag opleidingsniveau (maximaal lager secundair onderwijs), (3) hun gezinstaal is niet Nederlands en (4) ze wonen in een buurt met een hoge mate van schoolse vertraging. Op basis van de gemiddelde OKI-score deelden we de scholen in in drie groepen: hoge OKI-compositie, medium OKI-compositie of lage OKI-compositie.

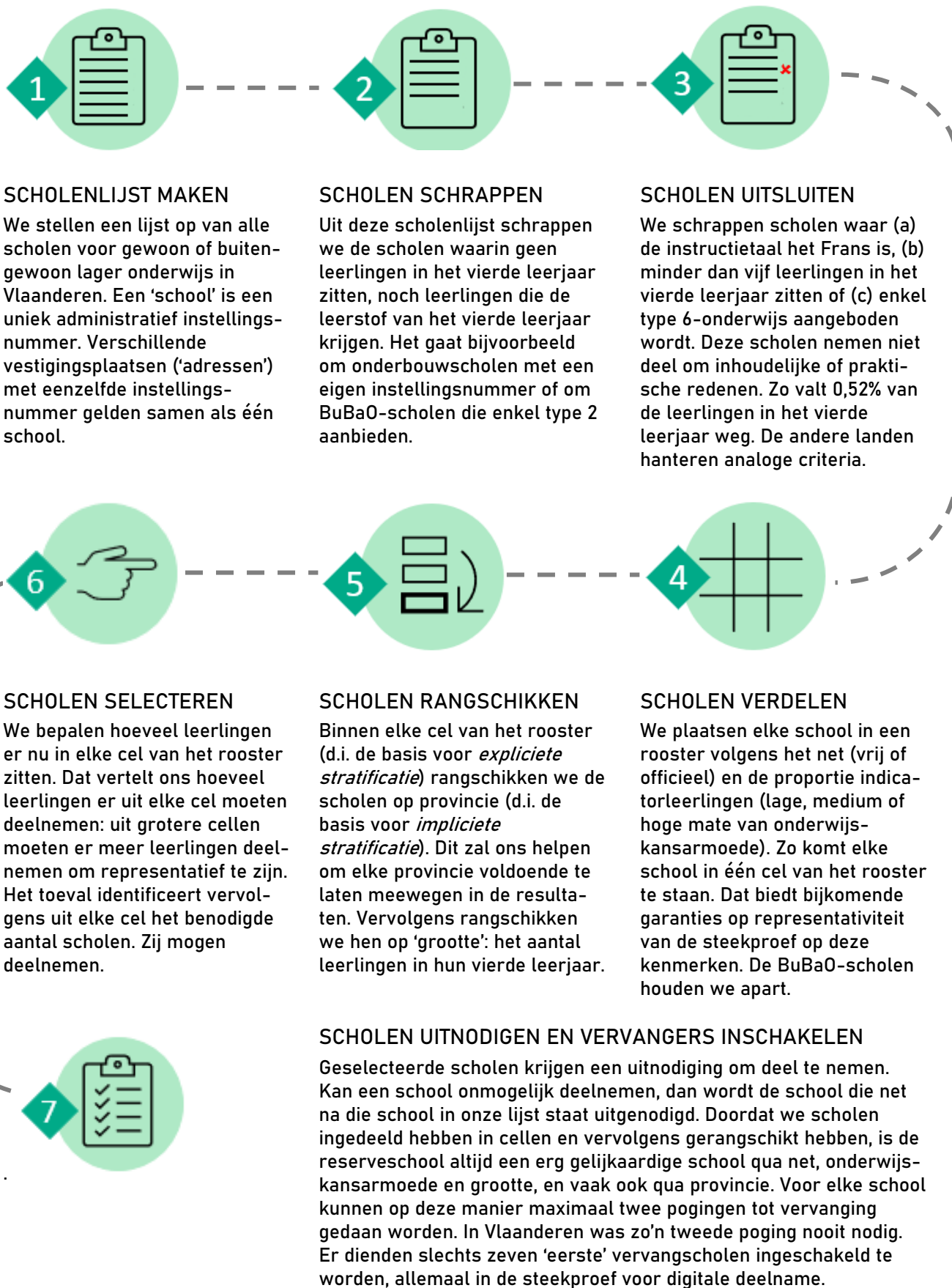
In de analyses maken we gebruik van gewichten ter correctie van een eventuele ondervertegenwoordiging door non-respons. Daardoor zijn de resultaten **representatief** voor alle scholen voor gewoon lager onderwijs in Vlaanderen wat betreft de mate van onderwijskansarmoede van hun leerlingen en de nettengroep waartoe ze behoren.

Ook scholen voor **buitengewoon basis-onderwijs (BuBaO)** zaten in de steekproef voor PIRLS 2021. Ze namen deel met alle leerlingen die leesonderwijs krijgen op een niveau dat ongeveer overeenstemt met dat van een vierde leerjaar in het gewoon onderwijs. We hebben deze scholen evenwel **verwijderd uit de analyses** die in dit rapport toegelicht worden omdat ze de resultaten kunnen vertekenen.

Verder is het goed om te weten dat de voorliggende **analyses enkel gebruik maken van de data van leerlingen die digitaal deelnamen**. 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 worden echter tot dusver niet gebruikt voor inhoudelijke analyses. Vandaar dat ze ook in dit rapport niet opgenomen zijn.

Als gevolg van de schrapping van de BuBaO-scholen, zijn de analyses in dit rapport gebaseerd op de gegevens van **136 scholen en 5 047 leerlingen**. Het gaat om 81 scholen in het vrij gesubsidieerd onderwijs (VGO), 33 scholen in het officieel gesubsidieerd onderwijs (OGO) en 22 scholen in het gemeenschapsonderwijs (GO). Ter vergelijking: voor PIRLS 2016 nemen we 88 scholen in het VGO op, 35 scholen in het OGO en 18 scholen in het GO.

Figuur 1: De verschillende stappen in beide steekproeftrekkingen (digitaal en pen-en-papier)



2.2. Bepalen van een vaardigheidsscore

De gestandaardiseerde PIRLS-leestoetsen werden **digitaal** afgenomen van voltallige klassen. De leerlingen werden ondersteund door een externe, centraal opgeleide begeleider. Elke toets bestond uit twee toetsblokken. Een toetsblok is een geheel van één verhalende of informatieve tekst met daarbij een reeks **open vragen en meerkeuzevragen**.

In een eerste stap ter verwerking van de toetsen werden de antwoorden van de leerlingen beoordeeld aan de hand van een sleutel (voor de meerkeuzevragen) of een gestandaardiseerde scoringsgids (voor de open vragen). Zo beschikten we voor elke leerling over een reeks **scores**: 0 voor een fout antwoord en veelal 1 voor een juist antwoord. Voor sommige vragen werd verder gedifferentieerd tussen volledig juiste antwoorden en gedeeltelijk juiste antwoorden door ook een score 2 of zelfs 3 toe te kennen.

Op de resulterende databestanden werden analyses toegepast volgens de principes van de **item-responstheorie** (IRT). Dat proces leidde tot een set van vijf 'plausibele' leesvaardigheidsscores per leerling. Hoe groter de leesvaardigheidsscore van de leerling, hoe sterker hij of zij is in begrijpend lezen. In onze analyses gebruiken we steeds elk van deze **vijf leesvaardigheidsscores** apart, waarna we de resultaten samenvoegen. Hiermee volgen we de werkwijze die aangeraden wordt door de IEA.

2.3. Bepalen van achtergrondkenmerken

Naast toetsen zet PIRLS ook **vragenlijsten** in. Die verzamelen gegevens over de leerlingen en hun gezin, over de leerkrachten en de scholen. In het kader van dit onderzoek maken we enkel gebruik van elementen uit de oudervragenlijst en de leerlingvragenlijst, zoals de geboortedatum en het geslacht van

de leerling en de mate waarin het Nederlands de thuistaal is.

Aanvullend doen we een beroep op **administratieve gegevens** uit de database van het Departement Onderwijs. Op basis van het instellingsnummer en het stamboeknummer van de leerlingen, verkrijgen we twee indicatoren voor onderwijskansarmoede (OKI): of de moeder al dan niet beschikt over een diploma van het secundair onderwijs, en of het gezin al dan niet begunstigde is van een schooltoelage.

2.4. Analyses

De beschikbaarheid van vaardigheidsscores voor elke leerling laat ons toe om **gemiddelde scores** te berekenen voor bijvoorbeeld alle jongens, voor leerlingen die graag lezen, of voor eender welke andere groep.

In dit rapport bespreken we concreet de gemiddelde score van de leerlingen in (a) het gemeenschapsonderwijs (GO), (b) het officieel gesubsidieerd onderwijs (OGO), en (c) het vrij gesubsidieerd onderwijs (VGO).

Op basis van statistische analyses gaan we na of er verschillen zijn in de gemiddelde leesvaardigheid van de leerlingen tussen de verschillende netten. Dat doen we op twee manieren:

- ① We schatten via een leeg model de **ruwe verschillen** zonder rekening te houden met achtergrondkenmerken van de leerlingen of scholen.
- ② We schatten via een 'nettomodel' de **verschillen na controle voor achtergrondkenmerken** van de leerlingen en scholen.

We zetten deze tweede stap omdat de werkwijze uit stap één geen bijzonder faire vergelijking maakt. De leesvaardigheid van leerlingen staat immers in verband met heel wat verschillende eigenschappen van de leerling, de thuissituatie, de klas, de school en zo verder. Wanneer het ene net meer leerlingen telt die eigenschappen hebben die samenhangen met een lagere score dan het andere, is het aangewezen om dit in rekening te brengen. Dat doen we aan de hand van een statistisch '**nettomodel**'. Het laat toe om te onderzoeken wat de rol van het net is wanneer de leerlingen in andere opzichten gelijk zijn aan elkaar. In het nettomodel nemen we concreet verschillende achtergrondkenmerken op waarvoor we willen controleren omdat we onder meer dankzij het eerste PIRLS-rapport (Denies et al., 2023) weten dat ze samenhangen met de resultaten.

2.5. Netten vergelijken

Het is belangrijk om vergelijkingen van gemiddelde scores correct te interpreteren. Daarvoor is kennis nodig van enkele belangrijke statistische concepten.

PIRLS is gebaseerd op een steekproef.

Het is niet mogelijk om alle leerlingen uit de onderzoekspopulatie te laten deelnemen aan een onderzoek zoals PIRLS. Daarom werden de PIRLS-toetsen afgenomen bij een beperkte deelgroep van de onderzoekspopulatie (zie 2.1). Op basis van de resultaten van deze representatieve steekproef doen we uitspraken over het leesbegrip van de leerlingen in heel Vlaanderen.

Betrouwbaarheidsinterval. Hoe kleiner het deel van de onderzoekspopulatie is dat heeft deelgenomen aan een onderzoek, hoe minder zeker we ervan kunnen zijn dat de gemiddelde toetsresultaten precies overeenkomen met de resultaten die de hele

onderzoekspopulatie zou behalen. Er is met andere woorden sprake van statistische onzekerheid. Om deze onzekerheid concreet weer te geven berekenen we rond elk gemiddelde een betrouwbaarheidsinterval. Deze intervallen geven het gebied aan waarbinnen het werkelijke gemiddelde hoogstwaarschijnlijk (met name: met 95% zekerheid) valt. Een betrouwbaarheidsinterval is vaak groter wanneer de steekproef kleiner is en/of wanneer de scores binnen de betrokken deelgroep verder uit elkaar liggen.

Significantie. Wanneer we willen bepalen of de resultaten van de leerlingen in één net verschillen van de resultaten van de leerlingen in een ander net, is het belangrijk om na te gaan of we met voldoende zekerheid kunnen besluiten dat de resultaten van beide volledige onderzoekspopulaties daadwerkelijk van elkaar verschillen. Een verschil tussen de gemiddelde PIRLS-resultaten van beide netten is statistisch significant als de betrouwbaarheidsintervallen niet overlappen. Wanneer de betrouwbaarheidsintervallen wel overlappen, kunnen we een verschil doorgaans niet (statistisch) significant noemen en mogen we dus niet aannemen dat er in het Vlaamse vierde leerjaar als geheel een betekenisvol verschil bestaat.

3. Ruwe verschillen tussen netten

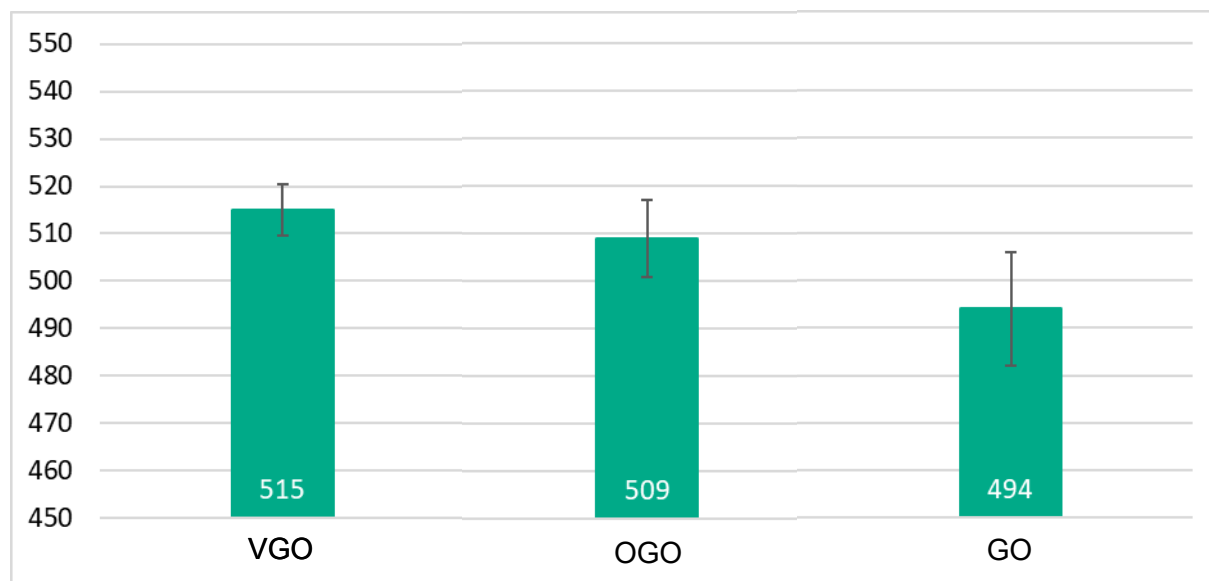
3.1. Situatie in 2021

Uit Figuur 2 blijkt dat de gemiddelde score van leerlingen in het VGO op PIRLS 2021 515 bedroeg. Daarmee was het VGO het hoogst scorende net in PIRLS 2021. Leerlingen uit het OGO scoorden gemiddeld 6 punten lager. De leerlingen uit het GO behaalden met 494 punten de laagste gemiddelde score. Als we de richtlijnen van Rindermann (2007) volgen, kunnen we het verschil tussen het hoogst en het laagst scorende net zeer voorzichtig schetsen als een gemiddelde achterstand van bijna een half jaar begrijpend-leesonderwijs.

Merk echter op dat de betrouwbaarheidsintervallen van het VGO en het OGO overlappen met elkaar, net zoals de betrouwbaarheidsintervallen van het GO en het OGO. Vooral het betrouwbaarheidsinterval voor het GO, dat vertegenwoordigd werd door de kleinste groep leerlingen en scholen (zie 2.1), is erg groot.

We kunnen op basis van deze resultaten niet besluiten dat er statistisch significante scoreverschillen zijn tussen de netten, behalve tussen het VGO en het GO. De leerlingen in het VGO behaalden in 2021 gemiddeld een significant hogere score dan de leerlingen in het GO.

Figuur 2: Prestaties begrijpend lezen opgedeeld naar net voor 2021, zonder statistische controle voor andere kenmerken



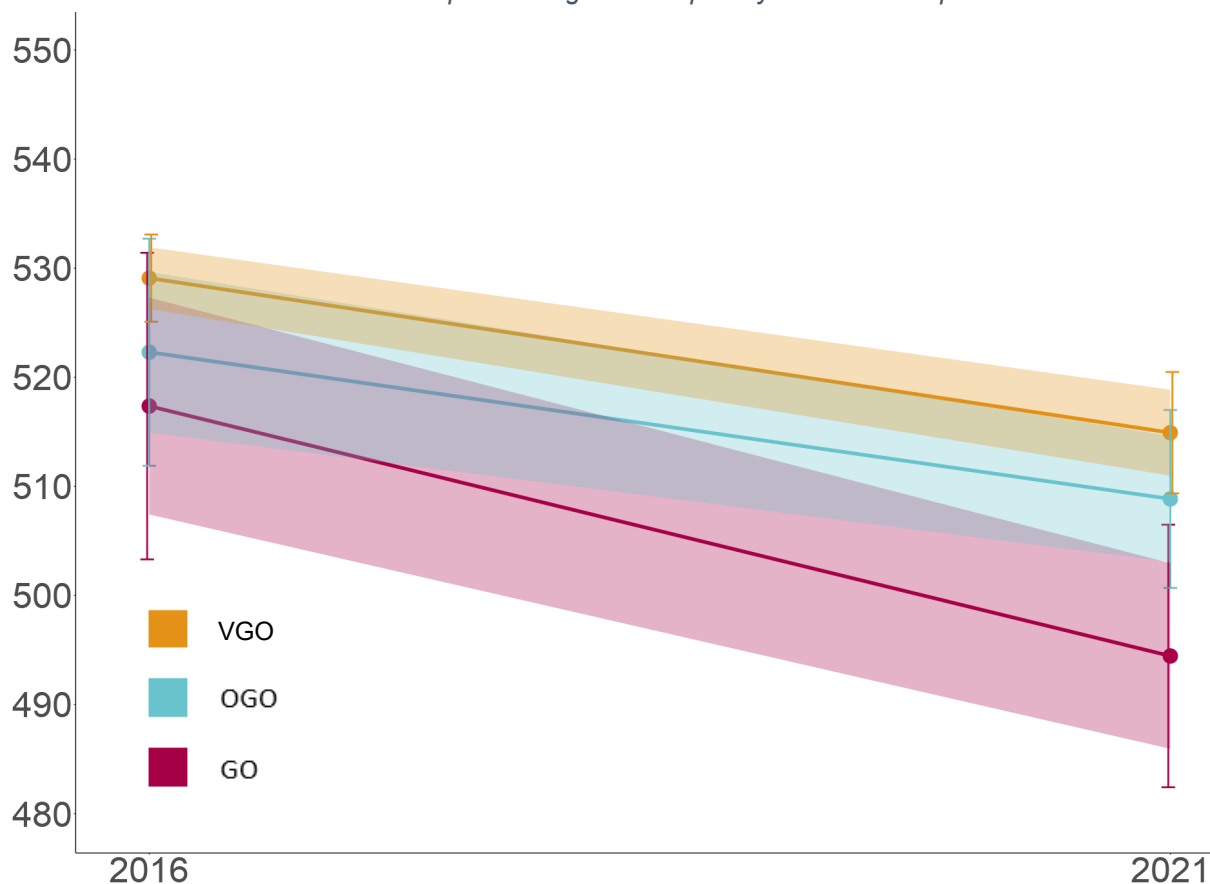
3.2. Evolutie ten opzichte van 2016

De situatie in 2021 blijkt gelijkaardig te zijn aan die van 2016. De netten waren toen op dezelfde manier gerangschikt. In 2016 waren er evenwel geen significante onderlinge verschillen.

Figuur 3 toont links de gemiddelde score per net in 2016, en rechts de gemiddelde score per net in 2021. De verticale strepen rond de puntscores geven de betrouwbaarheidsintervallen aan. Ze tonen dus de waarschijnlijke limieten van de puntmeting in 2016 en 2021. De gemiddelde score van de leerlingen ligt met 95% waarschijnlijkheid binnen beide uitersten (dus: ergens op deze verticale lijn). De gekleurde linten geven het 95% betrouwbaarheidsinterval voor de trend weer. De trendlijn loopt met andere woorden met 95% zekerheid binnen dit lint.

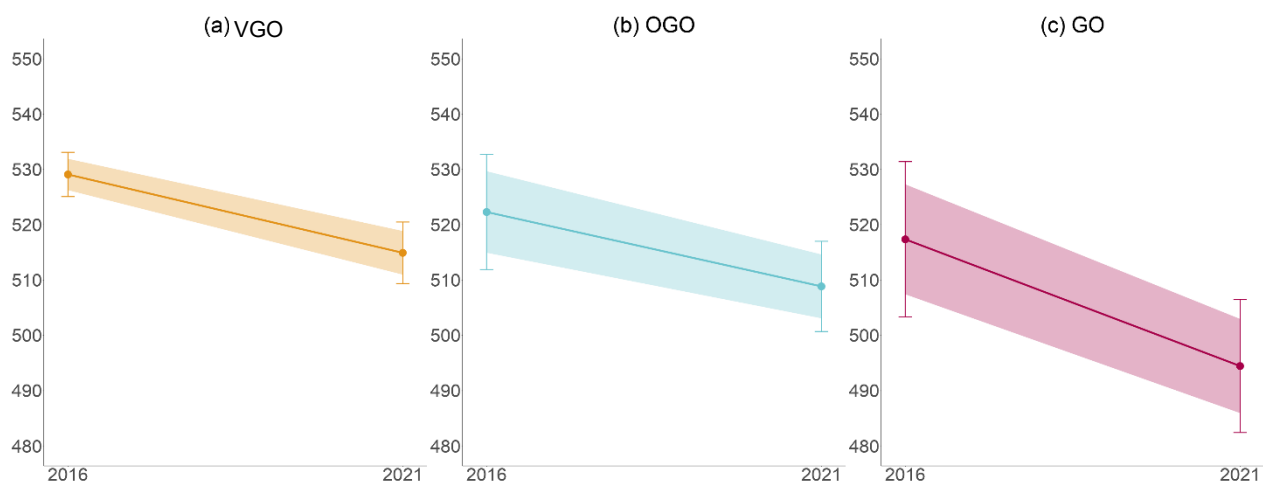
De lijnen in deze figuur doen uitschijnen dat de achteruitgang bij het GO nog groter was dan in de andere netten. De donkerroze lijn loopt namelijk steiler af dan de blauwe en de oranje (of: ze heeft een grotere hellingsgraad). Ook bij dit gegeven wijzen we echter op de grote betrouwbaarheidsintervallen als belangrijke kanttekening. We kunnen uit deze gegevens niet besluiten dat de achteruitgang in het volledige GO daadwerkelijk groter is dan in de andere netten. De donkerroze lijn kan immers in realiteit ook een stuk horizontaler lopen. Bovendien wijzen we er alvast op dat de steekproef van scholen die deelnam in 2021 toevalligerwijs een hogere score voor onderwijskansarmoede had dan in 2016, wat zeker een verklarende rol kan spelen in deze resultaten.

Figuur 3: Prestaties voor begrijpend lezen opgedeeld per net voor 2016 en 2021, zonder statistische controle voor andere kenmerken, met indicatie van het betrouwbaarheidsinterval voor de trend en de betrouwbaarheidsintervallen voor de puntmetingen – let op: de y-as start niet op 0.



Er is dermate veel overlap tussen de gekleurde linten dat we Figuur 4 toevoegen om de individuele betrouwbaarheidsintervallen beter zichtbaar te maken. Op basis hiervan kunnen we vaststellen dat er met 95% zekerheid sprake is van een daadwerkelijke daling in de gemiddelde scores van het GO, het VGO en het OGO. Ook voor het OGO kunnen we immers – op een haartje na – net géén perfect horizontale lijn trekken binnen de grenzen van het blauwe lint. Met andere woorden: elk net daalt significant.

Figuur 4: Prestaties voor begrijpend lezen opgedeeld per net voor 2016 en 2021, zonder statistische controle voor andere kenmerken, met indicatie van het betrouwbaarheidsinterval voor de trend – let op: de y-as start niet op 0.

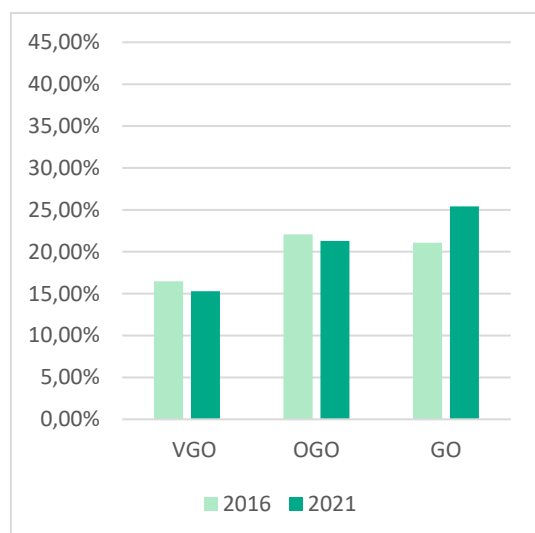


4. Verschillen tussen netten na correctie voor achtergrondkenmerken

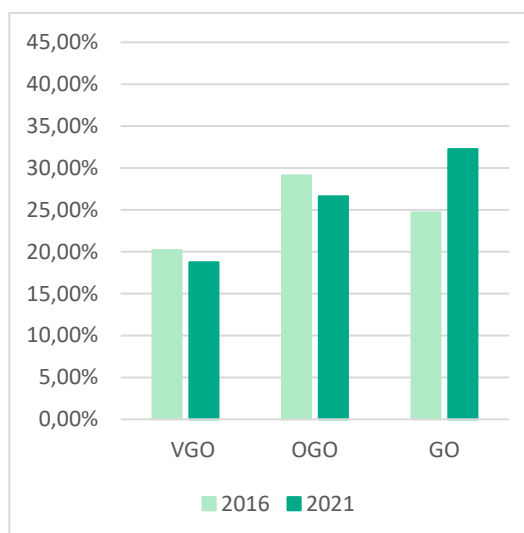
4.1. Verantwoording

Het is aan te bevelen om de analyses uit Hoofdstuk 3 te repliceren met controle voor relevante achtergrondkenmerken van de leerlingen. Op basis van de publiek beschikbare gegevens (Departement Onderwijs, 2016 en 2021) stellen we namelijk vast dat er zowel in PIRLS 2016 als in PIRLS 2021 verschillen tussen de netten waren wat betreft de gemiddelde score voor onderwijskansarmoede (OKI) van de deelnemende scholen. Figuur 5 tot Figuur 8 geven de verschillen weer. Let op: het betreft de gegevens specifiek voor de deelnemende scholen, niet die voor Vlaanderen als geheel.

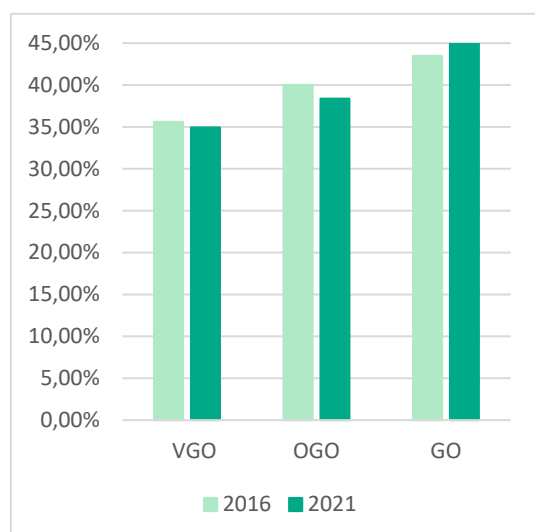
Figuur 5: Gemiddelde OKI-score voor de indicator 'moeder heeft geen diploma van het secundair onderwijs' per net in PIRLS 2016 en PIRLS 2021



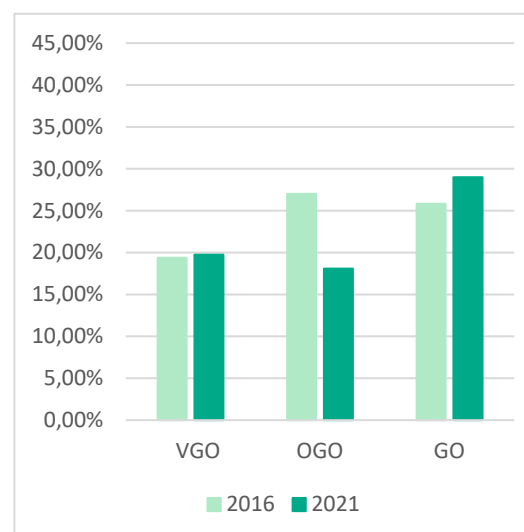
Figuur 6: Gemiddelde OKI-score voor de indicator 'de thuistaal is niet Nederlands' per net in PIRLS 2016 en PIRLS 2021



Figuur 7: Gemiddelde OKI-score voor de indicator 'gezin van de leerling is begunstigde van een schooltoelage' per net in PIRLS 2016 en PIRLS 2021



Figuur 8: Gemiddelde OKI-score voor de indicator 'leerling woont in een buurt met hoge mate van schoolse vertraging' per net in PIRLS 2016 en PIRLS 2021



De scholen die deelnamen aan PIRLS 2021 scoren met 0,89 gemiddeld het laagst op onderwijskansarmoede in het VGO. Zij hebben dus de meest gunstige gemiddelde uitgangspositie. De hoogste gemiddelde score van 1,33 is voor de deelnemende scholen uit het GO, met het OGO (1,04) er tussenin. Deze rangschikking stemt overeen met wat we zien in het Vlaams lager onderwijs als geheel.

De evolutie tussen PIRLS 2016 en PIRLS 2021 in de OKI-scores van de deelnemende scholen is bovendien anders naargelang het net:

- > In de groep deelnemende scholen uit het VGO blijven de OKI-scores vrijwel gelijk: 0,92 in 2016 en 0,89 in 2021.
- > Bij de deelnemers uit het OGO daalt de gemiddelde OKI-score van 1,18 in PIRLS 2016 naar 1,04 in PIRLS 2021.
- > In het GO stijgen de gemiddelde OKI-scores van PIRLS 2016 naar PIRLS 2021, respectievelijk van 1.15 naar 1.33. De GO-scholen die deelnamen in 2021 hadden dus gemiddeld een leerlingenpubliek dat vaker aantikte op de OKI dan de GO-scholen die deelnamen in 2016.

Om een eerlijke vergelijking mogelijk te maken, is het dus belangrijk om statistisch te controleren voor enkele achtergrondkenmerken van de leerlingen en voor enkele indicatoren van de kenmerken van het leerlingenpubliek in de scholen als geheel.

4.2. Variabelen in het nettomodel

Tabel 1 toont de volledige lijst van kenmerken die we opnemen in het nettomodel als controlevariabelen.

Wat de leerlingkenmerken betreft, zijn we beperkt tot de variabelen die ook in 2016 beschikbaar waren. Voor PIRLS 2016 beschikken we niet over de OKI-indicatoren voor individuele leerlingen. Daarom moeten we werken met vergelijkbare indicatoren uit de achtergrondvragenlijsten voor ouders.

Als indicator van de schoolcompositie nemen we de gemiddelde scores voor de indicatoren 'moeder heeft geen diploma van het secundair onderwijs' en 'schooltoelage' op.

Tabel 1: Overzicht van opgenomen controlevariabelen

Leerlingkenmerken	Schoolkenmerken
> Geslacht	> Compositie naar opleidingsniveau moeder (OKI)
> Geboortemaand	> Compositie naar schooltoelage (OKI)
> Geboortjaar (voorsprong of achterstand)	
> Aantal boeken thuis	
> Thuis taal (aandeel van Nederlands thuis)	
> Hoogste opleidingsniveau ouders	

We repliceerden deze analyses voor 2021 ook op een model waarin we wél gebruik maken van de individuele OKI-scores van de leerlingen, in de plaats van de (zo veel mogelijk) equivalente informatie uit de achtergrondvragenlijsten. De resultaten zijn zeer gelijkaardig.

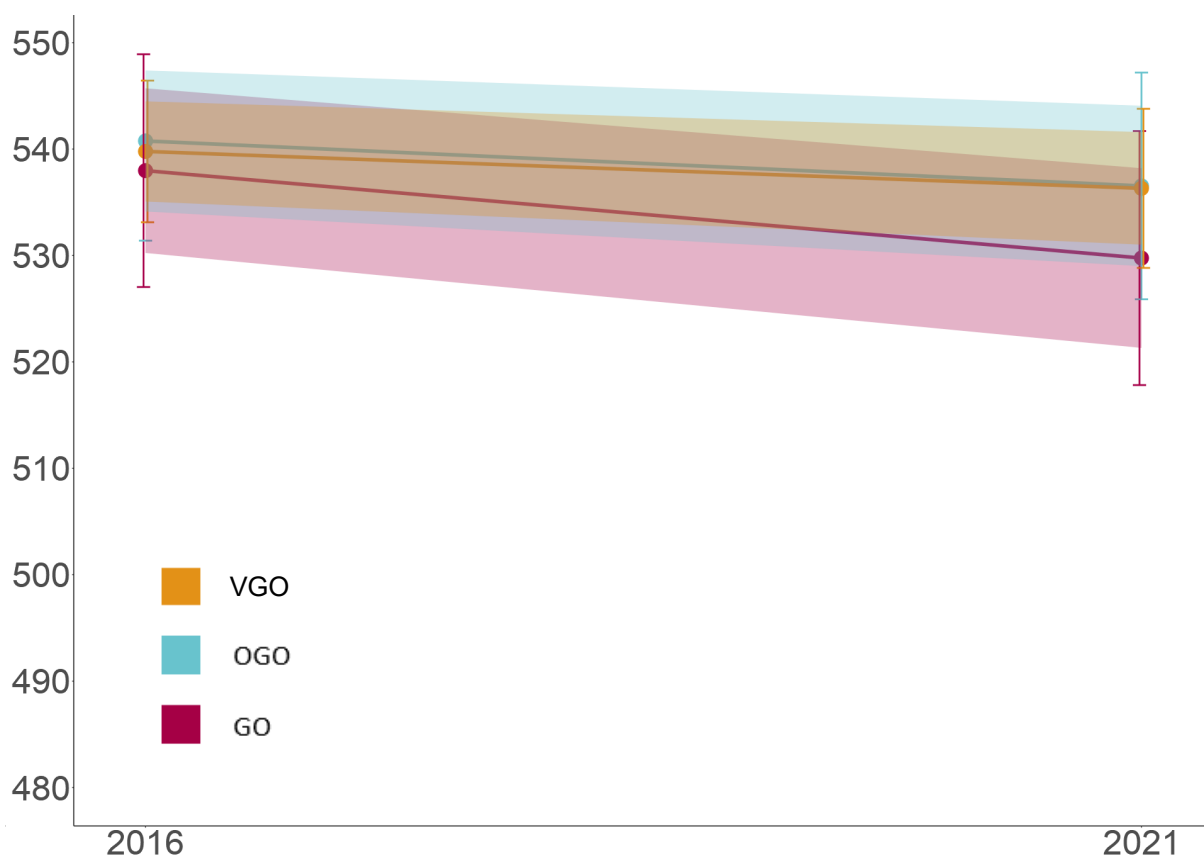
4.3. Verschillen tussen de netten in 2016 en 2021 in het 'nettomodel'

Figuur 9 toont de verschillen tussen de netten na controle voor de variabelen uit Tabel 1. Concreet worden de intercepten uit het nettomodel voor 2016 en het nettomodel 2021 weergegeven, aangepast per net. Deze scores liggen veel hoger dan in het lege model omdat het over de scores voor de 'referentieleerling' gaat. Dat is een jongen die geboren is in juli, geen schoolse vertraging of voorsprong heeft, thuis altijd Nederlands spreekt en er een kast vol boeken heeft, en ouders met een universitair diploma heeft. In zijn school zijn er geen leerlingen die een schooltoelage krijgen of een laagopgeleide moeder hebben.

Wat meteen opvalt, is dat de verschillen tussen de netten nu veel minder uitgesproken zijn. De betrouwbaarheidsintervallen overlappen ruim meer dan in het lege model. Een vrij groot deel van de verschillen tussen de netten is immers te wijten aan de verschillen in schoolcompositie en in de achtergrondkenmerken van de leerlingen die deelnamen.

Zo berekenen we dat in 2021 42.6% van de verschillen tussen de scholen verklaard worden door de schoolkenmerken 'compositie naar opleidingsniveau moeder' en 'compositie naar schooltoelage'. Wanneer we daarbij de netten toevoegen als voorspellers, stijgt dit percentage slechts naar 44.7%. De netten verklaren bovenop de schoolcompositiekenmerken dus slechts weinig van de verschillen tussen de scholen in 2021.

Figuur 9: Prestaties voor begrijpend lezen opgedeeld per net voor 2016 en 2021 in het 'nettomodel', met indicatie van het betrouwbaarheidsinterval voor de trend – let op: de y-as start niet op 0.



5. Conclusie

Dit verdiepende rapport bij PIRLS 2021 in Vlaanderen wierp meer licht op de eventuele verschillen tussen de drie netten.

We stelden vast dat de gemiddelde PIRLS-score van elk net significant daalde tussen 2016 en 2021. In 2016 vonden we geen significante verschillen tussen de netten, en in 2021 scoorde enkel het GO significant lager dan het VGO.

De significant hogere gemiddelde score voor het VGO dan voor het GO in 2021 wordt echter wegverklaard wanneer we de achtergrondkenmerken van de deelnemende leerlingen en de schoolcompositie op het vlak van onderwijskansarmoede in rekening brengen. Bovenop de opgenomen schoolcompositiekenmerken, verklaart het net waar de school deel van uitmaakt slechts 2,1% van de verschillen tussen scholen.

We kunnen dus niet stellen dat er, na controle voor achtergrondkenmerken van de leerlingen en compositiekenmerken van de scholen, een net is dat significant beter scoorde op PIRLS 2021 dan de andere netten.¹ We kunnen evenmin stellen dat er een net is dat significant sterker dan de andere netten achteruit ging ten opzichte van PIRLS 2016.

¹: Merk op dat de voorliggende analyses enigszins gehinderd worden door de vrij kleine steekproef in het GO. Dit is een logisch gevolg van de steekproefaanpak waarbij er representativiteit nagestreefd werd naargelang de onderwijsnetten. In de basisonderwijs loopt namelijk een relatief klein aandeel van de leerlingen school in het GO. De betrouwbaarheidsintervallen voor het GO zijn hierdoor vrij groot.

Referenties

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: Centrum voor Onderwijseffectiviteit en -evaluatie. https://onderwijs.vlaanderen.be/sites/default/files/2023-05/PIRLS2021_Brochure_KULeuven-UGent.pdf

Mullis, I. V. S., von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021 International Results in Reading*. Boston College, TIMSS & PIRLS International Study Center. <https://pirls2021.org/results>

Onderwijs Vlaanderen. (2016). *Leerlingenkenmerken Overzicht 2015-2016 BaO*. Geraadpleegd op 24 mei 2023, van <https://www.onderwijs.vlaanderen.be/nl/onderwijsstatistieken/themas-onderwijsstatistieken/leerlingenkenmerken>

Onderwijs Vlaanderen. (2021). *Leerlingenkenmerken Overzicht 2020-2021 BaO*. Geraadpleegd op 24 mei 2023, van <https://www.onderwijs.vlaanderen.be/nl/onderwijsstatistieken/themas-onderwijsstatistieken/leerlingenkenmerken>

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, 21(5), 667-706. <https://doi.org/10.1002/per.634>

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

