

# **Van Groei tot Bloei**

Een onderzoek naar de relatie tussen groei in rekenvaardigheid en groei in taalvaardigheid  
tussen groep 3 en groep 5

Student: Marit Bulthuis (S3704041)

Begeleider: dr. N. Frans

Tweede beoordelaar: dr. L. Baams

Rijksuniversiteit Groningen, faculteit Gedrags- en Maatschappijwetenschappen

Afdeling Pedagogische en Onderwijswetenschappen

Masterthesis

Datum: 31-05-2024

Aantal woorden: 7995 woorden

### Samenvatting

Nationaal en internationaal onderzoek wijzen uit dat de basisvaardigheden van Nederlandse leerlingen, ondanks enorme investeringen, sterk onder druk blijven staan. Aangezien de basisvaardigheden taal en rekenen op verschillende manieren met elkaar samenhangen, moeten problemen in de ontwikkeling van zowel taal- als rekenvaardigheden vroeg worden opgespoord om ervoor te zorgen dat de rekenbasis goed is. De periode tussen groep 3 en 5 kan hierbij als kritische periode worden gezien. Daarnaast kunnen bepaalde kindkenmerken, net als leesprestaties, een voorspellende waarde hebben. Het lijkt aannemelijk dat een dergelijke voorspellende waarde ook is weggelegd voor groei in leesvaardigheid. Groei blijkt namelijk uit het verschil in prestaties tussen twee toetsmomenten. Dit is onderzocht middels een multilevel regressieanalyse. Hiervoor werd gebruik gemaakt van in 2014 op meerdere momenten verzamelde data van 1402 leerlingen op 59 scholen door Nederland. Tussen groei in begrijpend lezen en groei in rekenen komt een zwak positief verband naar voren. Mogelijk kan een deel van de groei in rekenvaardigheid dus verklaard worden door de groei in begrijpend leesvaardigheid. Door middel van een logistische regressieanalyse is bovendien geprobeerd een dergelijke relatie aan te tonen voor stagnatie, oftewel uitblijvende groei. Hieruit kwamen geen significante resultaten naar voren. Een opvallende nevenbevinding van dit onderzoek is dat er op leerlingniveau weinig systematische verschillen in rekengroei blijken te zijn. Dat wijst erop dat elke leerling min of meer volgens dezelfde tendens groeit. Vervolgonderzoek zou moeten uitwijzen hoe onderscheid gemaakt kan worden tussen leerlingen die systematisch achterlopen en leerlingen die een sprongsgewijze ontwikkeling doormaken.

### **Abstract**

National and international research show that the basic skills of Dutch students are still under pressure, despite enormous investments. Since the basic skills of language and arithmetic are related in different ways, problems in the development of both language and arithmetic skills must be detected early to ensure a good arithmetic foundation. The period between group 3 and 5 can be regarded as critical. In addition, certain child characteristics, as well as reading performance, may have a predictive value. It seems plausible that a similar predictive value is reserved for growth in reading skills, because growth is reflected in the difference in performance between two test moments. This was investigated using a multilevel regression analysis. Data from 1402 students at 59 schools throughout the Netherlands which was collected at several times in 2014 was used. A weak positive relationship is found between growth in reading comprehension and growth in arithmetic skills. Possibly, part of the growth in arithmetic skills therefore can be explained by the growth in reading comprehension. A logistic regression analysis was used to demonstrate a similar relationship for stagnation or a lack of growth. No significant results were found. A remarkable incidental finding are the few systematic differences at student level in growth in arithmetic skills. This indicates that every student grows more or less according to the same trend. Follow-up research should find out how a distinction can be made between students who are systematically lagging behind and students who are experiencing a step-by-step development.

## Inleiding en theoretisch kader

Zowel nationaal onderzoek (Inspectie van het Onderwijs, 2023) als internationaal onderzoek (Meelissen et al., 2023) laat zien dat de basisvaardigheden van Nederlandse leerlingen sterk onder druk staan. Uit Peil.Taal en rekenen 2021-2022 blijkt bijvoorbeeld dat aan het einde van het basisonderwijs vooral de beheersing van het streefniveau voor rekenen achterblijft bij de gestelde doelen. Aan het einde van het speciaal (basis)onderwijs blijven de rekenprestaties ook achter ten opzichte van de prestaties op andere leergebieden (Inspectie van het Onderwijs, 2023). Uit PISA-2022 blijkt eveneens dat de prestaties van Nederlandse 15-jarigen een achteruitgang laten zien op wiskunde ten opzichte van vier jaar geleden. Echter, vooral voor leesvaardigheid is de afname in prestatieniveau in Nederland sterker dan in andere landen. De Nederlandse daling in leesvaardigheid tussen 2018 en 2022 is ook groter dan de daling tussen 2015 en 2018 en voor de eerste keer scoort Nederland ruim onder het gemiddelde van de landen die behoren tot de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) (Meelissen et al., 2023).

De aandacht voor het taal- en rekenonderwijs is sinds de invoering van de Wet referentieniveaus Nederlandse taal en rekenen (2010) enorm toegenomen. Bovendien maakte het kabinet Rutte IV in het coalitieakkoord structureel één miljard euro beschikbaar om onder andere de basisvaardigheden te verbeteren en werd boven op dat bedrag meermaals tijdelijk extra geld beschikbaar gesteld (NOS Nieuws, 2023). Desalniettemin blijven de basisvaardigheden van Nederlandse leerlingen sterk onder druk staan.

De basisvaardigheden taal en rekenen hangen op verschillende manieren met elkaar samen. Zo zijn veel rekenmethoden die in het Nederlandse basisonderwijs gebruikt worden (deels) gebaseerd op de realistische rekendidactiek. Dat houdt in dat rekenopgaven in rijke, talige contexten worden geplaatst om leerlingen te activeren en om hun betrokkenheid te verhogen (Prenger, 2011). De contexten kunnen er tevens voor zorgen dat leerlingen beter grip krijgen op begrippen en relaties, op nuances en accenten en op situaties en handelingen (Janson, 2014). Een aantal jaren geleden waren veel mensen daarom van mening dat ook tijdens de rekenlessen aandacht besteed zou moeten worden aan taalvaardigheid om aan het einde van de basisschool het gewenste niveau te kunnen bereiken (Prenger, 2011).

Tegenwoordig nemen steeds meer mensen een kritische houding aan ten opzichte van de realistische rekendidactiek, omdat empirische evidentie ontbreekt. Bovendien zou de realistische rekendidactiek, in tegenstelling tot de traditionele rekendidactiek, te weinig aandacht besteden aan het automatiseren van de basiskennis en een te groot beroep doen op vaardigheden die niet direct met rekenen te maken hebben zoals taalvaardigheid (Schmeier,

2017). Daardoor kunnen kinderen problemen ervaren bij het begrijpen van teksten in rekenmethoden, bij het begrijpen van wat er tijdens de rekenles gezegd wordt en bij het zelf gebruiken van taal tijdens de rekenles (Van Eerde, 2009).

Uit eerder correlatieve onderzoek (Krajewski & Schneider, 2009; LeFevre et al., 2010) blijkt daarnaast dat bij kinderen met een normale taalontwikkeling zowel cognitieve factoren (intelligentie, werkgeheugen en benoemsnelheid) als taalfactoren (fonologisch bewustzijn en grammaticale vaardigheid) een rol spelen bij het getalbegrip in de kleuterklas en later ook bij de totstandkoming van het aanvankelijk rekenen. In allerlei dagelijkse situaties komen kinderen in aanraking met klanken, symbolen en betekenissen die naar getallen verwijzen. Denk aan: telversjes, leeftijden, klok en kalender, huisnummers en het tellen en herkennen van kleine hoeveelheden zoals de ogen van een dobbelsteen (Veltman & Van den Heuvel-Panhuizen, 2015). Doordat kinderen tijdens het oplossen van rekenkundige problemen telstrategieën moeten gebruiken en de fonologische vorm van een telwoord moeten ophalen uit hun werkgeheugen (Simmons et al., 2008), wordt tijdens het oplossen van bijvoorbeeld optel- en aftreksommen tot tien een groot beroep gedaan op het fonologisch bewustzijn en het doorzien van structuren en regels (Kleemans et al., 2012). Op basis van diverse neuropsychologische studies naar de ontwikkeling van rekenvaardigheid waarbij gebruik gemaakt werd van functionele MRI-scans die de activiteit van de hersenen zichtbaar maken (e.g. Dehaene et al., 2003), lijkt eveneens dat het inzicht in en het werken met rekenkundige basisprincipes in sterke mate afhankelijk zijn van taalvaardigheid. Veel rekenproblemen bij kinderen op de basisschool blijken dan ook voort te komen uit onderontwikkelde beginnende geletterd- en gecijferdheid (Geary, 1990).

Verder wordt taal historisch gezien door verschillende gedrags- en maatschappijwetenschappers zoals de ontwikkelingspsycholoog Vygotsky (1986) en de pedagoog Langeveld (1934) gezien als bron voor het denken en leren van individuen. Daarmee zou ook rekenen een communicatieve activiteit zijn die erop gericht is betekenis te geven aan de kwantitatieve, relationele en ruimtelijke aspecten van de wereld (Van Eerde, 2009). Hierbij kan onderscheid gemaakt worden tussen spreek- en schrijftaal en reken- en wiskundetaal. De reken- en wiskundetaal komt voort uit de spreek- en schrijftaal, maar bevat ook eigen symbolen, uitdrukkingen en formules en een eigen grammaticale structuur (Ruijsenaars et al., 2004). Getallen hebben afhankelijk van de situatie de betekenis van een aantal, een getal om te tellen, een maat, een naam/nummer of een getal om mee te rekenen. In het begin staan de functies van getallen nog los van elkaar en kan de verscheidenheid aan betekenissen bij jonge kinderen tot verwarring leiden. Als je een kind dat vijf schelpen in een

emmertje heeft gedaan bijvoorbeeld vraagt of je drie schelpen mag, antwoordt het kind mogelijk dat het niet meer weet welke schelp drie was. Het kind denkt dan dat er naar het naamgetal gevraagd wordt, terwijl er in werkelijkheid naar het aantal wordt gevraagd (Veltman & Van den Heuvel-Panhuizen, 2015). Aan de hand van algemene begrippen die verwijzen naar hoeveelheden, relaties tussen hoeveelheden en handelingen met aantallen en aan de hand van specifieke begrippen zoals de telrij leren kinderen al op jonge leeftijd om de reken- en wiskundetaal te gebruiken om hun nabije realiteit te ordenen en te beschrijven (Ruijsenaars et al., 2004).

Daarentegen hebben numerieke concepten volgens Gelman en Butterworth (2005) en Passolunghi en Mammarella (2010) een ontwikkelingsfysiologische en neurale basis die onafhankelijk is van taal. Volgens hen is kennis van telwoorden dan ook niet noodzakelijk voor het begrijpen van numerieke concepten en kunnen mensen met (ernstige) taalproblemen wel over goede numerieke vaardigheden beschikken. Deze opvatting kan echter niet zonder meer vertaald worden naar de context van het schoolse rekenen. Voor schools rekenen is kennisoverdracht immers noodzakelijk. Dat gebeurt bijvoorbeeld door kinderen een rijke leeromgeving te bieden, maar nog veel vaker middels het gebruik van taal (Veltman & Van den Heuvel-Panhuizen, 2015).

Aangezien taalvaardigheid sterk verbonden is aan andere schoolse vaardigheden, moeten problemen in de ontwikkeling van zowel taal- als rekenvaardigheden vroeg worden opgespoord om ervoor te zorgen dat de rekenbasis goed is (Van Luit, 2002). De periode tussen groep 3 en 5 kan hierbij als kritische periode worden gezien, omdat zich in deze periode specifieke leerproblemen voordoen die vragen om snelle signalering (Geelhoed & Vieijra, 2009). Uit bestaand onderzoek kunnen ook diverse kindkenmerken afgeleid worden die voorspellend kunnen zijn ten aanzien van het schools presteren van kinderen in het basisonderwijs (Verachtert et al., 2007).

Allereerst scoren jongens over het algemeen beter op het gebied van rekenen en meisjes op het gebied van lezen en taalverzorging (Vogelzang, 2018). Daarnaast beginnen kinderen uit de hogere sociale klassen meestal met een voorsprong en kinderen uit de lagere sociale klassen met een achterstand (Schmeier, 2017). Uit PISA-2022 blijkt ook dat kinderen hogere scores halen voor wiskunde en leesvaardigheid naarmate hun ouders hoger opgeleid zijn. Verder blijken kinderen die thuis Nederlands spreken hogere scores te halen dan kinderen die thuis een andere Europese of niet-Europese taal spreken (Meelissen et al., 2023). Hickendorff en Janssen (2009) deden eveneens onderzoek naar de kloof in rekenvaardigheid tussen autochtone en allochtone kinderen. Zij keken daarbij naar de invloed van contexten in

rekenopgaven. Leerlingen uit groep 3, 4 en 5 maakten zowel een taak die uit contextopgaven bestond als een taak met alleen kale rekenopgaven. Uit het onderzoek bleek dat leerlingen die thuis geen Nederlands spreken bij de contextopgaven een grotere prestatieachterstand hadden ten opzichte van leerlingen die thuis wel Nederlands spreken dan bij de kale opgaven.

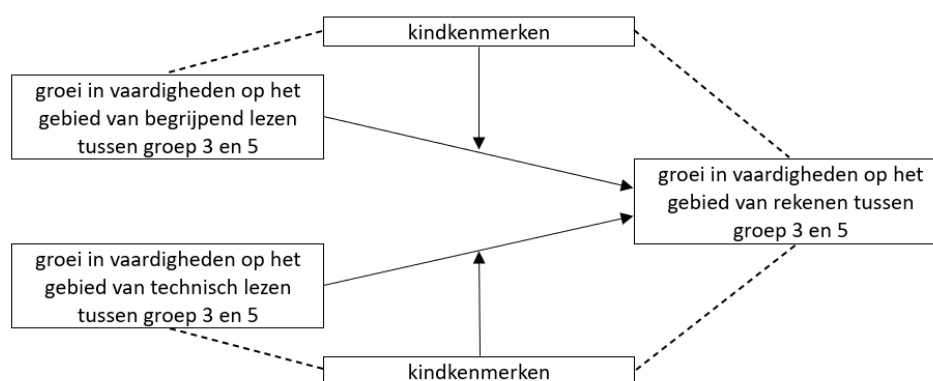
Een meer formele werkwijze van signaleren is het inzetten van gestandaardiseerde toetsen. Op die manier kan een vergelijking gemaakt worden met een norm of met een vooraf opgesteld criterium (Ruijssenaars et al., 2004). Hierbij is het, naast het bepalen van het niveau van een kind, belangrijk om tevens te bepalen of het kind zich voldoende ontwikkelt (Hollenberg & Veerbeek, 2014). Dit kan door te kijken naar het verschil in prestaties tussen twee toetsmomenten, ook wel groei genoemd (Schmeier, 2017). In het nieuwste leerlingvolgsysteem van Cito (2021) is de nadruk dan ook sterk op groei komen te liggen. Voor sommige kinderen geldt dat zij, ondanks het intensiveren van de leertijd en de instructie, achterblijven ten opzichte van hun leeftijdsgenoten en zich de leerstof moeilijk eigen maken. In dergelijke gevallen kan er sprake zijn van ernstige rekenproblemen of mogelijk zelfs van dyscalculie (Schmeier, 2017). Stagnerende vaardigheidsscores kunnen hiervoor een aanwijzing vormen (Naaktgeboren & Prins, 2014). Bovendien komen ernstige rekenproblemen vaak samen voor met ernstige lees- en spellingsproblemen. De vastgestelde comorbiditeit loopt echter uiteen van 11% tot 70% (Ghesquière & Ruijssenaars, 2021). Verder blijkt de samenhang tussen lees- en rekenprestaties het sterkste te zijn (een Pearson-correlatiecoëfficiënt van ,86) wanneer het gaat om automatiseren, oftewel om het vlot kunnen lezen van woorden dan wel het snel kunnen oplossen van simpele sommen (Ruijssenaars et al., 2004). Wanneer de algehele ontwikkeling van een kind dus goed gevolgd wordt, kan het kind gerichter en daarmee ook beter begeleid worden (Nix, 2022).

Aangezien lees- en rekenprestaties met elkaar samenhangen (Ghesquière & Ruijssenaars, 2021; Ruijssenaars et al., 2004) en groei blijkt uit het verschil in prestaties tussen twee toetsmomenten (Schmeier, 2017), lijkt het aannemelijk dat er naast een samenhang in niveau eveneens een verband bestaat tussen groei in leesvaardigheid en groei in rekenvaardigheid. Hiervoor bestaat alleen nog weinig tot geen empirische evidentie. Daarom wordt door middel van dit onderzoek getracht een bijdrage te leveren aan de wetenschappelijke kennis over een dergelijk verband. Op basis van een dergelijk verband zouden problemen in de taal- en rekenontwikkeling namelijk eerder opgespoord en gesignaleerd kunnen worden. De hoofdvraag luidt dan ook als volgt: *In hoeverre is er een relatie tussen groei in vaardigheden op het gebied van rekenen tussen groep 3 en 5 en groei in vaardigheden op het gebied van lezen tussen groep 3 en 5?* Omdat stagnaties in lees- en/of

rekenvaardigheden kunnen wijzen op ernstige leerproblemen (Naaktgeboren & Prins, 2014) en bepaalde kindkenmerken voorspellend kunnen zijn wat betreft schools presteren (Verachtert et al., 2007), zijn de volgende twee deelvragen opgesteld om deze hoofdvraag te kunnen beantwoorden: 1) *In hoeverre is er tussen groep 3 en 5 een relatie tussen stagnatie in rekenvaardigheid en stagnatie in leesvaardigheid?* en 2) *In hoeverre zijn kindkenmerken zoals geslacht, opleiding ouders en migratieachtergrond van invloed op deze relatie?* Het bijbehorende conceptuele model is te zien in Figuur 1.

## Figuur 1

### Conceptuele model



## Methode van onderzoek

### Design

Het doel van dit onderzoek was het in kaart brengen van de relatie tussen groei in rekenvaardigheid en groei in leesvaardigheid tussen groep 3 en 5. Dit gebeurde aan de hand van bestaande data die in 2014 op meerdere momenten werden verzameld voor een eerder onderzoek naar de bruikbaarheid van de kleutertoetsen van Cito voor het signaleren van taal- en/of rekenproblemen (Frans, 2019). Er was dus sprake van een correlatieel, retrospectief en longitudinaal onderzoek.

### Populatie en Steekproef

Dit onderzoek richtte zich op de populatie leerlingen in groep 3 tot en met 5 van het reguliere Nederlandse basisonderwijs. Om iets over deze populatie te kunnen zeggen, werd een geclusterde gelegenheidssteekproef getrokken van 1402 leerlingen op 59 verschillende scholen door heel Nederland, waarbij de volgende inclusiecriteria werden gehanteerd: de leerling zat op een school die gebruik maakte van het leerlingvolgsysteem van Cito (Cito-LOVS) en de leerling doubleerde tussen groep 3 en 5 niet. Het tweede inclusie criterium werd



toegevoegd om het berekenen van groei over de periode tussen groep 3 en 5 mogelijk te maken. Wanneer binnen drie maanden dezelfde toets meerdere keren werd afgenomen, werden alleen de toetsgegevens van de eerste toetsafname bewaard om eventuele test-hertest effecten zoveel mogelijk te beperken. De uiteindelijke steekproefgrootte week dan ook iets af van de hierboven beschreven aantallen.

## **Variabelen en Instrumenten**

### ***(On)afhankelijke variabelen***

In dit onderzoek vormde groei in leesvaardigheid de onafhankelijke variabele en groei in rekenvaardigheid de afhankelijke variabele. Aangezien lezen een complex samenspel is tussen hogere orde processen waarin het begrijpen van geschreven teksten centraal staat en lagere orde processen waarin het om de techniek van het lezen gaat (Sijstra, 2003), werd leesvaardigheid opgesplitst in begrijpend lezen en technisch lezen. Zowel groei in leesvaardigheid als groei in rekenvaardigheid werd in kaart gebracht aan de hand van toetsgegevens uit het Cito-LOVS. Het Cito-LOVS maakte gebruik van gestandaardiseerde, normgerichte toetsen die in het midden van het schooljaar (M) en aan het einde van het schooljaar (E) door de leerkracht afgenomen werden. Voor dit onderzoek werd uitgegaan van de volgende zes toetsmomenten: M3/E3/M4/E4/M5/E5. De leerlingen maakten de toetsen in groepsverband op papier of in een individuele setting op de computer.

De leesvaardigheid van een leerling werd bepaald op basis van de toets Begrijpend lezen en de Drie-Minuten-Toets (DMT). De toets Begrijpend lezen bestond uit meerkeuzevragen op basis waarvan gemeten kon worden in hoeverre de leerling begrip had van geschreven teksten en werd veelal niet afgenomen op de toetsmomenten midden groep 3 (M3) en eind groep 5 (E5) (Feenstra et al., 2010). De DMT bestond uit verschillende leeskaarten waarop woorden stonden die door de leerling hardop moesten worden gelezen en werd afgenomen om vast te kunnen stellen in welke mate de leerling vorderingen maakte bij het accuraat en vlot decoderen van woorden (Krom et al., 2010).

De rekenvaardigheid van een leerling werd bepaald op basis van de toets Rekenen-Wiskunde. De toets Rekenen-Wiskunde bestond voornamelijk uit open vragen waarbij van de leerling een kort antwoord in de vorm van een getal werd verwacht. Zo werd gemeten in hoeverre de leerling beschikte over fundamentele vaardigheden, in hoeverre deze fundamentele vaardigheden door de leerling gebruikt konden worden in praktische toepassingssituaties, of de leerling in staat was om eenvoudige verbanden, regels, patronen en structuren op te sporen en in hoeverre de leerling redeneer- en onderzoeksstrategieën kon gebruiken (Janssen et al., 2010).

Voor de toetsen Begrijpend lezen en Rekenen-Wiskunde gold dat alle opgaven aan de hand van IRT-modellen op dezelfde vaardigheidsschaal werden gekalibreerd tot een vaardigheidsscore (Feenstra et al., 2010; Janssen et al., 2010). Voor de DMT gebeurde dit aan de hand van een variant op een IRT-model (Krom et al., 2010). Elke toets leverde dan ook een vaardigheidsscore op die vergeleken kon worden met nationale normen in de vorm van vaardigheidsniveaus en met eerdere vaardigheidsscores van de leerling. Door te kijken naar het verschil tussen de vaardigheidsscores van twee toetsmomenten, ook wel vaardigheidsgroei genoemd, kon bepaald worden in hoeverre de leerling zich ontwikkelde (Schmeier, 2017).

Voor de toets Begrijpend lezen was de laagste betrouwbaarheidscoëfficiënt ,84 (Feenstra et al., 2010) en voor de DMT en de toets Rekenen-Wiskunde ,92 (Krom et al., 2010; Janssen et al., 2010). Van alle toetsen was de meetnauwkeurigheid volgens het criterium van de Commissie Testaangelegenheden Nederland (COTAN, 2009) dan ook goed te noemen. Er waren tevens aanwijzingen voor een hoge begripsvaliditeit van de toetsen (Feenstra et al., 2010; Krom et al., 2010; Janssen et al., 2010).

### ***Kindkenmerken***

Variabelen die mogelijk invloed hadden op de relatie tussen groei in rekenvaardigheid en groei in leesvaardigheid waren geslacht (jongen/meisje), opleiding ouders ((zeer) laag/normaal) en migratieachtergrond (ja/nee). Wat betreft de variabele ‘opleiding ouders’ werden de categorieën gebaseerd op de gewichtenregeling van het Centraal Bureau voor de Statistiek (CBS). Voor dit onderzoek werden de gewichten ‘laag’ en ‘zeer laag’ samengenomen vanwege de kleine omvang van beide gewichten. Een kind kreeg het gewicht ‘(zeer) laag’ als beide ouders beschikten over een diploma op het niveau van maximaal vmbo-k. Alle andere kinderen kregen het gewicht ‘normaal’ (CBS, 2017). Voor de categorieën met betrekking tot de variabele ‘migratieachtergrond’ werd uitgegaan van het al dan niet hebben van een aanduiding voor een niet-Nederlandse culturele achtergrond (NNCA). Een leerling met een NNCA was een leerling die niet uit Nederland, Indonesië, een Noord-Europees land of een Engelstalig land kwam (Besluit bekostiging WPO, 2022). Alle informatie met betrekking tot deze kindkenmerken kwam uit het leerlingadministratiesysteem van de school en was opgegeven door de leerkracht.

### **Procedure**

Dit onderzoek maakte gebruik van bestaande data uit het Cito-LOVS van de participerende scholen. Deze data werden in 2014 verzameld door Frans (2019). Uit de contactgegevens die de Dienst Uitvoering Onderwijs (DUO) van basisscholen had, werd een

willekeurige selectie gemaakt van de besturen van 1116 scholen. Deze besturen ontvingen een mail met daarin een uitleg over de aard en de doelstellingen van het onderzoek. Na twee weken ontvingen zij een herinnering en uiteindelijk gaven 84 scholen aan mee te willen werken aan het onderzoek. Een aantal van de scholen die niet mee wilden werken gaf aan hier geen tijd voor te hebben. Dezelfde reden werd opgegeven door 25 scholen die de data niet op tijd aanleverden. De 59 scholen die overbleven werden vooraf schriftelijk op de hoogte gesteld van het doel van het onderzoek en het gebruik van de data. Van de kinderen die in september 2014 waren gestart in groep 6 werden met terugwerkende kracht toetsgegevens vanaf de kleuterklas door de school aangeleverd. Dit gebeurde in samenwerking met de onderzoeker. Om de anonimiteit van de kinderen te kunnen waarborgen werden namen, geboortedagen en andere privacygevoelige informatie niet opgeslagen. Het onderzoek werd goedgekeurd door de Ethische Commissie van de afdeling Pedagogische en Onderwijswetenschappen van de Rijksuniversiteit Groningen.

### **Analyseplan**

De data werden geanalyseerd met behulp van R-versie 4.4.0 (R Core Team, 2023). Voorafgaand aan het uitvoeren van de analyses werden de leerlingen die hebben gedoubleerd verwijderd uit de dataset om groei over de periode tussen groep 3 en 5 te kunnen berekenen. Ook werden alle toetsgegevens gematched op basis van de verschillende toetsmomenten, zodat iedere toetsafname gekoppeld werd aan het bijbehorende afnamemoment. Vervolgens werden een groeivariabele en een stagnatievariabele aangemaakt voor begrijpend lezen, technisch lezen en rekenen. Voor de groeivariabele werd het verschil in vaardigheidsscore tussen twee opeenvolgende toetsmomenten berekend. Wanneer de periode tussen twee toetsmomenten minder dan drie maanden besloeg, werd ervan uitgegaan dat het om groei over dezelfde periode ging en werden alleen de toetsgegevens van de eerste toetsafname bewaard. De stagnatievariabele werd aangemaakt in de vorm van een dummyvariabele. Hierbij werd onderscheid gemaakt tussen kinderen met een positieve groei (0) en kinderen die stagneren (1), oftewel geen groei in vaardigheidsscore vertoonden. Om patronen in missende data te analyseren werden tevens dummyvariabelen aangemaakt voor begrijpend lezen, technisch lezen en rekenen, waarbij 0 stond voor vaardigheidsscore wel aanwezig en 1 voor vaardigheidsscore niet aanwezig. Daarna werd gekeken naar de relaties tussen deze dummyvariabelen en de andere variabelen om eventuele patronen in de missende data te onderzoeken. Hiervoor werd gebruikgemaakt van frequentietabellen en kruistabellen.

Om de kenmerken van de steekproef te omschrijven werden de kindkenmerken beschreven in een frequentietabel, werden de vaardigheidsscores en de groeivariabelen

beschreven aan de hand van centrum- en spreidingsmaten, werd de groei per toetsmoment beschreven aan de hand van boxplots en werd de stagnatie per toetsmoment beschreven in een kruistabel. Vervolgens werden met scatterplots en bijbehorende correlaties de bivariate relaties onderzocht (zie horizontale pijlen en stippellijnen in Figuur 1).

Ten slotte werd, in verband met de op verschillende niveaus geclusterde data (school-, leerling- en toetsniveau), een multilevel regressieanalyse uitgevoerd om de hoofdvraag te kunnen beantwoorden. Hierbij werd uitgegaan van een random intercept model dat voor iedere leerling een aparte regressielijn schatte, waarbij het intercept (de gemiddelde groei in rekenvaardigheid) kon verschillen per leerling en per school. Er werden twee modellen geschat. Het eerste model was een leeg random intercept model om de verdeling van de variantie tussen scholen, leerlingen en toetsmomenten te beschrijven. Vervolgens werd het tweede model geschat op basis van het conceptuele model in Figuur 1 door groei in begrijpend lezen en groei in technisch lezen aan het eerste model toe te voegen. Aangezien groei in begrijpend lezen en groei in technisch lezen gecorreleerd waren, werd ook gekeken naar modellen waarin alleen groei in begrijpend leesvaardigheid dan wel technische leesvaardigheid werd meegenomen. Aan de volgende drie assumpties moest voldaan worden: lineariteit, homoscedasticiteit en normaliteit van de residuen op school-, leerling- en toetsniveau. De eerste twee assumpties werden gecontroleerd aan de hand van een scatterplot van de gefitte waarden uitgezet tegen de residuen op toetsniveau en de laatste assumptie aan de hand van Q-Q plots van de verschillende residuen. Om te onderzoeken hoe de kans op stagnatie voorspeld werd, werd ook een logistische regressieanalyse uitgevoerd. Hierbij werden dezelfde stappen doorlopen als tijdens de multilevel regressieanalyse. De assumptie betreffende normaliteit van de residuen op toetsniveau verviel, maar verder moest eveneens aan dezelfde assumpties voldaan worden. Een deviance test werd uitgevoerd om de fit van de verschillende modellen met elkaar te kunnen vergelijken, waarbij een grotere deviance staat voor een slechtere fit van het model (Snijders & Bosker, 2012). De significantie werd steeds bepaald door de standaarderror te vermenigvuldigen met 1,96. Een waarde werd als statistisch significant gezien als de uitkomst hiervan kleiner was dan de absolute waarde van bijbehorende regressiecoëfficiënt (Snijders & Bosker, 2012).

## **Resultaten**

### **Problemen in de dataset**

Voor één school bleken de momenten waarop de verschillende toetsen zijn afgenomen niet goed gematched te kunnen worden. Om die reden is ervoor gekozen de data van die school uit de dataset te verwijderen. Het gaat hierbij om de data van 31 kinderen.

In Tabel 1 is voor de onafhankelijke variabelen ‘begrijpend lezen’ en ‘technisch lezen’ en voor de afhankelijke variabele ‘rekenen’ te zien wat de proporties en de absolute aantallen niet-missende en missende vaardigheidsscores zijn. Hierin is te zien dat er de meeste vaardigheidsscores missen voor de variabele ‘begrijpend lezen’, doordat deze toets niet op alle toetsmomenten wordt afgenomen. Voor de variabele ‘begrijpend lezen’ missen er dan ook aanzienlijk veel vaardigheidsscores voor de toetsmomenten midden groep 3 (M3; *prop.* = ,970 en *n* = 1183) en eind groep 5 (E5; *prop.* = ,926 en *n* = 1182).

**Tabel 1**

*(Niet-)Missende Vaardigheidsscores per Vaardigheidsgebied*

Variabele	Niet-missend <i>prop. (n)</i>	Missend <i>prop. (n)</i>
Begrijpend lezen	,602 (4542)	,398 (3005)
Technisch lezen	,823 (6210)	,177 (1337)
Rekenen	,992 (7488)	,008 (59)

Tabel 2 toont de relaties tussen niet-missende en missende data voor de vaardigheidsscores van de variabelen ‘begrijpend lezen’, ‘technisch lezen’ en ‘rekenen’ en de kindkenmerken ‘geslacht’, ‘opleiding ouders’ en ‘migratieachtergrond’. Hieruit blijkt dat de toets Begrijpend lezen bij kinderen met een niet-Nederlandse achtergrond iets vaker wordt afgenomen (66,3%) dan bij kinderen met een Nederlandse achtergrond (59,9%), maar verder lijkt er geen sprake te zijn van opvallende relaties tussen de kindkenmerken en missende data.

**Tabel 2***(Niet-)Missende Data per Vaardigheidsgebied uitgesplitst naar Kindkenmerken*

Variabele	Begrijpend lezen		Technisch lezen		Rekenen	
	Niet-missend <i>prop. (n)</i>	Missend <i>prop. (n)</i>	Niet-missend <i>prop. (n)</i>	Missend <i>prop. (n)</i>	Niet-missend <i>prop. (n)</i>	Missend <i>prop. (n)</i>
<b>Geslacht</b>						
jongen	,599 (2234)	,401 (1493)	,827 (3083)	,173 (644)	,995 (3710)	,005 (17)
meisje	,608 (2308)	,392 (1488)	,824 (3127)	,176 (669)	,995 (3778)	,005 (18)
<b>Opleiding ouders</b>						
(zeer) laag	,630 (386)	,370 (227)	,847 (519)	,153 (94)	,997 (611)	,003 (2)
normaal	,599 (3975)	,401 (2661)	,816 (5417)	,184 (1219)	,995 (6603)	,005 (33)
<b>Migratieachtergrond</b>						
Nederlands	,599 (4125)	,401 (2765)	,826 (5691)	,174 (1199)	,996 (6862)	,004 (28)
niet-Nederlands	,663 (408)	,337 (207)	,815 (501)	,185 (114)	,992 (610)	,008 (5)

**Beschrijving steekproef**

De uiteindelijke steekproef bestaat uit 1248 kinderen. Op het eerste toetsmoment was de gemiddelde leeftijd 6;10 jaar ( $SD = 5$  maanden) en op het laatste toetsmoment 9;2 jaar ( $SD = 5$  maanden). In Tabel 3 is een beschrijving van de steekproef te vinden met betrekking tot de kindkenmerken ‘geslacht’, ‘opleiding ouders’ en ‘migratieachtergrond’. Uit Tabel 3 blijkt dat de steekproef uit ongeveer evenveel jongens als meisjes bestaat en grotendeels kinderen met normaal opgeleide ouders en kinderen met een Nederlandse achtergrond omvat.

**Tabel 3***Verdeling kinderen over Kindkenmerken (N = 1248)*

Variabele	Proportie	Absolute aantal
Geslacht		
jongen	,496	619
meisje	,504	629
Opleiding ouders		
(zeer) laag	,082	103
normaal	,880	1098
Migratieachtergrond		
Nederlands	,904	1128
niet-Nederlands	,093	116

*Noot.* In verband met missende data tellen niet alle proporties op tot 1,000.

Tabel 4 toont voor de variabelen ‘begrijpend lezen’, ‘technisch lezen’ en ‘rekenen’ per toetsmoment de gemiddelde vaardigheidsscore en bijbehorende standaarddeviatie van zowel de steekproef als de populatie, waarbij de populatiewaarden afkomstig zijn uit de normeringsonderzoeken van Cito. Hierin valt allereerst op dat er voor technisch lezen en rekenen voor ieder toetsmoment sprake is van een toename in de gemiddelde vaardigheidsscore, oftewel van een gemiddelde groei. Voor begrijpend lezen is dit voor de meeste toetsmomenten ook het geval. Alleen op de toetsmomenten E3 en E5 is er voor begrijpend lezen gemiddeld gezien sprake van stagnatie. Ten tweede valt op dat de standaarddeviaties voor rekenen voor ieder toetsmoment kleiner worden. Voor rekenen neemt de variabiliteit van de individuele vaardigheidsscores ten opzichte van de gemiddelde vaardigheidsscore dus iets af met de tijd.

Wanneer de steekproef vergeleken wordt met de populatie, valt op dat de gemiddelde vaardigheidsscores van de steekproef aanzienlijk hoger liggen dan die van de populatie. Daarnaast valt op dat de standaarddeviaties van de steekproef op het gebied van rekenen verschillen van de populatiestandaarddeviaties. De individuele vaardigheidsscores binnen de steekproef wijken op het gebied van rekenen meer af van het gemiddelde dan binnen de populatie het geval is.

**Tabel 4***Gemiddelde Vaardigheidsscores en Standaarddeviaties per Toets(moment)*

Variabele	Toetsmoment	Steekproef		Populatie	
		Gemiddelde ( <i>n</i> )	Standaarddeviatie	Gemiddelde ( <i>n</i> )	Standaarddeviatie
Begrijpend lezen	M3	1,89 (37)	15,74	-	-
	E3	1,80 (972)	14,17	-2,40 (443)	15,63
	M4	15,01 (1086)	13,33	8,91 (785)	14,29
	E4	20,39 (1148)	14,44	13,18 (674)	15,28
	M5	26,96 (1205)	14,15	22,32 (819)	13,91
	E5	26,43 (94)	15,95	-	-
Technisch lezen	M3	25,80 (951)	14,69	23,28 (1025)	14,50
	E3	39,38 (907)	17,07	35,98 (942)	17,56
	M4	55,51 (1161)	19,14	52,75 (1018)	19,04
	E4	64,10 (1090)	17,79	61,45 (920)	18,15
	M5	72,19 (1112)	17,99	70,29 (954)	16,90
	E5	79,51 (989)	15,76	74,74 (811)	16,26
Rekenen	M3	35,41 (1218)	16,10	26,04 (1304)	14,60
	E3	45,37 (1187)	15,28	34,80 (1032)	14,60
	M4	54,52 (1267)	14,88	47,00 (1215)	14,68
	E4	64,87 (1254)	14,79	56,44 (948)	14,64
	M5	73,88 (1294)	14,72	69,00 (981)	14,48
	E5	81,22 (1268)	14,03	74,08 (970)	14,52

*Noot.* De toets Begrijpend lezen wordt veelal niet afgenomen op de toetsmomenten M3 en E5, waardoor er voor deze toetsmomenten geen normeringsgegevens beschikbaar zijn.

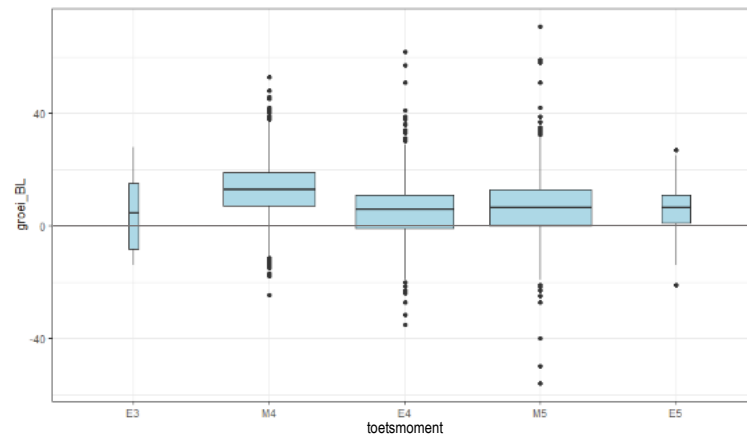
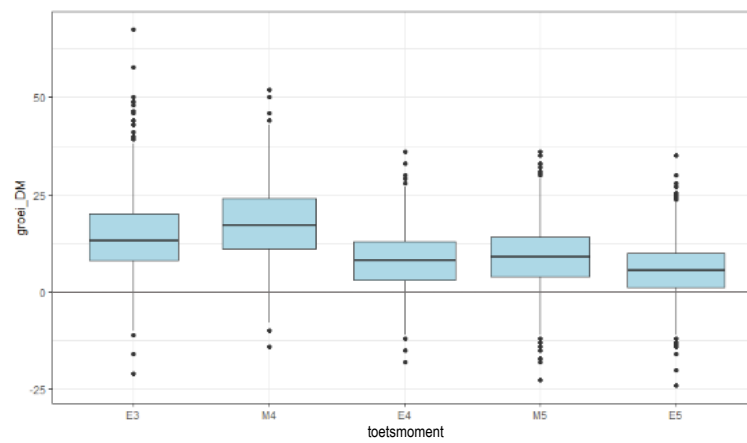
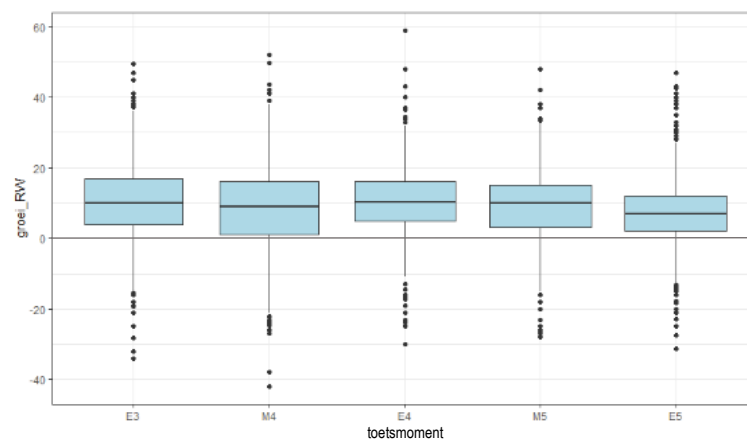
De centrum- en spreidingsmaten die de verdeling van de groei voor begrijpend lezen, technisch lezen en rekenen beschrijven staan in Tabel 5. Hierin is te zien dat de gemiddelde groei op alle vaardigheidsgebieden ongeveer gelijk is. Daarnaast is te zien dat alle minimumwaarden negatief zijn en dat er dus op alle vaardigheidsgebieden leerlingen zijn die een negatieve groei vertonen, oftewel stagneren. Voor begrijpend lezen is de proportie stagnatie ,115 ( $n = 547$ ), voor technisch lezen ,231 ( $n = 752$ ) en voor rekenen ,170 ( $n = 1050$ ). In verhouding stagneren dus de meeste leerlingen op het gebied van technisch lezen, maar de meest extreme waardes doen zich voor op het gebied van begrijpend lezen.



**Tabel 5***Centrum- en Spreidingsmaten Groeivariabelen*

Variabele	Gemiddelde	Mediaan	Minimum	Maximum	Standaarddeviatie
Begrijpend lezen	8,18	8,00	-56,00	71,00	11,00
Technisch lezen	10,63	10,00	-24,00	67,50	9,40
Rekenen	9,18	9,00	-42,00	59,00	10,09

In Figuur 2, 3 en 4 is voor de verschillende vaardigheidsgebieden de groei per toetsmoment beschreven. Uit deze figuren blijkt dat de mediaan voor alle groeivariabelen en toetsmomenten positief is. Met andere woorden, het grootste deel van de leerlingen maakt op alle vaardigheidsgebieden en voor alle toetsmomenten een positieve groei door. Voor rekenen is de groei per toetsmoment voor de meeste leerlingen redelijk constant. De groei per toetsmoment voor zowel begrijpend lezen als technisch lezen is voor de meeste leerlingen minder constant. Daarnaast blijkt uit de figuren dat er sprake is van veel spreiding op het gebied van begrijpend lezen op toetsmoment M5 (*min.* = -56,00 en *max.* = 71,00) en op het gebied van technisch lezen op toetsmoment E3 (*min.* = -21,00 en *max.* = 67,50). Uit zowel de figuren als Tabel 6 blijkt verder dat het grootste deel van de leerlingen op alle vaardigheidsgebieden en op alle toetsmomenten groeit en niet stagneert. Ook valt op dat er op het gebied van begrijpend lezen sprake is van een flinke toename in stagnaties op toetsmoment E4 ten opzichte van toetsmoment M4. Daarentegen is er op het gebied van rekenen op toetsmoment E4 sprake van een afname in stagnaties. Ten slotte blijkt uit de figuren dat er veel uitbijters zijn. Opvallend daaraan is dat er voor technisch lezen geen extreme uitbijters voorkomen aan de onderkant (groei < -30) en dat er aan de bovenkant de meeste extreme uitbijters voorkomen voor begrijpend lezen (groei > 50).

**Figuur 2***Groei Begrijpend Lezen per Toetsmoment***Figuur 3***Groei Technisch Lezen per Toetsmoment***Figuur 4***Groei Rekenen per Toetsmoment*

**Tabel 6***Stagnatie per Toets(moment)*

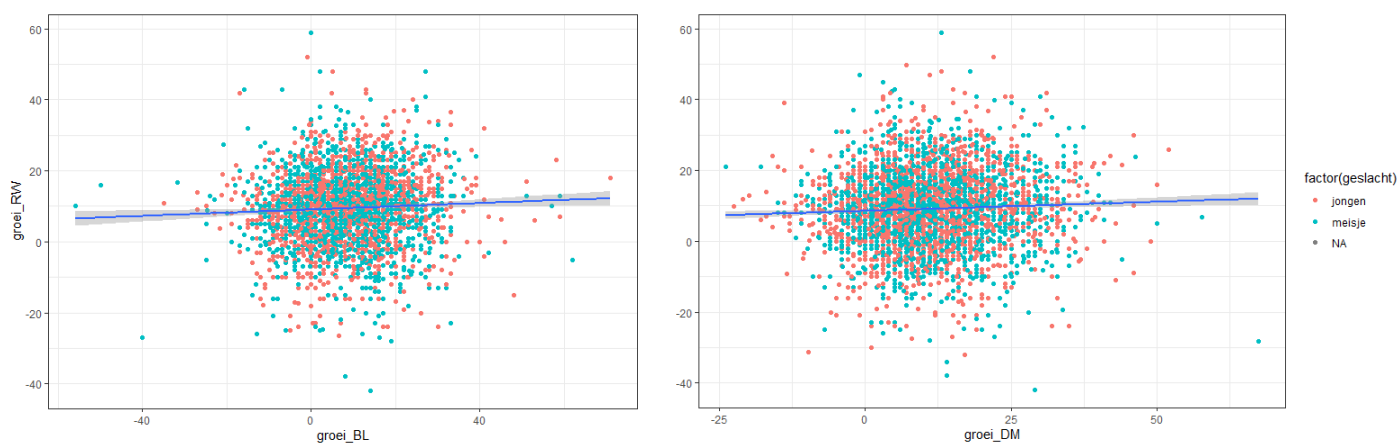
Toetsmoment	Begrijpend lezen		Technisch lezen		Rekenen	
	Groei	Stagnatie	Groei	Stagnatie	Groei	Stagnatie
	<i>prop. (n)</i>	<i>prop. (n)</i>	<i>prop. (n)</i>	<i>prop. (n)</i>	<i>prop. (n)</i>	<i>prop. (n)</i>
E3	,583 (7)	,417 (5)	,961 (780)	,039 (32)	,841 (994)	,159 (188)
M4	,905 (841)	,095 (88)	,972 (875)	,028 (25)	,761 (929)	,239 (292)
E4	,694 (737)	,306 (325)	,858 (914)	,142 (151)	,892 (1112)	,108 (134)
M5	,730 (842)	,270 (312)	,866 (867)	,134 (134)	,839 (1061)	,161 (204)
E5	,766 (72)	,234 (22)	,788 (760)	,212 (205)	,817 (1035)	,183 (232)

**Bivariate relaties**

Er is een zwak positief verband te zien tussen groei in begrijpend lezen en groei in rekenen ( $r = ,049$ ;  $p = ,006$ ) en tussen groei in technisch lezen en groei in rekenen ( $r = ,048$ ;  $p < ,001$ ). Tussen groei in begrijpend lezen en groei in technisch lezen is een wat sterker positief verband te zien ( $r = ,208$ ;  $p < ,001$ ). Deze relaties, met uitzondering van de laatstgenoemde, zijn weergegeven in Figuur 5, 6 en 7. Wat deze figuren ook laten zien is of de relaties verschillen wanneer deze uitgesplitst worden naar de kindkenmerken. De invloed van de kindkenmerken ‘geslacht’, ‘opleiding ouders’ en ‘migratieachtergrond’ op de relatie tussen groei in begrijpend lezen dan wel groei in technisch lezen en groei in rekenen lijkt op basis van de figuren minimaal.

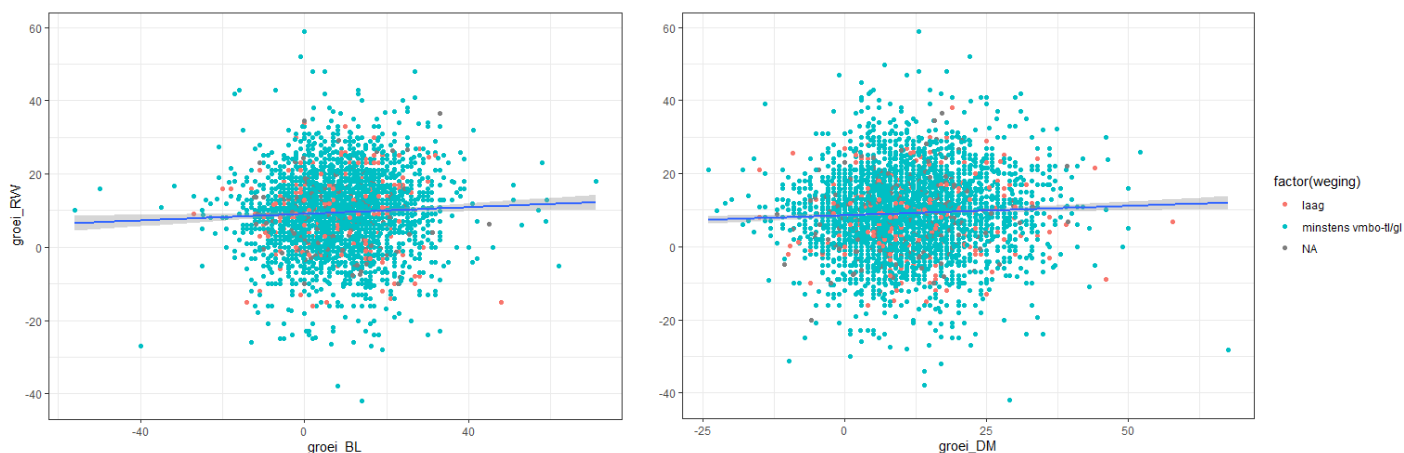
**Figuur 5**

*Relatie tussen Groei in Begrijpend Lezen/Technisch Lezen en Groei in Rekenen uitgesplitst naar Geslacht*

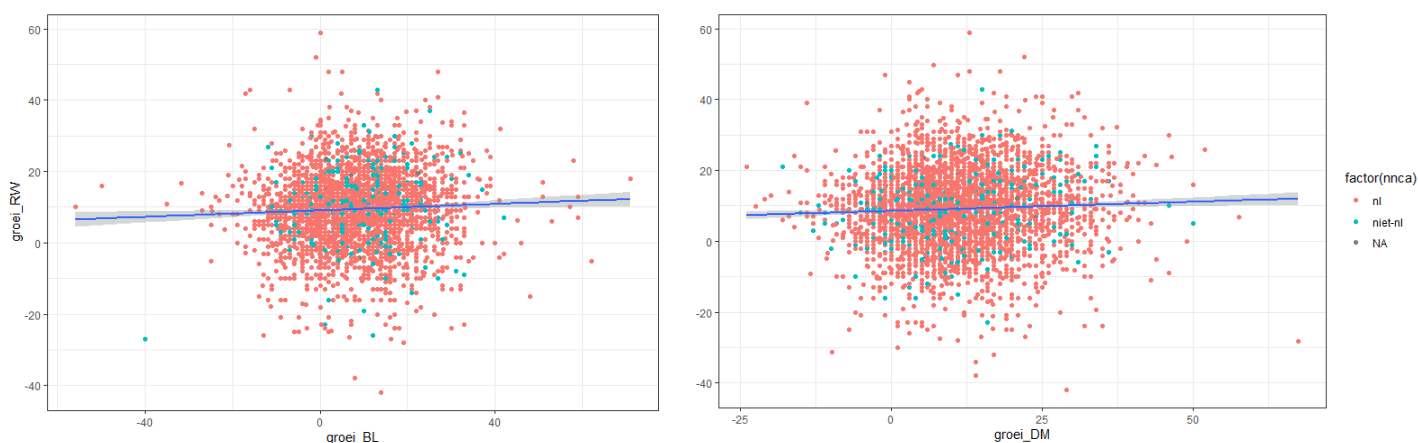


**Figuur 6**

*Relatie tussen Groei in Begrijpend Lezen/Technisch Lezen en Groei in Rekenen uitgesplitst naar Opleiding Ouders*

**Figuur 7**

*Relatie tussen Groei in Begrijpend Lezen/Technisch Lezen en Groei in Rekenen uitgesplitst naar Migratieachtergrond*



## Toetsresultaten

### *Multilevel regressieanalyse*

Uit het lege random intercept model (Model 1) blijkt dat de variantie op leerlingniveau op nul wordt geschat. Op leerlingniveau zijn er dus weinig systematische verschillen in gemiddelde groei in rekenvaardigheid. Elke leerling groeit min of meer volgens dezelfde tendens. Aangezien er op leerlingniveau geen variantie is om te verklaren, wijst dit er tevens op dat de kindkenmerken een minimale invloed hebben op de relatie tussen groei in begrijpend lezen dan wel groei in technisch lezen en groei in rekenen. Om die reden is ervoor gekozen het leerlingniveau niet mee te nemen in de modellen en de variantie in groei te beschrijven op school- en toetsniveau. De schattingen voor deze modellen (zonder leerlingniveau) zijn gerapporteerd in Tabel 7.

Uit Model 1 blijkt dat leerlingen een gemiddelde groei in rekenvaardigheid laten zien van 9,40 punten. Op een tweede toetsmoment is hun vaardigheidsscore dus met 9,40 punten toegenomen ten opzichte van een eerste toetsmoment. Op een gemiddelde school groeit de beste 16% gemiddeld 19,36 punten (zie Tabel 7, 1 *SD* boven het gemiddelde) en de slechtste 16% -0,56 punten (zie Tabel 7, 1 *SD* onder het gemiddelde), oftewel de slechtste 16% stagneert. Model 1 laat ook zien dat de variatie in gemiddelde groei in rekenvaardigheid veel groter is tussen toetsmomenten dan tussen scholen. Bij Model 1 hoort een intraclass correlatie van  $1,03^2/(1,03^2+9,96^2) = ,011$ . Daaruit blijkt dat toetsscores van verschillende leerlingen binnen dezelfde school nauwelijks met elkaar samenhangen.

Het model waaraan groei in begrijpend lezen en groei in technisch lezen toegevoegd zijn (Model 2), laat een significant betere fit zien dan Model 1 ( $\chi^2(2) = 7,443; p = ,024$ ) en haalt wat variantie weg tussen scholen en toetsmomenten. Met andere woorden, scholen en toetsmomenten lijken iets meer op elkaar in rekengroei wanneer rekening gehouden wordt met verschillen in groei op het gebied van begrijpend lezen en technisch lezen. De intraclass correlatie bij Model 2 is ,010 en is dus vergelijkbaar met de intraclass correlatie bij Model 1. Uit Model 2 blijkt dat leerlingen die geen groei vertonen in begrijpend lezen en technisch lezen een gemiddelde groei in rekenvaardigheid laten zien van 8,78 punten. Daarnaast blijkt dat zij op het gebied van rekenen gemiddeld 0,04 punten groeien voor elk punt dat zij groeien op het gebied van begrijpend lezen en 0,03 punten voor elk punt dat zij groeien op het gebied van technisch lezen, waarbij de laatstgenoemde schatting niet significant is. Uit de modellen waarin alleen groei in begrijpend leesvaardigheid dan wel technische leesvaardigheid meegenomen wordt, blijkt eveneens dat begrijpend lezen de sterkere voorspeller is. Verder blijkt dat de effecten iets zwakker worden geschat wanneer beide voorspellers in één model worden opgenomen.

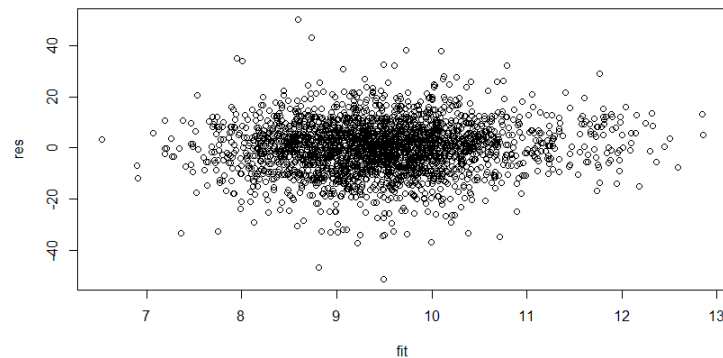
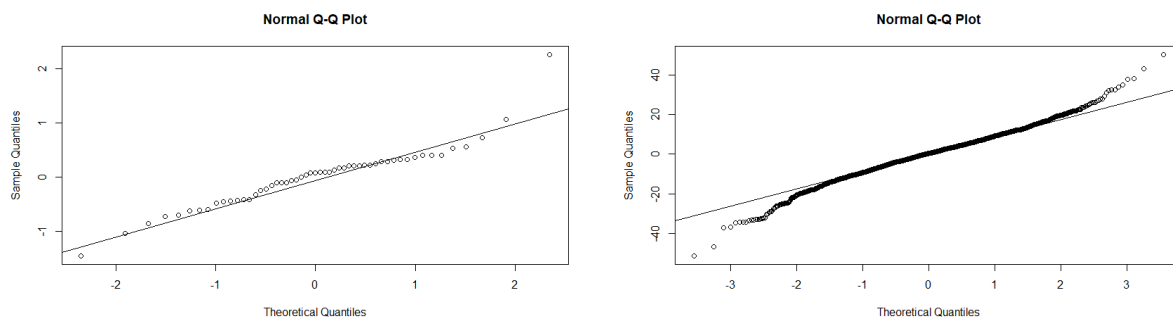
**Tabel 7**

*Schattingen voor Modellen met Schoolniveau (N = 53) en Toetsniveau (N = 2630) en Rekengroei als Afhankelijke Variabele*

Fixed effecten	Model 1		Model 2	
	Coëfficiënt	SE	Coëfficiënt	SE
Intercept	9,40*	0,255	8,78*	0,355
Groei begrijpend lezen			0,04*	0,018
Groei technisch lezen			0,03	0,022
Random effecten				
SD school	1,03		1,02	
SD toetsmoment	9,96		9,95	
Deviance	19.575,00		19.568,00	

*Noot.* Statistisch significante waarden zijn aangegeven met \*.

De assumpties van Model 2 zijn gecontroleerd aan de hand van Figuur 8 en 9. Het verband in Figuur 8 is goed te schatten met een rechte lijn. Bovendien is de spreiding in verticale richting ongeveer gelijk, waarbij de meeste extreme residuen aan zowel de boven- als onderkant afkomstig zijn van één school. Er is dus sprake van lineariteit. Voor hoger gefitte waarden lijkt de spreiding van de residuen echter iets kleiner, waardoor mogelijk sprake is van een lichte schending van de assumptie homoscedasticiteit. Figuur 9 laat de verdeling van de residuen op school- en toetsniveau zien. De residuen op toetsniveau lijken op basis van Figuur 9 normaal verdeeld. Daarentegen lijken de residuen op schoolniveau enigszins scheef verdeeld. Dit lijkt te komen door een uitbijter aan de rechterkant van de verdeling. Na verwijdering van deze uitbijter lijken de residuen op schoolniveau namelijk wel normaal verdeeld.

**Figuur 8***Controle Lineariteit en Homoscedasticiteit***Figuur 9***Controle Normaliteit van de Residuen op Schoolniveau (links) en Toetsniveau (rechts)***Logistische regressieanalyse**

De variantie op leerlingniveau bleek wederom op nul geschat te worden. Daarom is in de logistische modellen ook alleen uitgegaan van school- en toetsniveau, waarbij op toetsniveau geen variantie en standaarddeviatie bepaald konden worden. Tabel 8 toont de schattingen voor deze modellen (zonder leerlingniveau).

Uit de modellen blijkt dat de kans dat leerlingen stagneren op het gebied van rekenen  $\exp(-1,61)/(1+\exp(-1,61)) = 0,17$  is. De meeste leerlingen zullen dus niet stagneren op het gebied van rekenen. Daarnaast blijkt uit Log Model 2 dat stagneren op het gebied van rekenen gemiddeld  $\exp(0,04) = 1,04$  keer zo waarschijnlijk is als leerlingen stagneren op het gebied van begrijpend lezen en  $1/\exp(-0,08) = 1,08$  keer minder waarschijnlijk als ze stagneren op het gebied van technisch lezen. Dit blijkt ook zo te zijn wanneer gekeken wordt naar een model waarin alleen stagnatie in begrijpend leesvaardigheid dan wel technische leesvaardigheid wordt meegenomen. Echter, een kanttekening die hierbij gemaakt dient te worden is dat de schattingen waarop deze kansen zijn gebaseerd niet significant zijn. Het is eveneens lastig te bepalen welk model een betere fit laat zien, doordat er sprake is van een klein en niet significant verschil in deviance tussen de modellen ( $\chi^2(2) = 0,297; p = ,862$ ).

**Tabel 8**

*Schattingen voor Logistisch Modellen met Schoolniveau (N = 53) en Toetsniveau (N = 2630) en Rekengroei als Afhankelijke Variabele*

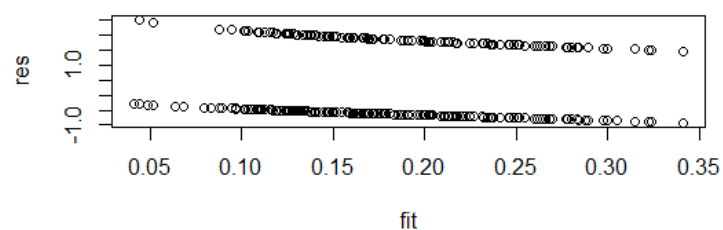
Fixed effecten	Log Model 1		Log Model 2	
	Coëfficiënt	SE	Coëfficiënt	SE
Intercept	-1,61*	0,07	-1,61*	0,081
Stagnatie begrijpend lezen			0,04	0,123
Stagnatie technisch lezen			-0,08	0,169
Random effecten				
SD school	0,32		0,32	
Deviance	2.378,30		2.378,00	

*Noot.* Statistisch significante waarden zijn aangegeven met \*.

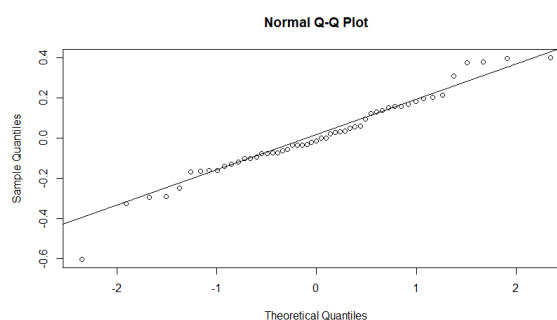
Aan de hand van Figuur 10 en 11 zijn de assumpties van Log Model 2 gecontroleerd. Figuur 10 levert geen aanwijzingen op voor een schending van de assumpties betreffende lineariteit en homoscedasticiteit. Op basis van Figuur 11 lijken de residuen op schoolniveau door eerdergenoemde uitbijter wederom enigszins scheef verdeeld. Na verwijdering van deze uitbijter lijken de residuen op schoolniveau normaler verdeeld.

**Figuur 10**

*Controle Lineariteit en Homoscedasticiteit*

**Figuur 11**

*Controle Normaliteit van de Residuen op Schoolniveau*





## Conclusie en discussie

Door middel van dit onderzoek is gezocht naar een relatie tussen groei in vaardigheden op het gebied van rekenen en groei in vaardigheden op het gebied van lezen tussen groep 3 en 5. Uit de resultaten komt een significant positief verband naar voren tussen groei in begrijpend lezen en groei in rekenen, oftewel een hogere groei in begrijpend lezen hangt samen met een hogere groei in rekenen en een lagere groei in begrijpend lezen met een lagere groei in rekenen. Echter, in vergelijking met de gemiddelde groei in rekenvaardigheid wordt slechts een relatief klein deel van de rekgroei verklaard door de groei in begrijpend lezen. De waarde van de standaarderror is bovendien bijna half zo groot als de waarde van de coëfficiënt behorende bij de groei in begrijpend lezen. Er is dus sprake van een zwak en onzeker verband. Tussen groei in technisch lezen en groei in rekenen is geen significante samenhang gevonden. Een relatie tussen stagnatie, oftewel uitblijvende groei, in rekenvaardigheid en stagnatie in leesvaardigheid kan op basis van dit onderzoek ook niet worden aangetoond (deelvraag 2). Verder blijken er op leerlingniveau weinig systematische verschillen te zijn, wat erop wijst dat elke leerling min of meer volgens dezelfde tendens groeit. Afwijkingen lijken dus veelal niet systematisch te zijn. De invloed van de kindkenmerken ‘geslacht’, ‘opleiding ouders’ en ‘migratieachtergrond’ op de relatie tussen groei in vaardigheden op het gebied van rekenen en groei in vaardigheden op het gebied van lezen lijkt daardoor eveneens minimaal (deelvraag 1).

Daarentegen waren de verwachtingen op basis van de literatuur dat er sprake zou zijn van een sterke positieve relatie tussen groei in rekenvaardigheid en groei in leesvaardigheid (Prenger, 2011; Krajewski & Schneider, 2009; LeFevre et al., 2010; Vygotsky, 1986; Langeveld, 1934), evenals tussen stagnatie in rekenvaardigheid en stagnatie in leesvaardigheid (Ghesquière & Ruijsenaars, 2021), en dat deze relatie het sterkste zou zijn wanneer het specifiek zou gaan om technische leesvaardigheid (Ruijsenaars et al., 2004). Mogelijk zijn de resultaten van dit onderzoek niet geheel in lijn met wat op basis van de literatuur werd verwacht, doordat de focus in de literatuur ligt op vaardigheidsniveau en in dit onderzoek op vaardigheidsgroei. In de meeste gevallen komt de op basis van een toets geschatte vaardigheidsscore van een leerling overeen met zijn werkelijke vaardigheidsniveau (Feenstra et al., 2010; Krom et al., 2010; Janssen et al., 2010). Er bestaat echter altijd een kans op een zekere meetonnauwkeurigheid. Doordat de vaardigheidsgroei van een leerling wordt bepaald op basis van twee vaardigheidsscores die van zichzelf al ietwat onzeker zijn, neemt de kans op een meetonnauwkeurigheid in vaardigheidsgroei toe. Mogelijk is er daardoor sprake van een sterkere positieve relatie tussen reken- en leesprestaties dan tussen

reken- en leesgroei en/of is die relatie makkelijker aan te tonen. Een mogelijke aanvullende verklaring voor het uitblijven van de relatie tussen stagnatie in rekenvaardigheid en stagnatie in leesvaardigheid is dat Ghesquière en Ruijsenaars (2021) onderzoek deden naar leerlingen met ernstige reken-, lees- en/of spellingsproblemen, terwijl dit onderzoek zich richtte op gemiddelde leerlingen. Voor dit onderzoek werd dan ook gebruik gemaakt van meetinstrumenten die bij alle leerlingen konden worden afgenomen. Mogelijk zijn deze meetinstrumenten minder geschikt voor het in kaart brengen van een relatie betreffende stagnatie dan bijvoorbeeld diagnostische meetinstrumenten.

Op basis van de literatuur (Vogelzang, 2018; Schmeier, 2017; Meelissen et al., 2023; Hickendorff & Janssen, 2009) werd bovendien verwacht dat de kindkenmerken ‘geslacht’, ‘opleiding ouders’ en ‘migratieachtergrond’ voorspellend zouden zijn ten aanzien van het schools presteren van kinderen in het basisonderwijs. Mogelijk komen de resultaten van dit onderzoek niet overeen met deze verwachting, doordat dergelijke kindkenmerken alleen een voorspellende waarde hebben wanneer het gaat om vaardigheidsniveau en vervalst deze wanneer gekeken wordt naar vaardigheidsgroei. Het vaardigheidsniveau van een zwakke leerling is immers lager dan dat van een sterke leerling, maar dit hoeft niet per definitie ook te resulteren in een kleiner verschil in vaardigheidsniveau tussen twee toetsmomenten. Mogelijk verklaart dat ook waardoor er op basis van dit onderzoek op leerlingniveau weinig systematische verschillen blijken te zijn. Een andere mogelijke verklaring die tevens in lijn zou zijn met de resultaten van het onderzoek van Hickendorff en Janssen (2009) is dat de voorspellende waarde van de kindkenmerken alleen van toepassing is wanneer het gaat om contextopgaven en niet, zoals bijvoorbeeld in het geval van de toets Rekenen-Wiskunde, om een combinatie van context- en kale opgaven. Op de taalvaardigheid van een leerling wordt namelijk alleen een beroep gedaan tijdens het maken van contextopgaven.

Voor dit onderzoek is gebruik gemaakt van meetinstrumenten die door de COTAN zijn beoordeeld als valide en betrouwbaar. Dat betekent dat waarschijnlijk gemeten is wat beoogd werd te meten en dat de resultaten waarschijnlijk hetzelfde zullen zijn wanneer het onderzoek op dezelfde manier nogmaals wordt uitgevoerd. Voorafgaand aan het uitvoeren van de analyses werden doubleerders uit de dataset verwijderd om groei over de periode tussen groep 3 en 5 te kunnen berekenen. De resultaten van dit onderzoek kunnen dan ook gegeneraliseerd worden naar de populatie niet-gedoubleerde leerlingen in groep 3 tot en met 5 van het reguliere Nederlandse basisonderwijs.

Uit de missende data analyse bleek dat er de meeste vaardigheidsscores missen voor de variabele ‘begrijpend lezen’. Dit komt doordat deze toets veelal niet werd afgenomen op

de toetsmomenten M3 en E5. Daarnaast bleek dat de toets Begrijpend lezen iets vaker werd afgenomen bij kinderen met een niet-Nederlandse achtergrond dan bij kinderen met een Nederlandse achtergrond, wat kan wijzen op de aanwezigheid van patronen in de missende data. De steekproef bevat echter veel meer kinderen met een Nederlandse achtergrond dan kinderen met een niet-Nederlandse achtergrond, waardoor dit waarschijnlijk weinig invloed heeft gehad op de conclusies.

De (logistische) modellen zijn gebaseerd op toetsmomenten waarvan data beschikbaar waren voor rekenen, begrijpend lezen én technisch lezen, omdat dit het mogelijk maakte de verschillende modellen met elkaar te vergelijken. Aangezien er relatief veel vaardigheidsscores missen voor de variabele 'begrijpend lezen', vielen er dus redelijk wat toetsmomenten weg. Door de grote steekproefomvang en de daaruit voortvloeiende statistische power was de kans op het vinden van kleine effecten daarentegen nog steeds vrij groot. Bij het trekken van conclusies op basis van dit onderzoek moet wel enige voorzichtigheid geboden worden. Er lijkt namelijk niet volledig voldaan te worden aan de assumpties voor een multilevel regressieanalyse. Zo is er mogelijk sprake van een lichte vorm van heteroscedasticiteit. Echter, er zijn minder datapunten beschikbaar voor hoger gefitte waarden, waardoor lastig te zeggen is in hoeverre dit daadwerkelijk het geval is. Daarnaast is er voor zowel de multilevel regressieanalyse als de logistische regressieanalyse mogelijk sprake van een lichte vorm van niet-normaliteit van de residuen op schoolniveau, maar dit lijkt in beide gevallen veroorzaakt te worden door een uitbijter.

Bij het schatten van de modellen is ervan uitgegaan dat groei in rekenvaardigheid voorspeld kan worden op basis van groei in leesvaardigheid. Een dergelijk causaal verband, zoals ook weergegeven is in Figuur 1, is op basis van dit onderzoek alleen lastig aan te tonen. Hiervoor moet namelijk ook gekeken worden naar de invloed van eventuele derde variabelen op de relatie. Aangezien er op leerlingniveau weinig systematische verschillen blijken te zijn, gaat het dan om derde variabelen die betrekking hebben op het toetsniveau.

Vervolgonderzoek zou bijvoorbeeld gedaan kunnen worden naar de rol van motivationele factoren in het proces van groei en stagnatie die fluctueren per moment en/of over tijd. Zo is van dyslectische leerlingen bekend dat lezen hen zoveel moeite kost dat ze het plezier erin verliezen en het lezen en mogelijk ook het rekenen vermijden, waardoor deze vaardigheden zich onvoldoende ontwikkelen (Van Koeven & Smits, 2020).

Een zeer opmerkelijke nevenbevinding van dit onderzoek is dat er op leerlingniveau weinig systematische verschillen blijken te zijn. Stagnatie is dus veelal een eenmalig fenomeen en er is een heel kleine groep leerlingen die echt systematisch achterloopt. Dat

maakt het moeilijk om de leerlingen die systematisch achterlopen te onderscheiden van de leerlingen die een sprongsgewijze ontwikkeling doormaken. Vervolgonderzoek zou dan ook moeten uitwijzen hoe de leerlingen die systematisch achterlopen gesignaleerd en goed in kaart gebracht kunnen worden.

Kortom, uit dit onderzoek blijkt dat er sprake is van een zwak en onzeker verband tussen groei in rekenvaardigheid en groei in begrijpend leesvaardigheid. Daarnaast blijkt het signaleren van leerlingen die systematisch achterlopen lastig te zijn. Dit onderstreept, evenals het feit dat de nadruk in het nieuwste leerlingvolgsysteem van Cito (2021) sterk op groei is komen te liggen, het belang van het verder onderzoeken van de mogelijkheden voor het aan de hand van groei vroegtijdig opsporen van problemen in de ontwikkeling van zowel taal- als rekenvaardigheden. Tot die tijd is het belangrijk dat leerkrachten niet alleen gebruik maken van toetsgegevens om informatie in te winnen over de voortgang van een leerling, maar bijvoorbeeld ook van observaties, gesprekken met ouders, etc.

## Referenties

- Besluit bekostiging WPO. (2022, 1 april). Geraadpleegd van [https://wetten.overheid.nl/BWBR0046159/2022-04-01/0#Hoofdstuk1\\_Artikel1](https://wetten.overheid.nl/BWBR0046159/2022-04-01/0#Hoofdstuk1_Artikel1)
- CBS. (2017, februari). *Herziening gewichtenregeling primair onderwijs*. CBS. Geraadpleegd op 17 december 2023, van [https://www.cbs.nl/-/media/\\_pdf/2017/06/gewichtenregeling-hoofdlijnenrapport.pdf](https://www.cbs.nl/-/media/_pdf/2017/06/gewichtenregeling-hoofdlijnenrapport.pdf)
- Cito. (2021). *Leerling in beeld*. Cito. Geraadpleegd op 18 december 2023, van <https://cito.nl/onderwijs/primair-onderwijs/leerling-in-beeld-lvs/>
- COTAN. (2009). *Beoordelingssysteem voor de kwaliteit van tests*. NIP; COTAN. [https://www.researchgate.net/publication/254836210\\_COTAN\\_Beoordelingssysteem\\_voor\\_de\\_kwaliteit\\_van\\_tests](https://www.researchgate.net/publication/254836210_COTAN_Beoordelingssysteem_voor_de_kwaliteit_van_tests)
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3-6), 487–506. <https://doi.org/10.1080/02643290244000239>
- Feenstra, H., Kamphuis, F., Kleintjes, F., & Krom, R. (2010). *Wetenschappelijke verantwoording: Begrijpend lezen voor groep 3 tot en met 6*. Cito. <https://docplayer.nl/370648-Wetenschappelijke-verantwoording-begrijpend-lezen-voor-groep-3-tot-en-met-6.html>
- Frans, N. (2019). *A captivating snapshot of standardized testing in early childhood: on the stability and utility of the Cito preschool/kindergarten tests*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen. <https://doi.org/10.33612/diss.95431744>
- Geary, D.C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, 49, 363-383. [https://doi.org/10.1016/0022-0965\(90\)90065-G](https://doi.org/10.1016/0022-0965(90)90065-G)
- Geelhoed, J.W., & Vieijra, J.P.M. (2009). Onderzoek naar leervorderingen. In Th. Kievit, J.A. Tak & J.D. Bosch (Red.), *Handboek: Psychodiagnostiek voor de hulpverlening aan kinderen* (pp. 439-477). De Tijdstroom.
- Gelman, R., & Butterworth, B. (2005). Language and conceptual development series: Number and language: How are they related? *Cognitive Sciences*, 9, 6-10. <https://doi.org/10.1016/j.tics.2004.11.004>

- Ghesquière, P., & Ruijsenaars, A.J.J.M. (2021). Kinderen en jongeren met een leerstoornis. In H. Grietens, J. Vanderfaellie & B. Maes (Red.), *Handboek jeugdhulpverlening: Een orthopedagogisch perspectief op kinderen en jongeren met problemen* (7de editie) (pp. 59-80). Acco.
- Hickendorff, M., & Janssen, J. (2009). De invloed van contexten in rekenopgaven op de prestaties van basisschoolleerlingen. *Reken-wiskundeonderwijs: onderzoek, ontwikkeling, praktijk*, 28, 3-11.  
[https://www.researchgate.net/publication/48345357\\_De\\_invloed\\_van\\_contexten\\_in\\_rekenopgaven\\_op\\_de\\_prestaties\\_van\\_basisschoolleerlingen](https://www.researchgate.net/publication/48345357_De_invloed_van_contexten_in_rekenopgaven_op_de_prestaties_van_basisschoolleerlingen)
- Hollenberg, J., & Veerbeek, G. (2014). Leerling volgen met vaardigheidsscores. *LBBO Beter Begeleiden*, 28-30. <https://cito.nl/media/ogjnmxez/leerling-volgen-met-vaardigheidsscores.pdf>
- Inspectie van het Onderwijs. (2023). *Peil.Taal en rekenen: Einde basisonderwijs en speciaal (basis)onderwijs 2021-2022*. Inspectie van het Onderwijs.  
<https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2023/05/10/peil-taal-en-rekenen-2021-2022/Rapport+Peil+Taal+en+rekenen+2021-2022.pdf>
- Janson, D. (2014). Op weg naar beter reken-wiskundeonderwijs: kansen en bedreigingen. *Reken-wiskundeonderwijs: onderzoek, ontwikkeling, praktijk*, 33, 125-127.  
[https://www.volgens-bartjens.nl/documenten/archief/bartjens/pp\\_33\\_125\\_-\\_127\\_janson\\_op\\_weg\\_naar\\_beter\\_reken-wiskundeonderwijs.pdf](https://www.volgens-bartjens.nl/documenten/archief/bartjens/pp_33_125_-_127_janson_op_weg_naar_beter_reken-wiskundeonderwijs.pdf)
- Janssen, J., Verhelst, N.D., Engelen, R., & Scheltens, F. (2010). *Wetenschappelijke verantwoording van de toetsen LOVS Rekenen-Wiskunde voor groep 3 tot en met 8*. Cito. <https://docplayer.nl/11108466-Wetenschappelijke-verantwoording-van-de-toetsen-lovs-rekenen-wiskunde-voor-groep-3-tot-en-met-8-j-janssen-n-verhelst-r-engelen-en-f.html>
- Kleemans, T., Segers, E., & Verhoeven, L. (2012). Voorspellers van rekenvaardigheid bij kinderen met ESM in groep 3. *VHZ*, 14-18.  
<https://repository.ubn.ru.nl/bitstream/handle/2066/103074/103074.pdf?sequence=1&isAllowed=y>
- Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and Instruction*, 19(6), 513-526. <https://doi.org/10.1016/j.learninstruc.2008.10.002>

- Krom, R., Jongen, I., Verhelst, N.D., Kamphuis, F., & Kleintjes, F. (2010). *Wetenschappelijke verantwoording: DMT en AVI*. Cito. <https://docplayer.nl/9490037-Wetenschappelijke-verantwoording-dmt-en-avi.html>
- Langeveld, M.J. (1934). *Taal en denken*. Wolters.
- LeFevre, J.A., Fast, L., Skwarchuk, S.L., Smith-Chant, B.L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to Mathematics: Longitudinal Predictors of Performance. *Child Development*, 81(6), 1753-1767. <https://doi.org/10.1111/j.1467-8624.2010.01508.x>
- Meelissen, M.R.M., Maassen, N.A.M., Gubbels, J., Van Langen, A.M.L., Valk, J., Dood, C., Derks, I., In 't Zandt, M., & Wolbers, M. (2023). *Resultaten PISA-2022 in vogelvlucht*. Universiteit Twente. <https://doi.org/10.3990/1.9789036559461>
- Naaktgeboren, M., & Prins, S. (2014). *Hulpwaaier dyscalculie: Tips en strategieën bij de hand*. Pica.
- Nix, M. (2022, 19 januari). *Volg je leerlingen en laat ze groeien*. Onderwijs van Morgen. Geraadpleegd op 3 april 2024, van <https://www.onderwijsvanmorgen.nl/volg-je-leerlingen-en-laat-ze-groeien/>
- NOS Nieuws. (2023, 10 mei). Nog steeds zorgen om basisvaardigheden leerlingen, 'scholen kunnen meer doen'. Geraadpleegd op 3 april 2024, van <https://nos.nl/artikel/2474543-nog-steeds-zorgen-om-basisvaardigheden-leerlingen-scholen-kunnen-meer-doen>
- Passolunghi, M.C., & Mammarella, I.C. (2010). Spatial and visual working memory ability in children with difficulties in arithmetic word problem solving. *European Journal of Cognitive Psychology*, 22, 944-963. <https://doi.org/10.1080/09541440903091127>
- Prenger, J. (2011). Werken aan taal via rekensommen. *Tijdschrift voor Remedial Teaching*, 5, 10-13. [https://www.slo.nl/publish/pages/6770/werken-aan-taal-via-rekensommen\\_tvrt2011-5.pdf](https://www.slo.nl/publish/pages/6770/werken-aan-taal-via-rekensommen_tvrt2011-5.pdf)
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Geraadpleegd op 11 december 2023, van <https://www.R-project.org/>
- Ruijsenaars, A.J.J.M., Van Luit, J.E.H., & Van Lieshout, E.C.D.M. (2004). *Rekenproblemen en dyscalculie: Theorie, onderzoek, diagnostiek en behandeling*. Lemniscaat.
- Schmeier, M. (2017). *Effectief rekenonderwijs op de basisschool*. Pica; Pelckmans.

- Simmons, F., Singleton, C., & Horne, J. K. (2008). Phonological awareness and visual-spatial sketchpad functioning predict early arithmetic attainment: Evidence from a longitudinal study. *European Journal of Cognitive Psychology*, 20, 711–722.  
<https://doi.org/10.1080/09541440701614922>
- Sijtstra, J. (2003). *Domeinbeschrijving Lezen*, intern memo. Cito.
- Snijders, T.A.B., & Bosker R.J. (2012). *Multilevel analysis. An introduction to basic and advanced multilevel modeling* (2nd edition). Sage Publications.
- Van Eerde, H.A.A. (2009). Rekenen-wiskunde en taal: een didactisch duo. *Reken-wiskundeonderwijs: onderzoek, ontwikkeling, praktijk*, 28(3), 19-32.  
<https://www.fisme.science.uu.nl/publicaties/literatuur/7214.pdf>
- Van Koeven, E., & Smits, A. (2020). *Rijke taal: Taaldidactiek voor het basisonderwijs*. Boom.
- Van Luit, J.E.H. (2002). Jonge kinderen met rekenproblemen. In A.J.J.M. Ruijsenaars & P. Ghesquière (Red.), *Dyslexie en dyscalculie: ernstige problemen in het leren lezen en rekenen. Recente ontwikkelingen in onderkenning en aanpak* (pp. 113-128). Acco.
- Veltman, A., & Van den Heuvel-Panhuizen, M. (2015). *Rekenen met hele getallen op de basisschool: Tussendoelen Annex Leerlijnen*. Noordhoff.
- Verachtert P., Gadeyne, E., Onghena P., & Ghesquière P. (2007). De rol van psychosociale variabelen in de predictie van rekenvaardigheid bij jonge kinderen. In R.J. Bosker, S. Doolaard & A.E. Jacobse (Red.), *Zorgvuldig en veelbelovend onderwijs: Proceedings van de 34e Onderwijs Research Dagen* (p. 173). RUG/GION.
- Vogelzang, M. (2018, 11 april). *De staat van het onderwijs: Onderwijsverslag 2016/2017*. Inspectie van het Onderwijs. Geraadpleegd op 19 december 2023, van  
[https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs/108126\\_IvhO\\_StaatvanhetOnderwijs\\_TG.pdf](https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2018/04/11/rapport-de-staat-van-het-onderwijs/108126_IvhO_StaatvanhetOnderwijs_TG.pdf)
- Vygotsky, L.S. (1986). *Thought and Language*. MIT Press.
- Wet referentieniveaus Nederlandse taal en rekenen. (2010, 29 april). Geraadpleegd van  
<https://wetten.overheid.nl/BWBR0027679/2022-08-01/0>