

Doelgericht aanbod verlengde leertijd in kleine groepen

Effectmeting kansrijke interventie in het po

Maart 2026

Annemarie van Langen
Isis Derks
Tessa Jenniskens
Maarten Wolbers

De interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' is onderzocht in het kader van een onderzoeksprogramma van het NRO: effectmeting kansrijke interventies. Het onderzoek is onafhankelijk uitgevoerd door A. van Langen, I. Derks, T. Jenniskens & M. Wolbers (KBA Nijmegen).

Dit project is (mede) gefinancierd door het NRO.



Projectnummer: K2022057

© 2026 KBA Nijmegen

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook, en evenmin in een retrieval systeem worden opgeslagen, zonder de voorafgaande schriftelijke toestemming van KBA Nijmegen.

No part of this book/publication may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

Inhoudsopgave

Managementsamenvatting	1
1 Introductie	3
1.1 Achtergrond van het onderzoek	3
1.2 De interventie	3
1.3 Probleemstelling en onderzoeksvragen	4
1.4 RCT-onderzoek en aanvullende analyses	5
1.5 Leeswijzer bij dit rapport	5
2 Literatuurstudie naar interventies verlengde leertijd	6
2.1 Inleiding	6
2.2 Verlengde leertijd ter bestrijding van kansenongelijkheid	6
2.3 Effecten en randvoorwaarden	8
2.4 Het meten van het effect van een VLT-interventie	9
2.5 Effectmeting interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen'	10
3 Opzet en methode	11
3.1 Onderzoeksdesign	11
3.2 Werving interventiescholen	12
3.3 Beschrijving onderzoeksinstrumenten	14
3.4 Randomisatieprocedure	17
3.5 Analyseplan	18
3.6 Poweranalyse	19
4 Beschrijvende resultaten	20
4.1 Inleiding	20
4.2 Beschrijving deelnemers en randomisatie-uitkomst	20
4.3 Beschrijving interventie	23
4.3.1 Implementatiekwaliteit	23
4.3.2 Pedagogisch-didactische kwaliteit	25
4.3.3 Betrokkenheid ouders en reguliere groepsleerkracht	27
5 Effect van de interventie en modererende condities	30
5.1 Bepalen leerwinst	30
5.1.1 Inleiding	30
5.1.2 Bivariate analyse	30
5.1.3 Multivariate analyse	33
5.2 Invloed van implementatiekwaliteit en pedagogisch-didactische kwaliteit	36
6 Conclusie en discussie	38
6.1 Uitgevoerd onderzoek	38
6.2 Beantwoording van de onderzoeksvragen	39
6.3 Reflectie en aanbevelingen	40

Literatuur		44
Bijlage 1	Ontwikkelde onderzoeksinstrumenten	46
Bijlage 2	Tabellen multivariate analyse	54

Managementsamenvatting

In 2022 heeft de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) een oproep gedaan tot effectonderzoek naar kansrijke interventies in het onderwijs. KBA Nijmegen heeft een voorstel ingediend voor het onderzoeken van de effectiviteit van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen'. Deze interventie is erop gericht achterstanden van leerlingen in groep 4 tot en met 8 weg te werken bij taal, technisch en begrijpend lezen, en rekenen-wiskunde. De deelnemende leerlingen krijgen twee keer per week les in groepjes van maximaal acht leerlingen gedurende een periode van tien weken. De lessen vinden onder of na schooltijd plaats buiten het reguliere klaslokaal. De leerkrachten die de interventie aanbieden zijn werkzaam op de school van de deelnemers ('interne coach') of worden ingehuurd ('externe coach'). De coaches dienen voor de inzet te worden geschoold door de ontwikkelaars van de interventie (Beuker Onderwijs) in het gebruik van de lesprogramma's, materialen, werkvormen en didactiek.

Het onderzoek dat in dit rapport wordt beschreven, is opgezet als een *randomized controlled trial* (RCT). Op de deelnemende scholen werden leerlingen van groep 6 die in aanmerking kwamen voor de interventie willekeurig toegewezen aan de interventie- of controlegroep. De effectmeting werd uitgevoerd aan de hand van een voormeting (eind groep 5) en twee nametingen (medio en eind groep 6) van de prestaties van de leerlingen bij begrijpend lezen en rekenen-wiskunde. De data werden grotendeels verzameld via zelf ontwikkelde instrumenten: twee leerlinginformatie-formulieren, een digitaal logboek over de geïmplementeerde activiteiten tijdens elke interventieles, een lesobservatieschema voor de observatie van interventielessen, en vragenlijsten voor ouders en reguliere groepsleerkrachten. Daarnaast is gebruikgemaakt van bestaande instrumenten, namelijk de Cito-LVS toetsen Begrijpend lezen en Rekenen-Wiskunde, en de Niet-Schoolse Cognitieve Capaciteitentest (NSCCT).

Een belangrijke kanttekening bij het uitgevoerde onderzoek is dat de interventieontwikkelaar kort voor de start van de interventie op de scholen besloot de samenwerking op te zeggen met de organisatie die de externe coaches zou leveren. Deze coaches bleken namelijk onvoldoende gekwalificeerd voor de begeleiding van basisschoolleerlingen. Het gevolg van deze situatie was dat van de negentien aangemelde interventiescholen er elf niet langer konden/wilden deelnemen aan de interventie en het interventieonderzoek, omdat er geen coach beschikbaar was om de interventielessen te verzorgen. Met het aantal deelnemende scholen nam ook het aantal deelnemende leerlingen af, wat negatieve gevolgen had voor de statistische power van de analyses. Om deze reden moeten de onderzoeksresultaten met de nodige voorzichtigheid worden geïnterpreteerd. Ofschoon het onderzoek is opgezet als een RCT-studie om causale effecten van de interventie te kunnen vaststellen, kunnen – gezien de (mogelijk selectieve) uitval van scholen en, als gevolg daarvan, eventuele heterogeniteitsproblemen bij de implementatie van de interventie en beperkte grootte van de onderzoeksgroep – geen harde conclusies getrokken worden over de werking van de interventie.

De uiteindelijke steekproef bestond uit acht scholen waar in totaal negentien interventiegroepjes zijn ingericht die tussen begin november 2023 en eind januari 2024 deelnamen aan de interventie, gericht op begrijpend lezen of rekenen-wiskunde. Uit onze analyses blijkt dat de deelnemers aan de rekeninterventie zowel op de korte als de middellange termijn significant sterker zijn gegroeid op de vaardigheid rekenen-wiskunde dan leerlingen in de controlegroep. De leesvaardigheid van deze leerlingen is niet sterker toegenomen dan die van de controlegroep. Voor zover de reken-groei is toe te schrijven aan de rekeninterventie is het effect van deze interventie dus

domeinspecifiek. Een vergelijkbare relatie vinden we niet voor de deelnemers aan de extra lessen begrijpend lezen in het kader van de interventie; hun leerwinst bij lezen en rekenen is niet significant groter dan van leerlingen in de controleconditie.

Daarnaast hebben we geconstateerd dat de kwaliteit van de pedagogisch-didactische interactie tussen de coach en de leerlingen tijdens de interventielessen van invloed lijken te zijn op de groei die interventiedeelnemers doormaken bij begrijpend lezen en rekenen-wiskunde, zowel op de korte als de middellange termijn (d.w.z. na enkele maanden). Een dergelijke relatie vinden we niet voor de implementatiekwaliteit van de interventie. De invloed van de betrokkenheid van ouders en reguliere groepsleerkrachten op de effectiviteit van de interventie kon niet worden vastgesteld, omdat de respons op de betreffende vragenlijsten daarvoor te laag was.

Bij vervolgonderzoek naar de effecten van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' is een belangrijk aandachtspunt dat er voldoende goed gekwalificeerde interne of externe coaches beschikbaar moeten zijn om de interventielessen te verzorgen, zodat het onderzoek meer (betrouwbare) data oplevert. Ook raden we aan de randomisatieprocedure in de toekomst niet op leerlingniveau, maar op klas- of schoolniveau uit te voeren.

Ondanks deze aandachtspunten en gegeven het voorbehoud over het gebrek aan statistische power waardoor we slechts indicatieve uitspraken over de effectiviteit van de interventie kunnen doen, lijken onze bevindingen te suggereren dat als scholen de rekeninterventie in het kader van 'Doelgericht aanbod verlengde leertijd in kleine groepen' aanbieden aan leerlingen in groep 6 die daarvoor in aanmerking komen, dit gunstig uitpakt voor het verkleinen van de leerachterstand van deze leerlingen op het gebied van rekenen-wiskunde.

1 Introductie

1.1 Achtergrond van het onderzoek

De Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) zet zich in voor een sterk wetenschapsstelsel in Nederland door kwaliteit en vernieuwing in de wetenschap te bevorderen. In 2022 heeft NWO aan de hand van een prijsvraag veertien veelbelovende interventies geselecteerd, waarvan verwacht wordt dat ze bijdragen aan het terugdringen van cognitieve en/of sociaal-emotionele vertragingen in het funderend onderwijs (po en vo). Onderzoekers werden daarna uitgenodigd een voorstel in te dienen om effectmetingen van deze interventies uit te voeren.

Eén van de geselecteerde interventies is 'Doelgericht aanbod verlengde leertijd in kleine groepen', gericht op het verkleinen van leerachterstanden bij taal, lezen en rekenen van leerlingen in het basisonderwijs. De betreffende interventie is ontwikkeld door Beuker Onderwijs en in het schooljaar 2023-2024 geïmplementeerd met financiering van NWO. KBA Nijmegen is aangewezen om het effect van deze interventie te onderzoeken. In dit rapport wordt verslag gedaan van dat onderzoek, dat gekenmerkt wordt door een RCT-design (d.w.z. met een random toegewezen interventie- en controlegroep). Naast dit onderzoek zijn aanvullende analyses met extra controlegroepen uitgevoerd waarvan verslag wordt gedaan in een addendum bij dit rapport.

Onze dank gaat uit naar het NWO dat ons de mogelijkheid bood dit onderzoek uit te voeren. Ook willen we graag de medewerkers van Beuker Onderwijs en de betrokken scholen, teamleden, leerlingen en ouders bedanken voor hun medewerking aan het onderzoek.

1.2 De interventie

Bij de start van het implementatie- en onderzoekstraject heeft Beuker Onderwijs de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' uitgebreid beschreven en toegelicht. In deze paragraaf geven we een beknopte samenvatting daarvan.¹ De interventie is opgezet om gericht achterstanden van leerlingen weg te werken bij taal, technisch en begrijpend lezen en rekenen-wiskunde. Per domein en leerjaar bestaat de interventie uit lesprogramma's van twintig lessen, uitgewerkt in een leerkrachthandleiding, een werkboek voor de leerlingen en een aantal spelmaterialen. De twintig lessen zijn opgedeeld in tien A-lessen (aanbod nieuwe lesstof) en tien B-lessen (herhaling en oefening). De deelnemende leerlingen krijgen twee keer per week les in kleine groepjes (maximaal acht leerlingen) gedurende een periode van tien weken. De lessen vinden onder of na schooltijd plaats buiten het reguliere klaslokaal. De leerkrachten die de interventie aanbieden zijn werkzaam op de school van de deelnemers ('interne coach') of worden ingehuurd ('externe coach'). Deze coaches dienen voor de inzet eerst te worden geschoold (door de medewerkers van Beuker Onderwijs) in het gebruik van de lesprogramma's, materialen, werkvormen en didactiek. Volgens de interventieontwikkelaar zijn de lessen zodanig uitgeschreven dat de coaches zich daarna nog nauwelijks hoeven voor te bereiden en veel mogelijkheden krijgen aangereikt om te differentiëren. De lessen sluiten goed aan bij het reguliere onderwijsaanbod, maar hebben een andere insteek met veel aandacht voor directe instructie, coöperatieve

¹ Zie ook <https://beukeronderwijs.nl/>

werkvormen, spellen en positieve feedback, waardoor de leerlingen uitgedaagd worden tot actieve deelname. Ook bevat het lesmateriaal veel verrijksaanbod en keuzeonderdelen.

De doelgroep van de interventie bestaat uit leerlingen in groep 4 tot en met 8 van het basisonderwijs met een leerachterstand in taal, lezen of rekenen ten opzichte van andere leerlingen in dezelfde jaargroep. Het beoogde cognitieve effect bestaat eruit dat deze leerlingen hun achterstand inlopen door harder vooruit te gaan op de basisvaardigheid die in de verlengde leertijd is aangeboden dan (vergelijkbare) andere leerlingen die niet deelnemen aan de interventie.

1.3 Probleemstelling en onderzoeksvragen

Tijdens de coronapandemie en de daarmee gepaard gaande schoolgebouwsluitingen en lesuitval hebben leerlingen in het basis- en voortgezet onderwijs leerachterstanden opgelopen; in het bijzonder degenen die tot de doelgroepen van onderwijsachterstandenbeleid behoren (Inspectie van het Onderwijs, 2022). In het kader van het Nationaal Programma Onderwijs (NP Onderwijs) ontvingen scholen in de jaren na corona substantiële extra middelen ('NPO-middelen') om maatregelen te kunnen nemen om de opgelopen achterstanden in te halen. Deze middelen konden zij tot en met schooljaar 2024-2025 inzetten.

De veertien veelbelovende interventies die NWO heeft geselecteerd en waarvan 'Doelgericht aanbod verlengde leertijd in kleine groepen' er één is, zijn specifiek uitgekozen op grond van de verwachting dat zij cognitieve en/of sociaal-emotionele leervertragingen in de Nederlandse onderwijscontext, mede ontstaan ten gevolge van de coronacrisis, kunnen helpen terugdringen. Door het financieren van de implementatie van deze interventies in schooljaar 2024-2025, alsmede de effectmetingen naar de interventies, wil NWO duurzame kennis ontwikkelen over de interventie-effectiviteit.

De interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' heeft volgens de ontwikkelaar positieve effecten op de leerprestaties omdat het voldoet aan de voorwaarden van effectieve naschoolse programma's (zie hoofdstuk 2 voor een literatuuroverzicht). Of deze effecten daadwerkelijk optreden hangt echter mede af van de mate waarin de interventie op de deelnemende scholen wordt uitgevoerd zoals bedoeld ('*treatment fidelity*' of implementatiekwaliteit). Twee condities die volgens de literatuur (zie opnieuw hoofdstuk 2) ook van invloed zijn op het succes van verlengde leertijd-interventies zijn de mate waarin de pedagogische-didactische vaardigheden van de leerkrachten aansluiten op het reguliere onderwijs en de leerdoelen en onderwijsbehoeften van de leerlingen, en de mate waarin de reguliere groepsleerkrachten en ouders bij de interventie worden betrokken.

In het onderzoek waarvan in dit rapport verslag wordt gedaan, is onderzocht in hoeverre de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' heeft geleid tot het verkleinen van de leerachterstanden bij begrijpend lezen en rekenen-wiskunde van leerlingen in groep 6 van het basisonderwijs. Met het onderzoek hopen we bij te dragen aan de wetenschappelijk evidentie voor de effectiviteit van de betreffende interventie, inclusief de condities die de effectiviteit modereren.

De onderzoeksvragen luiden als volgt:

1. *Wat is het effect van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' op de leerachterstand van de deelnemende leerlingen in relatie tot andere leerlingen in dezelfde jaargroep?*
2. *In hoeverre zijn de effecten van de interventie afhankelijk van de implementatiekwaliteit, c.q. de mate waarin het programma wordt uitgevoerd zoals bedoeld?*
3. *In hoeverre zijn de effecten van de interventie afhankelijk van de pedagogisch-didactische kwaliteit van het interactieproces tussen de (interventie)leerkrachten en de leerlingen?*
4. *In hoeverre zijn de effecten van de interventie afhankelijk van de mate waarin de reguliere groepsleerkrachten en ouders bij de interventie worden betrokken?*

1.4 RCT-onderzoek en aanvullende analyses

Het onderzoek dat is uitgevoerd naar de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' bestaat uit een kwantitatieve effectevaluatie in de vorm van een *Randomized Controlled Trial* (RCT) waarin het effect van de interventie op de relatieve leerachterstand van de deelnemers is onderzocht en een kwalitatieve procesevaluatie waarin is onderzocht hoe de implementatiekwaliteit, de pedagogisch-didactische interactie en de betrokkenheid van de groepsleerkrachten en ouders van invloed zijn op de effectiviteit van de interventie. Het RCT-design in het onderzoek geeft (in theorie) de beste garantie dat goed gefundeerde kennis wordt verkregen over wat werkt, voor wie, in welke omstandigheden en waarom; kennis die van belang is voor onderwijsprofessionals die de juiste keuze willen maken om met beperkte middelen hun leerlingen zo goed mogelijk te ondersteunen. In het onderhavige rapport is het RCT-onderzoek, inclusief alle haken en ogen die daaraan zitten in de praktijk van onderwijsonderzoek, beschreven.

Om het inzicht in het effect van de interventies verder te kunnen vergroten zijn ook *aanvullende analyses* uitgevoerd waarbij extra controlegroepen van scholen en leerlingen betrokken zijn. Deze aanvullende analyses zijn beschreven in een addendum bij dit rapport.

1.5 Leeswijzer bij dit rapport

In hoofdstuk 2 geven we een overzicht van de wetenschappelijke literatuur over interventies met verlengde leertijd. In hoofdstuk 3 volgt een verantwoording van de opzet en het verloop van het onderzoek. Hoofdstuk 4 bevat de beschrijvende resultaten. Hoofdstuk 5 gaat in op het effect van de interventie en de condities die dit effect modereren. In hoofdstuk 6 ten slotte wordt het onderzoek samengevat en voorzien van enige conclusies.

2 Literatuurstudie naar interventies verlengde leertijd

2.1 Inleiding

In dit hoofdstuk geven we een beknopt overzicht van de huidige wetenschappelijke kennis over (de effectiviteit van) interventies rondom verlengde leertijd. De literatuurstudie die aan de basis ligt van dit hoofdstuk is uitgevoerd in het voorjaar van 2023 en geactualiseerd in het voorjaar van 2025. Het schema hierna geeft een overzicht van de geraadpleegde databanken en gehanteerde zoektermen (figuur 2.1). De geraadpleegde literatuur dateert van de afgelopen vijftien tot twintig jaar.

Figuur 2.1 – Overzicht databanken en zoektermen

Databanken & zoekmachines	Zoektermen Nederlands	Zoektermen Engels
NRO Google Scholar ERIC	Verlengde leertijd Verlengde schooldag Uitbreiding onderwijstijd Leertijdverlenging Extra lestijd Naschoolse programma's Verlengde schooltijdprogramma's Buitenschoolse programma's <u><i>Al dan niet in combinatie met:</i></u> Basisonderwijs Leerachterstanden Onderwijsachterstand Verbetering leerprestaties Prestatiekloof	Extended learning time Lengthened school day Increased instructional time Time-on-task extension Additional instruction time After-school programs Extended day programs Out-of-school time programs (OST) Primary schools/education Learning loss/Educational gap Educational disadvantage Improving academic achievement Achievement gap

2.2 Verlengde leertijd ter bestrijding van kansenongelijkheid

Introductie

Basisscholen kiezen er soms voor bepaalde leerlingen extra uren onderwijs aan te bieden, boven op de reguliere onderwijsuren die zij samen met de rest van de klas krijgen. Deze ‘verlengde leertijd’ (VLT) is meestal bedoeld om leerlingen met een leerachterstand (een achterstand in het beheersen van de leerstof ten opzichte van leeftijdsgenoten in dezelfde onderwijsfase) te ondersteunen en ervoor te zorgen dat hun achterstand kleiner wordt of geheel verdwijnt. Leerachterstanden kunnen bijvoorbeeld ontstaan door langdurige afwezigheid door ziekte of een leer- en ontwikkelingsstoornis, maar de oorzaak kan ook liggen in het feit dat niet elk kind vanuit thuis evenveel cognitieve en culturele bagage meekrijgt (CBS, 2024). Leerlingen met een minder kansrijke gezinsachtergrond profiteren daardoor vaak minder van het onderwijs dan hun leeftijdsgenoten, waardoor ook een leerachterstand kan ontstaan. Het aanbieden van VLT aan deze leerlingen kan een middel zijn om deze kansenongelijkheid tegen te gaan. Ook de coronapandemie, inclusief de schoolgebouwen sluiting en lesuitval waarmee deze gepaard ging, heeft ervoor gezorgd dat veel leerlingen in het primair en voorgezet onderwijs vertraging in de cognitieve ontwikkeling – vaak ook in de sociaal-emotionele ontwikkeling – hebben opgelopen. Aan het einde van

het schooljaar 2020-2021 vertoonden basisschoolleerlingen volgens het Nationaal Programma Onderwijs (2021) een gemiddelde leervertraging van ongeveer 10 weken in rekenen-wiskunde, zeven weken in begrijpend lezen en zes weken in spelling. De negatieve effecten van de coronacrisis waren zichtbaar bij alle leerlingen, maar golden in versterkte mate voor leerlingen uit gezinnen met een lage sociaaleconomische status (SES) (Inspectie van het Onderwijs, 2022). Volgens het NCO (2021) was de vertraging in de leergroei anderhalf keer zo groot bij leerlingen met een lage SES als bij leerlingen met een hoge SES. Ook hier kan dus gesproken worden van leerachterstanden ten gevolge van kansenongelijkheid die eventueel met VLT zouden kunnen worden aangepakt. In dit literatuuroverzicht ligt de focus voornamelijk op publicaties over VLT-interventies ter bestrijding van leerachterstanden die gerelateerd zijn aan kansenongelijkheid.

Soorten verlengde leertijd (VLT)

Verlengde leertijd kan allerlei vormen aannemen, zoals zomerscholen, weekendscholen, kopklassen en schakelklassen (Beekhoven, 2017). Ook binnen deze categorieën zijn er nog vele varianten mogelijk. De programmaduur van een *zomerschool* kan bijvoorbeeld verschillen van een paar dagen tot enkele weken en ook het aanbod van een *weekendschool* kan in lengte variëren van enkele maanden tot enkele jaren. *Kopklassen* zijn bedoeld voor leerlingen in de fase tussen groep 8 en het voortgezet onderwijs, om hen eventueel naar een hoger niveau in het voortgezet onderwijs te laten doorstromen (Kuiken, 2016). Bij VLT in de vorm van een *schakelklas* – bijvoorbeeld gedurende één heel schooljaar – krijgen leerlingen voltijds of deeltijds onderwijs aangeboden in een aparte klas tijdens reguliere schooltijden en/of in extra lesuren na school. De meest voorkomende vorm van VLT is de deeltijdvariant van de schakelklas. Leerlingen die hieraan deelnemen zitten gemiddeld acht uur per week apart van hun ‘reguliere’ klasgenoten (Beekhoven, 2017).

VLT verwijst dus naar een breed scala aan onderwijsvormen waarmee extra uren onderwijs worden aangeboden aan bepaalde leerlingen. Ook de domeinen die in de VLT aandacht krijgen variëren. Zomerscholen leggen in het Nederlands basisonderwijs vaak de focus op taal- en rekenachterstanden en sociaal-emotionele vaardigheden (Kortekaas-Rijlaarsdam, 2020). Volgens Appelhof (2009) ligt in de praktijk bij de meeste VLT-interventies de focus op taal.

Deelnemers

Doorgaans geldt dat deelname aan VLT op vrijwillige basis is. In het basisonderwijs zijn het meestal de leerkrachten, intern begeleiders en/of ouders die een leerling aanmelden voor deelname aan VLT. Vaak worden de deelnemers uitgekozen op basis van verschillende kenmerken, zoals de ernst/omvang van de taal- en/of rekenachterstand, de motivatie van de leerling en/of van de ouders om deel te nemen en kenmerken van de gezinsachtergrond (Beekhoven, 2017). Vaak speelt ook het (onbenutte) leerpotentieel van de leerling een rol; met name de vraag of de leerkrachten of intern begeleiders van mening zijn dat er ‘meer inzit dan er nu uitkomt’. Dit geldt bijvoorbeeld bij de selectie van deelnemers aan kopklassen, waarin leerlingen immers ondersteund worden om eventueel door te stromen naar een hoger niveau in het voortgezet onderwijs dan zonder kopklas mogelijk zou zijn. Ook voor de verlengde schooldagvariant van de schakelklas wordt leerpotentieel – naast leerachterstand – vaak benoemd als selectie criterium (Beekhoven, 2017).

2.3 Effecten en randvoorwaarden

Volgens de geraadpleegde literatuur kunnen systematische effecten van VLT-interventies op leerlingprestaties lang niet altijd onomstotelijk worden vastgesteld (Kennisrotonde, 2019; Meyer & Klaveren, 2013; Kidron & Lindsay, 2014). Het Education Lab Netherlands (2021) stelt dat VLT-interventies behoorlijke leerwinsten kunnen opleveren, maar ook ineffectief kunnen zijn. Kidron en Lindsay (2014) concluderen dat er onvoldoende bewijs is te suggereren dat VLT een positief effect heeft op de leerprestaties van alle leerlingen in alle settings, ongeacht hun onderwijsbehoeften, leerjaar of sociaaleconomische achtergrond. Daarentegen vinden zij wel een positief effect bij leerlingen die een achterstand hebben ten opzichte van het gewenste prestatieniveau. Ook het Centraal Planbureau (CPB, 2018) stelt dat VLT specifiek voor leerlingen met een leerachterstand positieve leerresultaten oplevert. Kortekaas-Rijlaarsdam et al. (2020) benadrukken dat er in Nederland over het algemeen weinig effect kan worden aangetoond van het verlengen van schooldagen. De Engelse *Education Endowment Foundation* (2018) daarentegen stelt dat “leerlingen gemiddeld twee extra maanden per jaar vooruitgang boeken door verlengde schooltijd, met name door gericht gebruik te maken van voor- en naschoolse programma’s” (p. 14). Wedel (2021) concludeert op basis van internationaal vergelijkende data uit het TIMSS-onderzoek (*Trends in International Mathematics and Science Study*) dat een extra uur instructietijd gemiddeld genomen geassocieerd is met een stijging van 0,03 standaarddeviatie in de toetsscores van leerlingen, over alle landen heen.

De tegenstrijdige onderzoeksbevindingen over de effecten van VLT hangen samen met de grote variëteit in het VLT-aanbod. Het is daarom raadzaam bij het uitvoeren van effectmetingen ook te inventariseren hoe en welke kenmerken en condities van de VLT-programma’s een rol spelen bij de behaalde effecten (Kidron & Lindsay, 2014). Zomerscholen in het voortgezet onderwijs behalen bijvoorbeeld het beste resultaat als ze aaneengesloten tien dagen of langer duren (Haerlemans et al., 2018). Ook relevant is welke leerlingen worden geselecteerd voor de VLT. Zo blijken zomerscholen vooral effectief voor leerlingen uit gezinnen met een lage SES, die tijdens de vakanties thuis vaak minder gestimuleerd worden om educatieve activiteiten uit te voeren dan andere leerlingen (Heyden, 2021). Onderpresteerders profiteren sterker dan andere leerlingen van verlengde leertijd, omdat ze immers meer mogelijkheden hebben om te kunnen groeien (Mulder, 2008, 2009). Een andere bepalende factor is de intensiteit van de VLT. Er is een positief verband gevonden tussen het aantal uren VLT per week en de verbetering van de leerprestaties (Kennisrotonde, 2019), al geldt dat uiteraard maar tot op zekere hoogte (plafondeffect) en is onduidelijk welk aantal uren per dag of week het meest effectief is.

Kuiken (2013) benadrukt dat niet zozeer de hoeveelheid extra onderwijstijd het verschil maakt (*allocated time* of *administrative time*) als wel de tijd die leerlingen daadwerkelijk effectief aan schoolwerk besteden (*time on task*). Ook Wedel (2021) stelt dat naast de kwantiteit ook de kwaliteit van de instructie belangrijk is; haar analyses op de TIMSS-data laten zien dat een extra uur instructietijd aangeboden door een hooggekwalficeerde leraar samenhangt met een stijging van de gemiddelde toetsscore met 0,04 tot 0,05 standaarddeviatie. Het belang van gekwalificeerde leraren, die getraind zijn in pedagogisch-didactisch handelen, wordt in vrijwel alle bestaande onderzoeken naar het effect van VLT benoemd (Appelhof, 2009; Van de Burgwal, 2010; Kaplan & Chan, 2011; Kidron & Lindsay, 2014; Kuiken, 2016). Deze factor is dermate bepalend dat wanneer VLT-leraren onvoldoende gekwalificeerd zijn er geen positief effect van de VLT op de leerprestaties van leerlingen aanwezig is (Kidron & Lindsay, 2014; Kortekaas-Rijlaarsdam et al; 2020).

Een andere terugkerend thema in de geraadpleegde literatuur is dat de effecten van VLT-interventies het sterkst zijn als het aanbod specifiek is afgestemd op de onderwijsbehoefte van de leerling (Kortekaas-Rijlaarsdam et al., 2020). Reguliere leerkrachten moeten bij de VLT-interventie worden betrokken, ook als zij deze niet zelf aanbieden. Meerdere onderzoekers benadrukken het belang van de verbinding tussen de VLT-interventie en het reguliere onderwijs en ook van de aansluiting van de VLT-interventie op het bestaande curriculum van taal en/of rekenen (Appelhof, 2009; Driessen, Claassen & Smit, 2010). Het betrekken van ouders heeft eveneens een positief effect op de effectiviteit van de VLT-interventie (Van der Vegt, 2021). Ook Kim & Quinn (2013) constateerden dat het betrekken van ouders bij zomerscholen het succes van deze VLT vergrootte. Ouderbetrokkenheid kan verschillende vormen aannemen, zoals ouderparticipatie, contact tussen ouder en school en onderwijsondersteunend gedrag thuis (Heyden, 2021). In relatie tot VLT kan ouderbetrokkenheid zich manifesteren als ervoor zorgen dat het kind aanwezig is, belangstelling tonen en praten over de VLT, thuis oefenen met de leerstof, samen lezen of educatieve activiteiten ondernemen.

Ook de groepsgrootte bij de VLT is van invloed op de leerprestaties van leerlingen. Aangetoond is dat één-op-één begeleiding zeer effectief is voor het verbeteren van met name taal- en leesprestaties, mits gegeven door een gekwalificeerde begeleider (Kortekaas-Rijlaarsdam et al., 2020). Organisatorisch en financieel gezien is een individueel VLT-aanbod echter niet altijd haalbaar, en de meeste literatuur richt zich dan ook op VLT voor (kleine) groepen leerlingen. Het voordeel daarvan is dat leerlingen van elkaar kunnen leren, tegelijkertijd bestaat ook het risico dat minder goed kan worden aangesloten op individuele behoeftes (Kortekaas-Rijlaarsdam et al., 2020). Volgens Slavin et al., (2011) zijn de effecten van *remedial teaching* in kleine groepjes (twee tot vier leerlingen) slechts iets kleiner dan de effecten van één-op-één begeleiding. Ook uit een onderzoek van Cook et al. (2015) blijkt dat de effectiviteit van VLT het grootste is wanneer er in kleine groepjes van maximaal vier tot zes leerlingen wordt gewerkt.

2.4 Het meten van het effect van een VLT-interventie

Dat in veel onderzoek geen systematische effecten van VLT-interventies op leerprestaties kunnen worden vastgesteld, kan ook te wijten zijn aan gebrekkige onderzoeksmethodes. Patall, Cooper en Allen (2010) tonen met een reviewstudie aan dat de onderzoekdesigns voor het meten van de effecten van VLT-interventies over het algemeen zwak zijn. Smit, Driessen en Van Kuijk (2015) bevestigen dat het onderzoek naar naschoolse programma's "vaak niet voldoet aan gangbare methodologische eisen" (p. 41) waardoor er moeilijk uitspraken over de causale invloed van VLT-interventies op leerprestaties gedaan kunnen worden. De meeste VLT-programma's werken, bijvoorbeeld, met een vrijwillige deelname van de leerlingen of hun ouders wat kan zorgen voor een selectiebias; de deelnemers zijn vaak meer gemotiveerd dan een controlegroep van niet-deelnemers (Jez & Wassmer, 2015). Behalve niet gerandomiseerd zijn de interventie- en controlegroepen in veel onderzoeken ook van beperkte omvang (Smit, Driessen, & Kuijk, 2015). Daarnaast zijn de onderlinge verschillen tussen VLT-interventies groot en is het moeilijk een eenduidig effect aan te tonen van de VLT (Kortekaas-Rijlaarsdam et al, 2020). In 2010 concludeerde Kasenberg dat er geen duidelijke onderzoeksmatige onderbouwing bestond voor de effecten van VLT. Op enkele studies na is er een klein decennium later volgens de Kennisrotonde (2019) nog niet veel veranderd, al is er wel een veldexperiment uitgevoerd door Meyer en Van Klaveren (2013). Hierin werd aan willekeurige ouders van basisschoolleerlingen de mogelijkheid geboden hun kind op te geven voor een VLT-interventie van enkele maanden. Vervolgens werden deze

leerlingen vergeleken met leerlingen die niet hadden deelgenomen, afkomstig uit dezelfde kleine stad in Nederland. In dit onderzoek zijn geen significante effecten gevonden van de VLT-interventie op de lees- en rekenvaardigheden van de deelnemende leerlingen.

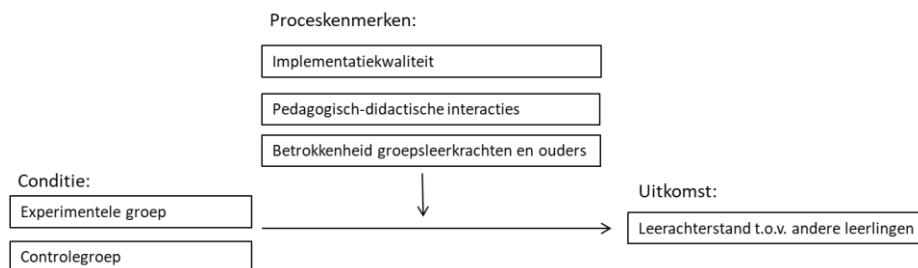
2.5 Effectmeting interventie ‘Doelgericht aanbod verlengde leertijd in kleine groepen’

Volgens Beuker Onderwijs, de ontwikkelaar van ‘Doelgericht aanbod verlengde leertijd in kleine groepen’, voldoet deze interventie aan de voorwaarden van een effectief VLT-programma; de interventie wordt aangeboden door een gekwalificeerde leerkracht (de interne of externe coach), in kleine groepen (maximaal 8 leerlingen) gedurende langere tijd (tweemaal per week gedurende een periode van 10 weken). Bovendien sluit het lesmateriaal aan op (het curriculum van) het reguliere onderwijs en wordt ouderbetrokkenheid gestimuleerd. Daarom is het volgens Beuker Onderwijs aannemelijk dat interventiedeelnemers een positief effect zal hebben op de leerprestaties van de deelnemers. Aangezien de interventie is geselecteerd was de NWO-jury het daar kennelijk mee eens. Of de aanname klopt dat het om een kansrijke interventie gaat hebben we onderzocht aan de hand van een onderzoek, dat bestaat uit een gecontroleerd experiment met aselechte toewijzing van proefpersonen aan de interventie- of controleconditie (RCT-design). In het volgende hoofdstuk lichten we de opzet en het verloop van het onderzoek verder toe en doen vervolgens verslag van de resultaten.

3 Opzet en methode

3.1 Onderzoeksdesign

Het onderzoek bestaat uit een **effectevaluatie** in de vorm van een *Randomized Controlled Trial* (RCT; gekenmerkt door een random toegewezen experimentele groep [of interventiegroep] en controlegroep), waarin het effect van de interventie op de relatieve leerachterstand van de deelnemers is onderzocht (onderzoeksvraag 1) en een **procesevaluatie** waarin is onderzocht hoe de implementatiekwaliteit en pedagogisch-didactische interacties tijdens de interventie, alsmede de betrokkenheid van de groepsleerkrachten en ouders, van invloed zijn op de effectiviteit van de interventie (onderzoeksvragen 2, 3 en 4). Figuur 3.1 vat het onderliggende causale onderzoeksmodel samen.



Figuur 3.1 – Causaal onderzoeksmodel voor de effectiviteit van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen'

Om de RCT zo zuiver mogelijk in te steken, is de interventie en het daaraan gekoppelde onderzoek beperkt tot leerlingen in één specifiek leerjaar. Daarmee voorkomen we dat de onderzoeksuitkomsten worden beïnvloed door de onderwijsfase van de leerlingen, aangezien deze variabele geen deel uitmaakt van het onderzoeksmodel. Gekozen is voor groep 6, onder meer omdat de interventieontwikkelaar meerdere lesprogramma's voor deze groep heeft ontwikkeld en vrijwel alle basisscholen toetsen (bijvoorbeeld van Cito LVS) gebruiken voor het meten van de leerprestaties van leerlingen direct vóór, in en na groep 6 (E5, M6, E6). De meeste leerlingen die tijdens de interventieperiode in groep 6 zouden zitten, bevonden zich bovendien tijdens de coronacrisis in leerjaar 2, 3 en 4; onze aanname was dat zij in die fase een relatief groot risico hebben gelopen op het ontwikkelen van leerachterstand.²

Een andere inperking is dat we uitsluitend de basisvaardigheden begrijpend lezen en rekenen onderzoeken, ofschoon de interventieontwikkelaar ook lesprogramma's voor taal in groep 6 aanbiedt. Ook de variatie in leerachterstand naar domein (taal, lezen, rekenen) is echter geen onderdeel van het RCT-onderzoeksmodel. Gekozen is voor begrijpend lezen en rekenen-wiskunde omdat juist bij deze domeinen al geruime tijd zorgen zijn over de leerachterstanden (mede ten

² Uit later onderzoek naar de effecten van de coronacrisis is gebleken dat leerlingen in deze leerjaren inderdaad relatief vaak leerachterstand hebben opgelopen (Haelermans et al., 2023).

gevolge van de coronacrisis) van leerlingen uit achterstandsgroepen (zie o.a. Inspectie van het Onderwijs, 2022). Daar komt bij dat Boekeloo in 2014 vaststelde dat het effect van de interventie op taal moeilijk te meten is, omdat het interventiemateriaal niet alleen aandacht besteedt aan spelling, terwijl de meest gebruikte LVS-taaltoetsen vooral inzicht geven in de prestaties voor spelling.

Om een eventueel effect van de interventie bij leerlingen in de interventieconditie vast te kunnen stellen en het effect met zekerheid toe te kunnen schrijven aan de deelname aan de interventie, dienen de leerlingen in de controleconditie het reguliere lesprogramma in groep 6 volgen. Dit betekent dat de leerlingen in interventie- en controleconditie in principe hetzelfde reguliere onderwijsprogramma volgen in groep 6 (*'care as usual'*). De leerlingen in de interventieconditie ontvangen als aanvulling daarop extra ondersteuning, in de vorm van doelgericht aanbod voor taal of rekenen, onderwezen door een coach in kleine groepjes. De leerlingen in de controleconditie ontvangen géén extra ondersteuning buiten het reguliere lesprogramma.³

3.2 Werving interventiescholen

Voor de werving van deelnemers aan de interventie en het onderzoek is gebruikgemaakt van de contacten die de interventieontwikkelaar (Beuker Onderwijs) al onderhield met twee schoolbesturen in Amsterdam. Gezamenlijk vormen deze twee schoolbesturen het bevoegd gezag over 33 scholen, die allemaal gelegen zijn in Amsterdam en waarvan een deel al gebruikmaakt van de diensten van Beuker Onderwijs. Al deze scholen zijn benaderd voor deelname aan het interventieonderzoek. KBA Nijmegen heeft samen met Beuker Onderwijs in de eerste helft van 2023 een aantal informatiebijeenkomsten voor de scholen en besturen georganiseerd, waarin zowel de interventie als het onderzoek zijn toegelicht. Daarnaast hebben de scholen ook schriftelijk informatiemateriaal ontvangen – zoals een stappenplan voor het onderzoek en informatiebrieven voor ouders – waarna ze telefonisch en per e-mail zijn benaderd met het verzoek zich aan te melden als deelnemer door een formulier in te vullen. Met dit aanmeldformulier verstrekten de scholen informatie over het gebruikte leerlingvolgsysteem, het aantal leerlingen in groep 6 in schooljaar 2023-2024 en de eventuele beschikbaarheid van een leerkracht uit het team voor het verzorgen van de interventielessen.

Vóór de zomervakantie van 2023 hadden zich op deze wijze 23 basisscholen aangemeld, die voldeden aan onze eis dat hun populatie voor een substantieel deel bestaat uit leerlingen behorende tot de doelgroepen van het Onderwijsachterstandenbeleid (OAB) en allemaal gebruikmaakten van hetzelfde leerlingvolgsysteem (Cito LVS) om de lees- en rekenprestaties van hun leerlingen te monitoren. Kort na de zomervakantie echter haakten vier scholen weer af. Zij gaven daarvoor verschillende redenen op, zoals personeelsgebrek, gelijktijdige deelname aan andere interventies en te weinig draagvlak in het team. Uit het aanmeldformulier van de resterende 19 kandidaat-deelnemers bleek dat de meeste andere scholen eveneens kampten met personeelsgebrek: slechts drie scholen gaven aan leden van het eigen schoolteam te kunnen inzetten voor de interventielessen (*'interne coaches'*); de andere 16 konden dat niet. Op deze scholen zouden de interventielessen dus moeten worden gegeven door externe coaches. De verantwoordelijkheid

³ Om na te gaan of de leerlingen niet een andere vorm van extra ondersteuning dan de interventie ontvangen (bijv. bijles), wordt per leerling informatie opgehaald met het leerlinginformatieformulier (zie paragraaf 3.3). Mocht deze vorm van 'non-compliance' zich voordoen bij leerlingen, dan wordt daar rekening mee gehouden in de analyse (zie paragraaf 4.2)

voor het leveren (en trainen) van deze externe coaches lag bij Beuker Onderwijs, die daarvoor de BijlesAcademie Amsterdam benaderde. Dit instituut was voor Beuker Onderwijs een bekende samenwerkingspartner, die al vaker was ingezet voor het verzorgen van extra lessen op scholen op basis van (een voorloper van) hetzelfde interventiemateriaal.

Vlak voor de start van de beoogde interventieperiode (oktober 2023) bleek echter dat de (meeste) coaches die de BijlesAcademie (BA) had geworven onvoldoende gekwalificeerd waren; het ontbrak hen aan leservaring en/of ervaring in het werken met kinderen in de basisschoolleeftijd. Dit werd geconstateerd door Beuker Onderwijs zelf tijdens de training van de BA-coaches en ook door de scholen die ter voorbereiding op de interventie contact hadden met deze coaches. Beuker Onderwijs zegde daarop de samenwerking met de BA op, met als gevolg dat er voor 16 van de aangemelde interventiescholen niet langer een coach beschikbaar was die de interventielessen kon verzorgen. Op twee scholen is dit opgelost door alsnog eigen teamleden als interne coaches in te zetten; op drie andere scholen hebben de twee medewerkers van Beuker Onderwijs zichzelf als externe coach opgeworpen. Voor 11 scholen waren dergelijke oplossingen niet haalbaar en is de deelname aan de interventie stopgezet. Uiteindelijk hebben er dus slechts **acht scholen** deelgenomen aan de interventie; drie met externe coaches en vijf met interne coaches. Een kanttekening hierbij is dat niet alle interne coaches gekwalificeerde leerkrachten waren; op enkele scholen is een stagiaire of onderwijsassistent ingezet. De externe coaches waren eerder overgekwalificeerd; beide medewerkers van Beuker Onderwijs hebben een lerarenopleiding en een universitaire opleiding afgerond, beschikken over een schat aan onderwijservaring inclusief ervaring als *remedial teacher* en hebben bovendien het interventiemateriaal zelf ontwikkeld en met de voorloper daarvan al vele jaren gewerkt.

Tabel 3.1 – Kenmerken van de acht interventiescholen, schooljaar 2023-2024*

School	Schoolgrootte (N leerlingen)	Aantal klassen groep 6	Aantal leerlingen groep 6	Interventie- domein	Coach intern/extern	Schoolgewicht
1	202	2	28	Begrijpend lezen	Intern	34,60
2	242	2	43	Begrijpend lezen	Extern	36,02
3	434	2	46	Rekenen-wiskunde	Intern	31,33
4	158	1	17	Rekenen-wiskunde	Extern	26,28
5	444	2	50	Begrijpend lezen	Extern	33,96
6	305	2	49	Rekenen-wiskunde	Intern	34,63
7	108	1	7	Begrijpend lezen	Intern	36,75
8	250	2	45	Beide	Intern	35,03

De acht interventiescholen konden zelf kiezen of de interventie bij hun leerlingen gericht zou worden op het domein begrijpend lezen of rekenen. Eén school koos zelfs voor beide domeinen; begrijpend lezen in de ene parallelklas van groep 6, rekenen in de andere. Voor de effectmeting is deze keuze overigens niet relevant, aangezien het effect van de interventie op de leerachterstand ongeacht het domein wordt vastgesteld (zie ook paragraaf 3.5). Tabel 3.1 geeft enkele

kenmerken van de acht interventiescholen. Het schoolgewicht in de laatste kolom is een indicatie van het aandeel OAB-doelgroep leerlingen op de school.⁴

3.3 Beschrijving onderzoeksinstrumenten

Tijdens schooljaar 2023-2024, het jaar waarin de interventie werd geïmplementeerd, is op de interventiescholen een groot aantal gegevens verzameld met behulp van verschillende, voornamelijk zelf ontwikkelde onderzoeksinstrumenten. Tabel 3.2 geeft een overzicht, en toont per instrument de verzamelde data en het afnamemoment. De interventie zelf vond plaats tussen begin november en eind januari.

Tabel 3.2 – Overzicht onderzoeksinstrumenten, verzamelde data en afnamemoment

Instrument	Verzamelde data	Afnamemoment	Interventieperiode
Leerlinginformatieformulier 1	Achtergrondgegevens en Cito LVS-scores* (eind groep 5) van alle leerlingen in groep 6	Sep 2023	
NSCCT	Leerpotentiescore alle leerlingen in groep 6	Nov 2023	
Logboek interventielessen	Verloop interventielessen per interventiegroepje volgens coach	Nov 2023 – Jan 2024	
Lesobservatie	Pedagogisch-didactische kenmerken geobserveerde interventielessen (2 lessen per interventiegroepje)	Dec 2023 – Jan 2024	
Vragenlijst ouders	Betrokkenheid ouders van leerlingen in experimentele conditie	Feb 2024	
Vragenlijst groepsleerkracht	Betrokkenheid groepsleerkracht van leerlingen in experimentele conditie	Feb 2024	
Leerlinginformatieformulier 2	Onderwijsbijzonderheden en Cito LVS-scores* (medio & eind groep 6) van alle leerlingen op Leerlinginformatieformulier 1	Jun 2024	

* Voor meer informatie over de Cito LVS-toetsen: zie paragraaf 4.2

Hierna volgt een korte omschrijving van het instrument en de gegevens die ermee zijn verzameld. In hoofdstuk 4 wordt het verloop van de dataverzameling en de opbrengst beschreven. In bijlage 1 zijn alle instrumenten opgenomen die specifiek voor dit onderzoek zijn ontwikkeld.

Leerlinginformatieformulier 1 en 2

Leerlinginformatieformulier 1 en 2 zijn digitale Excel-formulieren waarmee de interventiescholen op eenvoudige wijze informatie konden verstrekken over alle leerlingen in groep 6 in schooljaar 2023-2024. Het doel hiervan was een databestand op te bouwen aan de hand waarvan a) leerlingen konden worden geïdentificeerd die in aanmerking kwamen voor de interventie (zie paragraaf 3.2.4) en b) analyses konden worden uitgevoerd om de leergroei van de leerlingen in de interventie- en controleconditie met elkaar te vergelijken, onder controle van enkele achtergrondkenmerken.

⁴ Het schoolgewicht is een maat van de Inspectie van het Onderwijs die ligt tussen 20 en 40. Hoe hoger het schoolgewicht, hoe complexer de leerlingenpopulatie. Het schoolgewicht wordt berekend o.b.v. het opleidingsniveau en herkomstland van alle ouders, het gemiddelde opleidingsniveau en de verblijfsduur in Nederland van alle moeders, en het aandeel ouders in de schuldsanering.

- Met Leerlinginformatieformulier 1 werd informatie opgevraagd over de naam van de klas, geslacht, geboortemaand, geboortjaar en thuistaal en de vaardigheidsscores die de leerling eind groep 5 had behaald op de Cito-LVS toetsen Begrijpend lezen en Rekenen-Wiskunde. Ook is nagevraagd of de leerling volgens de school bij voorbaat niet in aanmerking kwam voor deelname aan de interventie, bijvoorbeeld vanwege een fysieke handicap, ernstig gedragsprobleem of onvoldoende beheersing van het Nederlands.
- Met Leerlinginformatieformulier 2 werden de medio en eind groep 6 behaalde vaardigheidsscores op de Cito LVS toetsen Begrijpend lezen en Rekenen-Wiskunde verzameld van alle leerlingen in groep 6 die ook voorkomen op Leerlinginformatieformulier 1.⁵ Daarnaast werd informatie opgevraagd over eventuele (onderwijs)bijzonderheden die zich hebben voorgedaan tijdens het schooljaar en die de lees- en/of rekenprestatie van de leerling mogelijk positief of negatief hebben beïnvloed; bijvoorbeeld langdurige afwezigheid door ziekte of het volgen van extra bijles.

Voor het opvragen van de leerlinggegevens via de Leerlinginformatieformulieren hoefden geen leerlingnamen uitgewisseld te worden. Bij het invullen van het eerste formulier is een uniek leerlingnummer toegekend aan elke leerling van groep 6. De scholen hebben hiermee voor eigen beheer een tweede versie van het eerste formulier aangelegd en bewaard. Alleen op deze 'schaduwlijst', uitsluitend toegankelijk voor de school, zijn leerlingnummers en leerlingnamen aan elkaar gekoppeld. Gedurende het hele onderzoek zijn alle gegevens over een leerling aan het bijbehorende leerlingnummer gekoppeld.

NSCCT

De Niet-Schoolse Cognitieve Capaciteiten Test (NSCCT) is het enige instrument in het onderhavige onderzoek dat we niet zelf hebben ontwikkeld. De NSCCT is ontwikkeld door onderzoekers van de Rijksuniversiteit Groningen en officieel goedgekeurd door de Commissie Test Aangelegenheden Nederland (COTAN).⁶ De (klassikale) test is afgenomen in groep 6 van alle interventiescholen, met als doel informatie te verzamelen over de leerpotentie van de leerlingen; een potentiële controlevariabele bij het analyseren van de leerwinst. De NSCCT voor groep 6 bestaat uit 93 meerkeuzeopgaven, verdeeld over vijf subtests: Figuur samenstellen, Exclusie, Getallenreeksen, Categorieën en Analogieën. De afname duurt ongeveer één uur. Tezamen geven de subtests een indruk van het leerpotentieel, uitgedrukt in een 'leerpotentiescore'. In het onderhavige onderzoek is de NSCCT afgenomen door de groepsleerkracht of intern begeleider (IB'er) met behulp van de handleiding die door KBA Nijmegen is verstrekt. De ingevulde papieren testboekjes zijn bij KBA Nijmegen nagekeken en verwerkt.⁷

Logboek interventiecoach

Het logboek is een digitale vragenlijst waarmee de coaches na elke interventieles konden registreren hoe de les was verlopen. Het doel hiervan was het meten van de implementatiekwaliteit van de interventie; oftewel de mate waarin de interventie is uitgevoerd zoals bedoeld (*fidelity of*

⁵ Over tussentijds ingestroomde leerlingen in groep 6 van de interventiescholen is dus geen informatie opgevraagd. Zij komen niet voor in dit rapport.

⁶ Meer informatie over de NSCCT is hier te vinden: <https://dia-nscc.nl/>.

⁷ De scholen hebben naderhand een terugkoppeling ontvangen. Dit schoolrapport toont de verdeling van de leerpotentiescores van de eigen leerlingen naast die van vergelijkbare scholen en een landelijke steekproef. Ook laat het rapport zien welke leerlingen risico lopen op onder- of overpresteren, o.b.v. de leerpotentiescores en de LVS-scores.

treatment). Voor zowel begrijpend lezen als rekenen-wiskunde bestaat de interventie uit twintig lessen, waarvan tien A-lessen (aanbod nieuwe lesstof) en tien B-lessen (herhaling en oefening). Per domein en per type les (A of B) verschilt de inhoud en structuur van de les, zodat er in feite vier lesvarianten bestaan. Het logboek sluit hierop aan door via routing alleen vragen te tonen die van toepassing zijn. Het gaat voornamelijk om gesloten vragen, bijvoorbeeld over de start- en eindtijd van de les, het behandelde hoofdstuk, het gebruikte materiaal, de leskwaliteit en bereikte lesdoelen volgens de coach zelf en eventuele bijzonderheden tijdens de les. Per interventiegroepje kreeg de betreffende coach via een unieke link toegang tot de vragenlijst en kon deze zo vaak als nodig – in principe na iedere interventieles – opnieuw invullen. De unieke link zorgde er ook voor dat bij het openen van het logboek automatisch een overzicht werd aangeboden van de leerlingen (leerlingnummers) in het betreffende interventiegroepje. De coach kon hierop aanvinken welke leerlingen tijdens de gelogde interventieles (niet) aanwezig waren.

Lesobservatie

Een observatieschema is ontwikkeld aan de hand waarvan getrainde onderzoeksassistenten van KBA Nijmegen enkele interventielessen van elk interventiegroepje systematisch konden observeren en beoordelen op een aantal pedagogisch-didactische aspecten. Het doel daarvan was zicht te krijgen op de pedagogisch-didactische kwaliteit van de interventieles. Op het observatieschema – bruikbaar voor de observatie van alle lessen, ongeacht domein of A-/B-les – worden vijf rubrieken onderscheiden. Vier rubrieken verwijzen naar de didactische lesopbouw volgens het EDI-model: Introductie, Instructie, Verwerking, Terugblik en afsluiting. De vijfde rubriek is Pedagogische aanpak en interactie. Binnen elke rubriek zijn op het observatieschema een aantal specifiekere activiteiten of handelingen benoemd die de coaches tijdens de interventielessen (kunnen) uitvoeren. De onderzoeksassistenten werden erin getraind tijdens de lessen te observeren of en hoe de coach deze activiteiten uitvoert en op basis daarvan een score toe te kennen. Daarnaast konden ze met een score de mate van betrokkenheid bij de geobserveerde interventieles van alle aanwezige leerlingen aangeven. Tot slot konden ze eventuele bijzondere situaties noteren die tijdens de geobserveerde interventieles voorgevallen zijn.

Oudervragenlijst

De oudervragenlijst is ontwikkeld om bij de ouders van de interventiedeelnemers informatie op te halen over hun betrokkenheid bij de interventie voor hun kind. De vragenlijst bestaat uit 12 stellingen, waarop de ouders konden reageren door te kiezen uit drie antwoordopties (Ja, Nee, Weet ik niet). Ook konden ze eventueel aanvullende opmerkingen noteren en aanvinken wie de vragenlijst heeft ingevuld. Bewust is de vragenlijst zeer kort en eenvoudig gehouden, op verzoek van de interventiescholen die aangaven dat veel van hun ouders weinig vaardig zijn in de Nederlandse taal.

Leerkrachtvragenlijst

De leerkrachtvragenlijst is ontwikkeld om bij de reguliere groepsleerkrachten van de leerlingen in de interventieconditie na te gaan in hoeverre zij betrokken zijn geweest bij het afstemmen van de interventie op het reguliere curriculum en de onderwijsbehoeften van de deelnemende leerlingen. De vragenlijst bestaat uit vijftien stellingen, verdeeld over drie rubrieken: Start, Kwaliteit en verloop en Opbrengst van de extra lessen. De groepsleerkrachten konden bij elke stelling kiezen uit vier antwoordopties (Waar, Neutraal, Niet waar, Weet ik niet/Niet van toepassing) en indien gewenst ook aanvullende toelichting of opmerkingen noteren. Voorafgaand aan de stellingen worden met de lijst twee vragen gesteld over de klasnaam en de coach van de interventielessen op de school (Ikzelf, Collega, Externe coach, Weet ik niet, Anders namelijk). Per school kregen de

groepsleerkrachten via een unieke link (doorgestuurd door de schoolcontactpersoon) toegang tot de vragenlijst. Als er meer dan een groepsleerkracht was, kon de schoolcontactpersoon zelf bepalen wie de leerkrachtvragenlijst het beste kon invullen.

3.4 Randomisatieprocedure

De toewijzing van leerlingen aan de experimentele groep (hierna verder aangeduid als interventiegroep) of de controlegroep is verlopen op basis van de door de interventiescholen verstrekte informatie op Leerlinginformatieformulier 1 (zie voorgaande paragraaf). Naast enkele achtergrondkenmerken van alle leerlingen in groep 6 betrof deze informatie ook de vaardigheidsscores die eind groep 5 waren behaald op de Cito-LVS toetsen Begrijpend lezen en Rekenen-Wiskunde. Op basis van deze vaardigheidsscores en het daarmee corresponderende vaardigheidsniveau⁸ konden alle leerlingen op de interventiescholen worden geïdentificeerd die, vanwege een lager vaardigheidsniveau (C, D of 'hoge' E⁹) op ten minste één van beide toetsdomeinen, in aanmerking kwamen voor de interventie. Dit waren er in totaal 175. Hierbij zijn geen leerlingen die volgens de school moesten worden uitgesloten van de interventie, bijvoorbeeld vanwege een fysieke handicap, gedragsprobleem of onvoldoende beheersing van het Nederlands.

Via een zo zorgvuldig mogelijke randomisatieprocedure zijn deze leerlingen aselekt toegewezen aan de interventie- of controlegroep. Op de grotere scholen was ruimte voor twee of drie interventiegroepjes, en dienden de leerlingen niet alleen te worden toegewezen aan de interventieconditie, maar ook aan een specifiek interventiegroepje¹⁰. In totaal zijn op de acht scholen 96 leerlingen (55% van de 175 leerlingen die in aanmerking kwamen) toegewezen aan de interventieconditie. Zij zijn verdeeld over negentien interventiegroepjes; dertien met de interventie begrijpend lezen, zes met de interventie rekenen-wiskunde. Negen interventiegroepjes kregen een externe coach toegewezen; tien een interne coach. Gemiddeld omvatten de groepjes dus vijf leerlingen, met een minimum van vier en een maximum van acht leerlingen. In totaal 79 leerlingen (45% van de 175) zijn aan de controlegroep toegewezen.

Dat niet precies evenveel leerlingen zijn toegewezen aan de interventieconditie als aan de controlegroep, komt doordat de interventiegroepjes per school en daarbinnen zoveel mogelijk ook per parallelklas¹¹ werden samengesteld. Als er een oneven aantal leerlingen (bijvoorbeeld zeven) in aanmerking kwam voor de interventie, is er één leerling meer toegewezen aan de interventieconditie (in het voorbeeld: vier) dan aan de controlegroep (in het voorbeeld: drie). Op één zeer kleine interventieschool (school 7 in tabel 3.1) zijn alle leerlingen die in aanmerking kwamen voor de interventie toegewezen aan de interventieconditie, in overleg met de school.

⁸ In het Cito-LVS corresponderen de vaardigheidsscores met een vaardigheidsniveau dat loopt van A t/m E. Dit geeft aan hoe de leerlingen – gegeven hun vaardigheidsscore – scoren ten opzichte van het landelijk gemiddelde. Leerlingen op niveau A en B scoren boven het landelijk gemiddelde, leerlingen op niveau C, D en E scoren daar (ruim) onder.

⁹ Leerlingen met een zeer lage score ('lage' E, vaak gemeten met een sterk afwijkende toetsversie, bijv. M4 in plaats van de eind groep 5 gebruikelijke versie E5) zijn op advies van Beuker Onderwijs buiten beschouwing gelaten, omdat zij naar verwachting onvoldoende kunnen profiteren van de interventie.

¹⁰ In de randomisatieprocedure werd gebruikgemaakt van een digitale tool en een dobbelsteen. Met de tool werden interventiedeelnemers random geselecteerd uit de totale groep leerlingen die daarvoor in aanmerking kwam; met de dobbelsteen werden de geselecteerde interventiedeelnemers eventueel ook nog toegewezen aan interventiegroepjes.

¹¹ Dit is grotendeels gelukt: in 16 van de 19 interventiegroepjes zitten alle deelnemers in dezelfde klas.

3.5 Analyseplan

De analyses in het onderzoek zijn – zoals eerder aangegeven – uitgevoerd volgens een RCT-design. Deze analyses hebben uitsluitend betrekking op data die zijn verzameld op en bij de acht interventiescholen, en die betrekking hebben op de leerlingen die in aanmerking kwamen voor de interventie (c.q. die random zijn toegewezen aan de interventie- of controleconditie).¹² Kortweg gaan we met deze analyses na of de leerlingen die aan de interventie hebben deelgenomen harder vooruit zijn gegaan dan hun klasgenoten met een vergelijkbaar startprestatieniveau die niet aan de interventie hebben deelgenomen, en in hoeverre het effect van de interventie wordt gemodereerd door bepaalde factoren. De analyse-aanpak bestaat uit drie stappen.

a. Beschrijving leerlingen

Als eerste stap beschrijven en vergelijken we de leerlingen in de interventie- en controleconditie op de interventiescholen aan de hand van de verzamelde achtergrondkenmerken (leeftijd, geslacht, thuistaal), leerpotentie, leerprestaties (LVS-toetsscores) in groep 5 en 6 en (onderwijs)bijzonderheden. Op basis van de vergelijking stellen we vast of de randomisatie succesvol is verlopen.

b. Beschrijving interventie

Zoals is beschreven in paragraaf 3.2, zijn in het onderzoek verschillende instrumenten ingezet om het proces van de interventie (de interventiecondities) in kaart te brengen. Als tweede analysestap voeren we op de data per instrument eerst enkele beschrijvende analyses uit om het verloop van de interventie in beeld te brengen. Op basis van de resultaten daarvan komen we tevens tot de constructie van indicatoren voor de implementatiekwaliteit, pedagogisch-didactische interactiekwaliteit en ouder- en leerkrachtbetrokkenheid, die worden meegenomen bij stap c.

c. Effect van de interventie en condities die dit effect modereren

De derde analysestap bestaat uit het bepalen van het effect van de interventie op het verkleinen van de leerachterstand en de mate waarin dit effect afhankelijk is van enkele condities. Daartoe analyseren we eerst de leerwinst van de leerlingen in de interventie- en controleconditie. De leerwinst van een leerling wordt bepaald door te kijken naar zijn of haar groei in vaardigheidsscore op de LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde tussen de verschillende meetmomenten (eind groep 5, midden en eind groep 6). Vervolgens stellen we het effect van de interventie op de leerwinst (c.q. het verkleinen van de leerachterstand) vast. We doen dit door *fixed effects* regressiemodellen te schatten.

De resultaten van analysestap a en b zijn beschreven in het volgende hoofdstuk. De resultaten van analysestap c komen aan de orde in hoofdstuk 5.

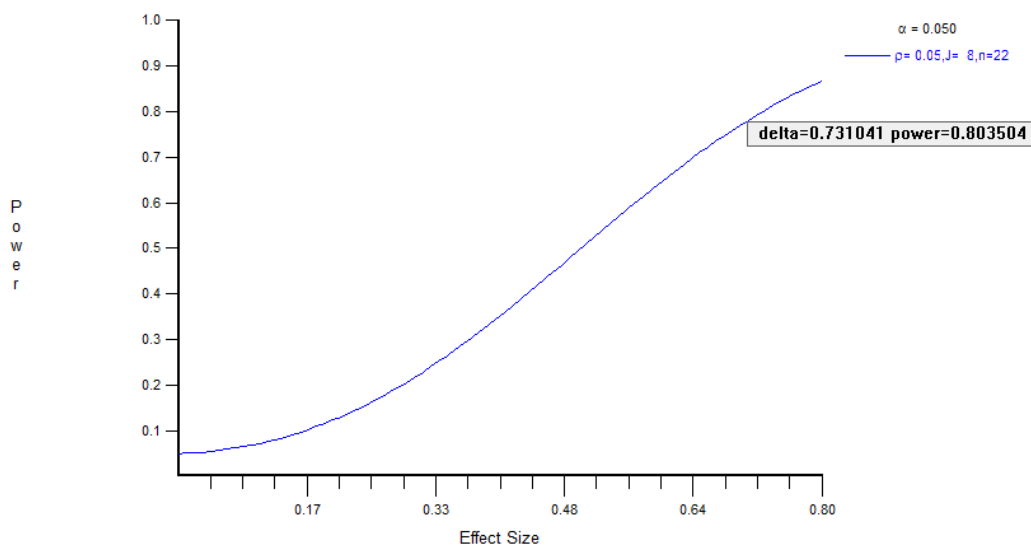
¹² De leerlingen op de interventiescholen die niet in aanmerking kwamen voor de interventie (bijvoorbeeld vanwege een te hoog startprestatieniveau) zijn dus niet meegenomen in deze analyses, ofschoon met sommige instrumenten (Leerlinginformatieformulier 1 & 2, NSCCT) wel gegevens over hen zijn verzameld. Dat was nodig voor het onderzoek en – in het geval van de teruggekoppelde NSCCT – ook zinvol voor de school zelf.

3.6 Poweranalyse

Bij de start van het schooljaar 2023-2024, waarin de interventie zou plaatsvinden, werd uitgegaan van een deelname van (ten minste) 20 scholen en minimaal 256 interventiedeelnemers. Deze benodigde aantallen kwamen overeen met de aantallen die door de interventieontwikkelaar en opdrachtgever voorgesteld werden. Dit betekent dat er per school zo'n 12 á 13 leerlingen zouden moeten deelnemen aan de interventie, met een ongeveer even groot aantal leerlingen in de controlegroep. Uitgaande van deze aantallen, een power van 0,80 en een significantieniveau van $\alpha=0,05$ (enkelzijdig getoetst) was de minimale effectgrootte (d) die aangetroffen zou moeten worden voor een statistisch significant effect ongeveer 0,25.

Zoals in paragraaf 3.2 reeds is beschreven, hebben we echter te maken gehad met onvoorziene omstandigheden waardoor een groot deel van de 23 aanvankelijk geworven scholen weer is uitgevallen. Uiteindelijk hebben er daardoor maar acht scholen deelgenomen aan de interventie. In totaal kwamen er op deze scholen 175 leerlingen in aanmerking voor deelname aan de interventie. Deze grote uitval van scholen (en leerlingen) leidde ertoe dat, uitgaande van een power van 0,80 en een significantieniveau van $\alpha=0,05$, de minimale effectgrootte die geobserveerd zou moeten worden voor een significant resultaat uitkomt op $d=0,43$.

In bovenstaande berekeningen is uitgegaan van een niet-geconditioneerde effectgrootte. Wanneer bij de poweranalyse ook nog rekening wordt gehouden met de clustering van leerlingen binnen klassen (er is feitelijk gerandomiseerd binnen klassen op scholen)¹³, dan wordt duidelijk dat de minimale effectgrootte die aangetoond zou moeten worden bij een power van 0,8 en een significantieniveau van $\alpha=0,05$ fors hoger ligt ($d=0,73$). We kunnen daarmee concluderen dat op basis van de gerealiseerde aantallen scholen en leerlingen het zeer lastig zal worden om een significant effect te vinden (zie onderstaand figuur).



Figuur 3.2 - Herziene powerberekening, 2-level CRT

¹³ Er wordt bij de berekening uitgegaan van een clustering van ongeveer 22 leerlingen ($n=22$) per school ($J=8$).

4 Beschrijvende resultaten

4.1 Inleiding

Ten behoeve van het onderzoek zijn tijdens schooljaar 2023-2024 op verschillende momenten gegevens verzameld om kenmerken van de leerlingen en van de interventie in kaart te brengen. In dit hoofdstuk presenteren we de resultaten daarvan. Eerst beschrijven en vergelijken we de leerlingen in de interventieconditie en de leerlingen in de controleconditie. Vervolgens beschrijven we de kenmerken van de interventie inclusief de wijze waarop de onderliggende data zijn verzameld en verwerkt.

4.2 Beschrijving deelnemers en randomisatie-uitkomst

Leerprestaties vóór de interventie en controlevariabelen

Van de 175 leerlingen die in aanmerking kwamen voor de interventie zijn er via een randomisatieprocedure 96 toegewezen aan de interventieconditie en 79 aan de controleconditie (zie ook paragraaf 3.4). Het bepalen welke leerlingen in aanmerking kwamen, gebeurde op basis van de vaardigheidsscores op de toetsen van Cito LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde die deze leerlingen eind groep 5 hadden behaald en die met behulp van Leerlinginformatieformulier 1 in september 2023 (bij de start van de leerlingen in groep 6) verzameld.

De Cito LVS-toetsen zijn betrouwbare en gevalideerde instrumenten, goedgekeurd door het College voor Toetsen en Examens (CvTE). Deze instrumenten worden landelijk op grote schaal gebruikt om de leervorderingen van de leerlingen op school te monitoren. Basisscholen zijn verplicht een leerlingvolgsysteem (LVS) te gebruiken om de ontwikkeling en resultaten van leerlingen systematisch bij te houden, om zo onderwijs af te kunnen stemmen op de behoeften van leerlingen. Bijna tachtig procent van de basisscholen gebruikt het leerlingvolgsysteem van Cito. De interventiescholen hebben zelf de Cito LVS-toetsen eind groep 5 (schooljaar 2022-2023) afgenomen, in het kader van het monitoren van de leervorderingen van de eigen leerlingen. Aangenomen mag worden dat ze daarbij de richtlijnen in de Cito-handleiding hebben gevolgd. De scores die met het Leerlinginformatieformulier zijn opgevraagd, zijn door ons gecontroleerd; waardes buiten de range van de toets zijn gecodeerd als 'ontbrekende waarde'. Met hetzelfde leerlinginformatieformulier zijn ook enkele achtergrondkenmerken van de leerlingen verzameld. In november 2023 is vervolgens de NSCCT afgenomen bij de leerlingen, resulterend in een leerpotentiescore.

In tabel 4.1 worden de leerlingen in de interventie- en controleconditie beschreven en vergeleken op alle voornoemde kenmerken. Het blijkt dat de leerlingen die zijn toegewezen aan de interventieconditie gemiddeld iets ouder zijn dan de leerlingen in de controleconditie. Het verschil is klein (twee maanden), maar statistisch significant ($p < 0,05$).¹⁴ Op alle overige kenmerken blijken de groepen niet significant van elkaar te verschillen. Daarmee kan in redelijkheid worden gesteld dat er sprake is van een random toewijzing van leerlingen aan de interventie- en controleconditie.

¹⁴ Getoetst via een t-toets voor onafhankelijke groepen (leeftijd, leerpotentie [NSCCT] en LVS-vaardigheidsscores) en een χ^2 -toets (geslacht en thuistaal).

Tabel 4.1 – Achtergrondkenmerken, leerprestaties eind groep 5 en leerpotentie; leerlingen in interventie- en controleconditie

	Leerlingen in interventieconditie (n=96)	Leerlingen in controleconditie (n=79)
Geslacht		
<i>Jongen</i>	38,50%	39,20%
<i>Meisje</i>	61,50%	60,80%
N excl. ontbrekende waarden	96	79
Thuistaal Nederlands		
<i>Ja</i>	50,00%	49,30%
<i>Nee</i>	27,20%	27,40%
<i>Overig (onbekend/niet eenduidig)</i>	22,80%	23,30%
N excl. ontbrekende waarden	92	73
Gemiddelde leeftijd in maanden op 1-9-2023 (sd) **	114,1 (6,5)	112,1 (5,9)
N excl. ontbrekende waarden	94	79
Gemiddelde LVS-score Begrijpend lezen eind groep 5 (sd)	136,14 (17,72)	137,51 (19,53)
N excl. ontbrekende waarden	96	79
Gemiddelde LVS-score Rekenen-Wiskunde eind groep 5 (sd)	192,82 (25,17)	195,70 (29,26)
N excl. ontbrekende waarden	76	60
Gemiddelde leerpotentiescore (sd)	86,7 (11,2)	87,2 (12,3)
N excl. ontbrekende waarden	90	72

** significant verschil in gemiddelden ($p < 0,05$)

Leerprestaties na de interventie

Aan het einde van schooljaar 2023-2024 zijn aan de hand van Leerlinginformatieformulier 2 nogmaals de vaardigheidsscores van de leerlingen op de Cito LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde opgevraagd; ditmaal de scores die zij medio groep 6 (één tot twee weken na afloop van de interventie) en eind groep 6 (ruim vier maanden na de interventie) hebben behaald (zie tabel 4.2). Ook deze Cito LVS-toetsen zijn door de scholen zelf afgenomen; de met het Leerlinginformatieformulier verzamelde vaardigheidsscores zijn door ons eerst weer gecontroleerd op buiten-range waarden. Deze scores dienen samen met de vaardigheidsscores van eind groep 5 als gegevensbron voor het vaststellen van de leergroei van de leerlingen (zie hoofdstuk 5).

Tabel 4.2 – Leerpresetaties midden/eind groep 6; leerlingen in interventie- en controleconditie

	Leerlingen in interven- tieconditie (n=96)	Leerlingen in controleconditie (n=79)
Gemiddelde LVS-score Begrijpend lezen medio groep 6 (sd)	155,71 (20,94)	156,47 (22,28)
N excl. ontbrekende waarden	91	77
Gemiddelde LVS-score Begrijpend lezen eind groep 6 (sd)	161,04 (22,59)	163,00 (25,27)
N excl. ontbrekende waarden	89	78
Gemiddelde LVS-score Rekenen-Wiskunde medio groep 6 (sd)	215,04 (29,21)	210,04 (33,12)
N excl. ontbrekende waarden	89	77
Gemiddelde LVS-score Rekenen-Wiskunde eind groep 6 (sd)	220,32 (29,84)	217,45 (36,42)
N excl. ontbrekende waarden	88	77

Tabel 4.2 laat zien dat de leerlingen in de interventiegroep en de controlegroep niet significant van elkaar verschillen wat betreft de behaalde scores op de LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde halverwege groep 6 en aan het einde van groep 6. Dit zegt overigens nog weinig over het effect van de interventie, onder andere omdat nog geen rekening is gehouden met non-compliance en overige uitval. De tabel beschrijft slechts de LVS-scores van de twee groepen leerlingen. In hoofdstuk 5 wordt het effect van de interventie op de leergroei nader onderzocht.

Experimentele uitval en non-compliance

Van de 96 leerlingen die gestart zijn met de interventie zijn er negen voortijdig weer uitgevallen; vaak al na één of enkele lessen. Volgens de coaches (die dit registreerden in het logboek) was de reden van deze uitval voor vijf leerlingen een verhuizing en/of schoolwissel. Eén leerling is door de coach en school uit de interventiegroep verwijderd wegens gedrags- en motivatieproblemen. Van de overige drie leerlingen is de reden van het voortijdig stoppen onbekend. Met behulp van een logistische regressieanalyse is nagegaan of de variabelen geslacht, thuistaal, leeftijd in maanden, leerpotentie en de scores op de LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde van eind groep 5 van invloed zijn op de kans op uitval. Deze (achtergrond)kenmerken van de leerlingen blijken echter niet significant van invloed te zijn op de kans op uitval, waarmee we kunnen concluderen dat de uitval random is. Niettemin worden deze negen leerlingen bij de analyses ter bepaling van het effect van de interventie (zie hoofdstuk 5) niet meegenomen.¹⁵

Naast uitval kan er ook sprake zijn van (onderwijs)bijzonderheden die leerlingen in de interventie- of controleconditie hebben meegemaakt in groep 6 en die positief of negatief van invloed zijn op hun ontwikkeling en leerpresetaties, bijvoorbeeld langdurige afwezigheid door ziekte of deelname aan bijles (anders dan de onderzochte interventie). Dergelijke omstandigheden (anders dan ‘care

¹⁵ Om onzuiverheid in het RCT-design te voorkomen is besloten deze leerlingen in zijn geheel uit de analyses te verwijderen. Van de meesten zijn geen LVS-scores van medio en eind groep 6 beschikbaar, zodat de leerwinst toch niet bepaald kon worden. Ook is onduidelijk in hoeverre deze leerlingen tijdens hun korte interventiedeelname extra ondersteuning boven op het reguliere onderwijsaanbod in groep 6 hebben ontvangen.

as usual') zouden de zuiverheid van het RCT-design kunnen aantasten. Met behulp van Leerling-informatieformulier 2 is bij de scholen nagegaan of zulke bijzonderheden zich hebben voorgedaan bij de betreffende leerlingen. Dat leverde echter geen informatie op waarmee rekening moet worden gehouden bij het analyseren van het effect van de interventie; dat geldt zowel voor de leerlingen in de interventieconditie als voor de leerlingen in de controleconditie.

Beide groepen hebben dus hetzelfde reguliere onderwijsaanbod in groep 6 ontvangen; de leerlingen in de interventieconditie hebben aanvullend daarop extra ondersteuning ontvangen in de vorm van de interventie. Daarmee wordt voldaan aan een belangrijke voorwaarde voor een zuivere effectmeting (zie ook paragraaf 3.1).

4.3 Beschrijving interventie

4.3.1 Implementatiekwaliteit

Dataverzameling en respons

De implementatiekwaliteit van de interventie is gemeten met het digitale logboek waarin de coaches van de interventiegroepen vragen beantwoord hebben over de aanpak en het verloop van de interventielessen. Tijdens de interventieperiode is herhaaldelijk gecontroleerd of het aantal gelogde interventielessen synchroon liep met het aantal volgens de planning uitgevoerde lessen. Coaches die achterbleven met het bijhouden van het logboek zijn daar per e-mail of telefonisch op gewezen. Uiteindelijk zijn er met het logboek 252 verschillende interventielessen geregistreerd.

Dataverwerking

Van twee interventiegroepjes zijn met het logboek alle twintig lessen geregistreerd waaruit de interventie maximaal bestaat. Van de overige zeventien interventiegroepjes zijn er minder dan twintig lessen geregistreerd. De redenen daarvoor zijn niet meer te achterhalen. Het kan zijn dat coaches na sommige lessen verzuimd hebben het logboek in te vullen, maar het is ook mogelijk dat coaches het lesmateriaal in minder dan twintig interventielessen hebben behandeld. Om die reden is besloten de resultaten van alle wél gelogde lessen van elk interventiegroepje te verwerken tot vergelijkbare scores en deze te middelen. Het resultaat verwijst dan naar de gemiddelde implementatiekwaliteit van de aangeboden interventie aan het betreffende interventiegroepje, ongeacht het aantal interventielessen op basis waarvan dit gemiddelde is bepaald. Alleen wanneer van een interventiegroepje minder dan tien interventielessen zijn gelogd is dat achterwege gebleven, omdat dat aantal te laag is om met voldoende zekerheid de gemiddelde implementatiekwaliteit te bepalen. Dit geldt voor twee interventiegroepjes; de logboekgegevens van deze groepjes zijn dus niet meegenomen in de multivariate analyses. Van de overige zeventien interventiegroepjes waren in totaal 241 logboekregistraties beschikbaar die in de analyses meegenomen zijn.

Met het logboek konden vier verschillende lesvarianten (A-lessen en B-lessen voor begrijpend lezen of rekenen-wiskunde) in kaart worden gebracht. Afhankelijk van de lesvariant zijn aan de coaches via routing deels andere vragen voorgelegd. Om de gelogde lessen toch vergelijkbaar te kunnen scoren zijn alleen de gemeenschappelijke aspecten bij de scoring betrokken en

meegenomen in de multivariate analyses.¹⁶ In het lesboek is elke interventieles, ongeacht lesvariant, opgebouwd volgens dezelfde structuur van Introductie (o.a. voorkennis activeren, lesdoelen benoemen), Instructie en oefenen (o.a. nieuwe strategie of woord aanbieden, opdrachten maken), en Afronding (o.a. nabespreken, vooruitblik). In het logboek zijn over deze drie lesonderdelen enkele vragen gesteld waarmee de coaches konden aangeven of de genoemde leselementen tijdens de interventieles behandeld zijn (ja=1, nee=0). Ook hebben de coaches in het logboek altijd vragen beantwoord over het behalen van de lesdoelen tijdens de gelogde les (nee=0, deels=1, ja=2) en aangegeven welke beoordeling ze geven aan de lesuitvoering als geheel (rapportcijfer tussen 1 en 10). Op basis van deze gemeenschappelijke onderdelen levert elk ingevuld logboek dus vijf scores op, voor respectievelijk Introductie, Instructie en oefenen, Afronding, Behalen lesdoelen en Algemeen oordeel uitgevoerde les. De optelling van deze scores leidt tot een totaalscore die verwijst naar de implementatiekwaliteit van de gelogde les en een range heeft van 0-17. Sommige onderdelen leggen daarbij meer gewicht in de schaal dan andere (zie tabel 4.3); bijvoorbeeld het onderdeel 'Instructie en oefenen' kan maximaal zes punten opleveren, aangezien de achterliggende items verwijzen naar de kern van de interventieles. Hoe hoger de score, hoe beter de coach zich (naar eigen zeggen) heeft gehouden aan het uitvoeren van de interventie zoals beoogd en des te hoger dus de implementatiekwaliteit. Vervolgens is per interventiegroepje (met uitzondering van de twee groepjes met onvoldoende gelogde lessen) ook de gemiddelde implementatiekwaliteit berekend, door de totaalscores van de afzonderlijke lessen te middelen. Deze scores worden meegenomen in de multivariate analyses.

Beschrijvende resultaten

Tabel 4.3 toont de scores van alle lessen van de zeventien interventiegroepjes waarvan meer dan tien lessen zijn gelogd. Ook zijn enkele contextkenmerken van de lessen weergegeven, eveneens met het logboek verzameld. De coaches hebben niet altijd alle vragen over de gelogde les beantwoord, waardoor het aantal lessen waarop de cijfers betrekking hebben soms lager is dan 241. Onderaan de tabel worden enkele contextkenmerken vermeld van de interventiegroepjes. Ook deze kenmerken zijn verzameld met het logboek, maar ze hebben betrekking op het groeps- en niet het lesniveau. Daarom zijn deze kenmerken beschikbaar over alle negentien interventiegroepjes, ook de twee waarvan minder dan tien lessen zijn gelogd.

Voor de zeventien interventiegroepjes waarvan voldoende interventielessen zijn gelogd, is ook de gemiddelde implementatiekwaliteit van de interventie berekend. Deze score, die wordt meegenomen in de multivariate analyses (zie hoofdstuk 5) bedraagt gemiddeld 11,5 met een range van 7,5-15,9.

¹⁶ Enkele conditionele logboekvragen (afhankelijk van het antwoord op voorafgaande vragen wel/niet voorgelegd) zijn dus buiten beschouwing gelaten; bijv. over gebruik spelmaterialen en aantal behandelde rekenopdrachten/leesteksten.

Tabel 4.3 – Implementatiekwaliteit en contextkenmerken interventielessen/interventiegroepjes

Implementatiekwaliteit interventielessen	Gemiddelde	sd	N (max=241)*
Introductie (0-2)	1,50	0,73	241
Instructie en oefenen (0-6)	4,04	2,15	241
Afronding (0-2)	1,41	0,81	241
Lesdoelen behaald (0-4)	2,45	1,14	236
Algemeen oordeel les (0-3)	2,27	0,70	236
Totaal (0-17)	11,66	4,12	235
Contextkenmerken interventielessen	Gemiddelde	sd	N (max=241)*
Duur interventielessen (gem.) <i>Range</i>	53,9 minuten <i>Min. 20, max. 75</i>	9,0	239
Aantal aanwezige leerlingen (gem.) <i>Range</i>	5,2 <i>Min. 4, max. 8</i>	0,8	241
Contextkenmerken interventiegroepjes	%		N (max=19)**
Interne coach	53%		10
Externe coach	47%		9
Tijdens schooltijd	42%		8
Na schooltijd	47%		9
Wisselend	11%		2

* Betreft 17 van de 19 interventiegroepjes / ** Betreft alle 19 interventiegroepjes

4.3.2 Pedagogisch-didactische kwaliteit

Dataverzameling en respons

De pedagogisch-didactische kwaliteit van de interventielessen is gemeten met lesobservaties die zijn uitgevoerd door getrainde onderzoeksassistenten van KBA Nijmegen aan de hand van een observatieschema. De lesobservaties zijn in overleg met de scholen ingepland en gefaciliteerd. Bij elk interventiegroepje zijn volgens planning twee lesobservaties uitgevoerd, in week 4 of 5 respectievelijk week 7 of 8 van de interventieperiode (die in totaal tien weken duurde). In totaal beschikken we dus over 38 ingevulde observatieschema's. Deze hebben betrekking op 26 interventielessen begrijpend lezen (twee lessen van dertien interventiegroepjes) en twaalf interventielessen rekenen-wiskunde (twee lessen van zes interventiegroepjes).

Dataverwerking

De onderzoeksassistenten hebben op het observatieschema scores toegekend aan de kwaliteit van een aantal door de coach uitgevoerde didactische en pedagogische activiteiten en handelingen tijdens de geobserveerde les (0=nee, niet gezien, 1=ja, matige kwaliteit/onvolledig, 2=goede kwaliteit/volledig, 9=niet van toepassing; deze laatste score is achteraf tot 'missing waarde' verklaard). Ook hebben ze scores toegekend aan de betrokkenheid van de individuele leerlingen bij de geobserveerde interventielessen (0=slecht, 1=matig, 2=goed). Voor alle 38 geobserveerde interventielessen zijn deze scores beschikbaar.

Op basis van de toegekende scores voor de geobserveerde activiteiten en handelingen van de coach zijn aan de hand van een factor- en betrouwbaarheidsanalyse vijf subschalen

geconstrueerd.¹⁷ Drie daarvan verwijzen vooral naar (de kwaliteit van) didactische aspecten in de lesaanpak van de coach: a) de introductie en afsluiting (7 items, Cronbachs alpha 0,78), b) de instructie (4 items, Cronbachs alpha 0,60) en c) de verwerking (6 items, Cronbachs alpha 0,90). Twee subschalen verwijzen vooral naar (de kwaliteit van) pedagogische aspecten in de lesaanpak van de coach: d) het lesklimaat (3 items, Cronbachs alpha 0,78) en e) het leerlinggedrag (2 items, Cronbachs alpha 0,76). Alle schalen lopen van 0 tot 2; hoe hoger, hoe beter de didactische of pedagogische kwaliteit.

Daarnaast hebben we op basis van de scores voor lesbetrokkenheid van alle afzonderlijke leerlingen een score berekend voor de gemiddelde lesbetrokkenheid van de leerlingen in elke geobserveerde les. Ook deze score loopt van 0 tot 2 en kan worden gezien als een aanvullende indicator voor de kwaliteit van de pedagogisch-didactische interactie tijdens de interventies.

Door middel van een tweede orde factoranalyse is vervolgens nagegaan of er op basis van de geconstrueerde subschalen één gezamenlijke, overkoepelende schaal voor de pedagogisch-didactische kwaliteit van (de interactie tijdens) de les gevormd kan worden. Dat bleek inderdaad mogelijk. De resulterende schaal, bestaande uit een optelling van de scores op de zes subschalen¹⁸, heeft een voldoende betrouwbaarheid (Cronbachs alpha 0,76) en is meegenomen in de multivariate analyses (zie hoofdstuk 5).

Beschrijvende resultaten

In tabel 4.4 worden de scores op de afzonderlijke zes schalen weergegeven, voor alle 38 geobserveerde interventielessen.

Tabel 4.4 – Pedagogisch-didactische kwaliteit geobserveerde interventielessen (n=38); scores op de subschalen en totaal (gemiddelde, standaarddeviatie)

	Gemiddelde	sd
Didactisch		
Introductie en afsluiting	1,39	0,54
Instructie	1,82	0,30
Verwerking	1,67	0,48
Pedagogisch		
Klimaat	1,86	0,30
Gedrag	1,88	0,30
Betrokkenheid leerlingen	1,65	0,39
Pedagogisch-didactische kwaliteit	10,22	1,63

¹⁷ De betrouwbaarheid van de schalen is bepaald met Cronbachs alpha. Een alpha van ten minste 0,70 is een indicatie voor een goede betrouwbaarheid; een alpha van 0,60-0,70 wordt gezien als voldoende (Kline, 1999). Enkele items uit het observatieschema zijn niet meegenomen in de schalen omdat ze de betrouwbaarheid verlaagden.

¹⁸ Bij drie lessen kon één van de subschaalscores niet worden berekend, omdat aan alle items de score 9 (niet van toepassing) was toegekend. Dit is opgelost door de score te imputeren die bij de andere geobserveerde les aan hetzelfde groepje was gegeven op de betreffende subschaal.

4.3.3 Betrokkenheid ouders en reguliere groepsleerkracht

Dataverzameling en respons

De contactpersonen van de acht interventiescholen hebben van de onderzoekers het verzoek gegeven om de vragenlijsten voor de ouders en de groepsleerkrachten van de interventiedeelnemers te verspreiden:

- De (papieren) oudervragenlijst kon de contactpersoon meegeven aan de leerlingen zelf (n=96). De vragenlijst zat in een envelop waarin ook een korte brief met uitleg en een antwoordenvelop waren bijgevoegd. Op elke vragenlijst was een uniek leerlingnummer afgedrukt; de school kon op het overzicht van namen gekoppeld aan nummers (dat op school werd bewaard) nazoeken welke vragenlijst bij welke leerling hoorde. De ouders konden de vragenlijst anoniem invullen en in de antwoordenvelop doen. Deze envelop konden ze zelf op de post doen of afgeven op school.
- De e-mail met de link naar de (digitale) leerkrachtvragenlijst en een korte uitleg kon de contactpersoon doorsturen naar de reguliere leerkrachten van groep 6. De bedoeling was dat uit elke parallelklas van groep 6 één groepsleerkracht de vragenlijst zou invullen; in totaal waren er veertien parallelklassen voor groep 6 op de acht interventiescholen.

Voor beide vragenlijsten zijn meerdere herinneringen uitgestuurd naar de schoolcontactpersonen. Uiteindelijk hebben we van de ouder(s) van 38 interventiedeelnemers een ingevulde oudervragenlijsten ontvangen, een respons van 39 procent. Van twee interventiescholen is geen enkele oudervragenlijst ontvangen. De leerkrachtvragenlijst is door zeven groepsleerkrachten van vijf verschillende interventiescholen ingevuld, een respons van vijftig procent. Van drie interventiescholen is helemaal geen leerkrachtvragenlijst ontvangen.

Dataverwerking

In de tabellen hierna worden de resultaten van de ouder- en leerkrachtvragenlijsten eenvoudig gepresenteerd, door per stelling te laten zien hoeveel procent van de respondenten deze stelling onderschreef (antwoord 'ja'). De respons op de beide vragenlijsten is echter dermate laag dat we het niet verantwoord achten om op basis van deze resultaten uitspraken te doen over de invloed van de ouder- en groepsleerkrachtbetrokkenheid op de effectiviteit van de interventie. De met de vragenlijsten verzamelde data worden daarom niet meegenomen bij de multivariate analyses (zie hoofdstuk 5) en onderzoeksvraag 4 kan op basis hiervan niet worden beantwoord. De ruwe resultaten per stelling zijn dan ook niet verwerkt tot indicatoren voor de betrokkenheid van de ouders en groepsleerkrachten.

Beschrijvende resultaten

Tabel 4.5 en 4.6 laten zien welk deel van de responderende ouders en groepsleerkrachten de voorgelegde stellingen heeft onderschreven. De vermelde percentages zijn berekend ten opzichte van alle respondenten, dus inclusief enkelen die kozen voor het antwoord 'weet ik niet/niet van toepassing'.

Tabel 4.5 – Resultaten oudervragenlijst; aantal en % ouders dat de stellingen onderschrijft (n=38)

Stelling	'Ja'	
	N	%
1 Vindt u het prettig/fijn dat uw kind extra lessen krijgt?	37	97
2 Bent u tevreden over het tijdstip (dag en tijd) van de extra lessen?	33	87
3 Vertelt de school/leerkracht u wat uw kind leert bij de extra lessen?	20	53
4 Vertelt de school/leerkracht u hoe het met uw kind gaat bij de extra lessen?	14	37
5 Praat u met uw kind over de extra lessen?	35	92
6 Vindt uw kind de extra lessen leuk?	32	84
7 Leert uw kind nieuwe dingen in de extra lessen?	33	87
8 Werkt uw kind hard tijdens de extra lessen?	28	74
9 Doet u thuis met uw kind de spelletjes uit de extra lessen?	12	32
10 Gaat het op school/in de klas beter met uw kind door de extra lessen?	16	42
11 Leest u thuis boekjes samen met uw kind?	25	66
12 Doet u thuis rekenspelletjes samen met uw kind?	30	79

Tabel 4.6 – Resultaten leerkrachtvragenlijst; aantal en % dat de stellingen onderschrijft (n=7)

Stelling	'Waar'	
	N	%
De start van de extra lessen		
1 Ik ben voldoende betrokken bij het besluit van mijn school om mee te doen aan de interventie/de extra lessen	4	57
2 Ik ben tevreden over de selectie van leerlingen uit mijn klas voor de extra lessen	7	100
3 Ik ben tevreden over de tijdstippen waarop de extra lessen plaatsvonden	3	43
4 Ik ben bij de start voldoende geïnformeerd over de inhoud en werkwijze van de extra lessen	2	29
5 Degene die de extra lessen gaf heeft van tevoren de specifieke onderwijsbehoeften van de deelnemende leerlingen uit mijn klas geïnventariseerd	3	43
De kwaliteit en het verloop van de extra lessen		
6 De inhoud van de extra lessen sloot goed aan bij de lesstof die in de klas werd behandeld	3	43
7 De extra lessen sloten goed aan bij de individuele onderwijsbehoeften van de deelnemende leerlingen uit mijn klas	3	43
8 Er werd in de extra lessen voldoende gedifferentieerd naar de verschillende niveaus	2	29
9 De deelnemende leerlingen uit mijn klas gingen graag naar de extra lessen	6	86
10 De didactische aanpak in de extra lessen sloot goed aan op de aanpak in de klas	2	29
De opbrengst van de extra lessen		
11 De achterstand van de deelnemende leerlingen uit mijn klas is door de extra lessen kleiner geworden	4	57
12 De deelnemende leerlingen doen door de extra lessen beter mee met de lessen in de klas	3	43
13 De deelnemende leerlingen uit mijn klas hebben door de extra lessen meer zin om te leren dan voorheen	3	43
14 Ook andere leerlingen uit mijn klas zouden moeten kunnen deelnemen aan deze extra lessen	5	71
15 De oefeningen, opdrachten e/o spelletjes uit het lesboek vd. extra lessen worden ook in de klas toegepast	1	14

De soms zeer hoge percentages in tabel 4.5 suggereren dat er in enige mate sprake is geweest van sociaal wenselijke antwoorden van de ouders, al valt ook op dat dit niet geldt voor bepaalde stellingen die misschien wel het meest verwijzen naar ouderbetrokkenheid (zie bijvoorbeeld de stellingen 3, 4 en 9). In tabel 4.6 blijkt dat de leerkrachtbetrokkenheid zelfs binnen deze kleine groep respondenten aardig varieert (zie bijvoorbeeld de stellingen 7 en 10). Een en ander maakt het des te spijtiger dat, vanwege de lage respons, in het volgende hoofdstuk niet wordt nagegaan in hoeverre de ouder- en leerkrachtbetrokkenheid een modererende rol spelen bij het effect van de interventie.

5 Effect van de interventie en modererende condities

5.1 Bepalen leerwinst

5.1.1 Inleiding

In paragraaf 4.2 constateerden we dat de toewijzing van leerlingen aan de interventie- en controlegroep willekeurig is verlopen – afgemeten aan het feit dat er geen verschillen zijn naar de gemeten achtergrondkenmerken tussen de interventie- en controlegroep (met uitzondering van de variabele leeftijd). Met die wetenschap brengen we in deze paragraaf de effecten van de interventie ‘Doelgericht aanbod verlengde leertijd in kleine groepen’ op de leerachterstand van de interventiedeelnemers zo goed mogelijk in kaart door de leerachterstand en de ontwikkeling daarin van deelnemers en niet-deelnemers te vergelijken. We beantwoorden daarmee de eerste onderzoeksvraag: *Wat is het effect van de interventie ‘Doelgericht aanbod verlengde leertijd in kleine groepen’ op de leerachterstand van de deelnemende leerlingen in relatie tot andere leerlingen in dezelfde jaargroep?*

Voorafgaand daaraan willen we hier eerst een belangrijke kanttekening maken. Zoals is toegelicht in hoofdstuk 3 is de beoogde onderzoeksgroep, kort voor de interventie van start zou gaan, meer dan gehalveerd door het tekort aan voldoende gekwalificeerde coaches. Door deze (mogelijk selectieve) uitval (en de mogelijk ontstane heterogeniteit in implementatiekwaliteit van de interventie) is de statistische power van de analyses die in dit hoofdstuk worden gepresenteerd aanzienlijk verlaagd. We kunnen met andere woorden slechts met beperkte zekerheid vaststellen of verschillen in leergroei van de interventie-deelnemers in vergelijking tot de leerlingen in de controlegroep daadwerkelijk zijn toe te schrijven aan de werking van de interventie ‘Doelgericht aanbod verlengde leertijd’. Desondanks wordt de term ‘effect’ in dit hoofdstuk voorzichtig gebruikt, in aansluiting op het gehanteerde (RCT-)design en de uitgevoerde analyses.

5.1.2 Bivariate analyse

In tabel 5.1 presenteren we de gemiddelde groei van leerlingen uit de interventie- en controlegroep op de LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde. We brengen zowel de gemiddelde groei tussen eind groep 5 en midden groep 6 (direct na de interventieperiode) als die tussen eind groep 5 en eind groep 6 (enkele maanden na de interventieperiode) op de twee domeinen in kaart. In de tabel zijn de gemiddelde groeiscoringen en standaardafwijkingen vermeld. De scores van de verschillende momenten (E5, M6 en E6) kennen eenzelfde metriek en daarmee kan er een groeiscoring berekend worden. Voor de scores van de verschillende toetsdomeinen geldt dat deze verschillende ranges kennen en niet met elkaar vergeleken kunnen worden.

Tabel 5.1 - Gemiddelde groei (standaarddeviatie tussen haakjes) op LVS-toetsen Begrijpend lezen en Rekenen-Wiskunde, naar groep)

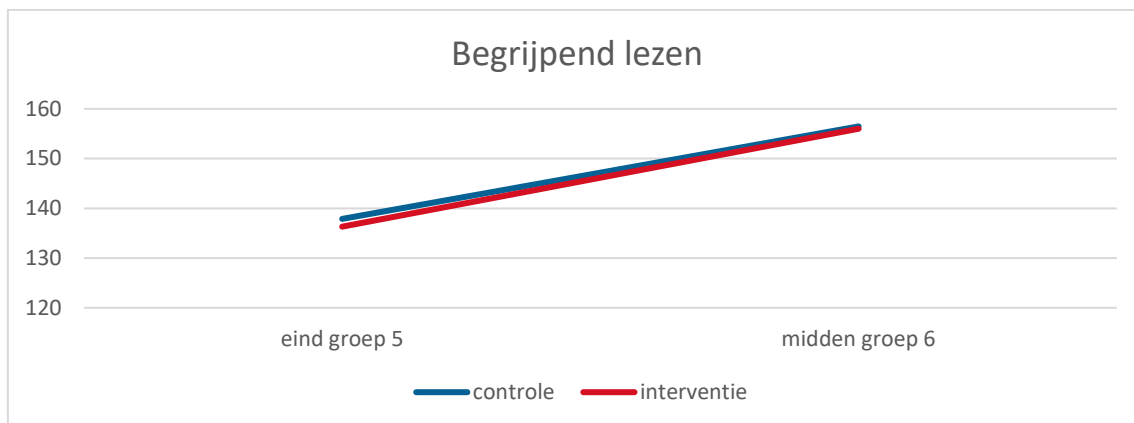
	Leerlingen in interventieconditie	Leerlingen in controleconditie
Begrijpend lezen		
Groei eind groep 5 – midden groep 6	19,69 (16,49)	18,59 (18,29)
N excl. ontbrekende waarden	86	77
Groei eind groep 5 – eind groep 6	24,93 (17,70)	25,27 (22,32)
N excl. ontbrekende waarden	85	78
Rekenen-Wiskunde		
Groei eind groep 5 – midden groep 6 ***	24,70 (17,56)	16,31 (19,44)
N excl. ontbrekende waarden	70	59
Groei eind groep 5 – eind groep 6	33,50 (17,56)	29,02 (22,02)
N excl. ontbrekende waarden	62	50

*** significant verschil in gemiddelden ($p < 0,01$)

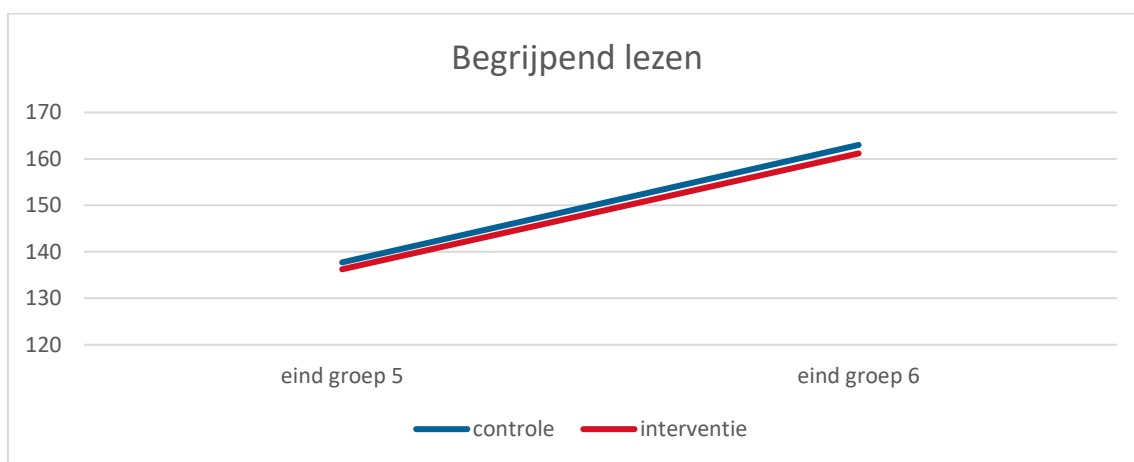
Door middel van een t-toets voor onafhankelijke groepen is nagegaan of de gemiddelde groeiscoringen van de controle- en interventiegroep significant van elkaar verschillen.¹⁹ Leerlingen in de interventieconditie zijn tussen de metingen van eind groep 5 en midden groep 6 significant sterker gegroeid in de rekenvaardigheid dan de leerlingen in de controleconditie (klein tot middelgroot effect, Cohen's $d=0,471$). Op de langere termijn, tussen eind groep 5 en eind groep 6, vinden we geen verschil in de mate van groei in rekenvaardigheid tussen leerlingen in de verschillende condities. Op het gebied van begrijpend lezen hebben leerlingen in de interventie- en controleconditie zich vergelijkbaar ontwikkeld: er is geen significant verschil in de groei van de leerlingen van beide groepen op de korte én middellange termijn.

In de figuren 5.1 t/m 5.4 wordt de groei van de leerlingen in de verschillende condities op het gebied van begrijpend lezen en rekenen-wiskunde grafisch weergegeven.

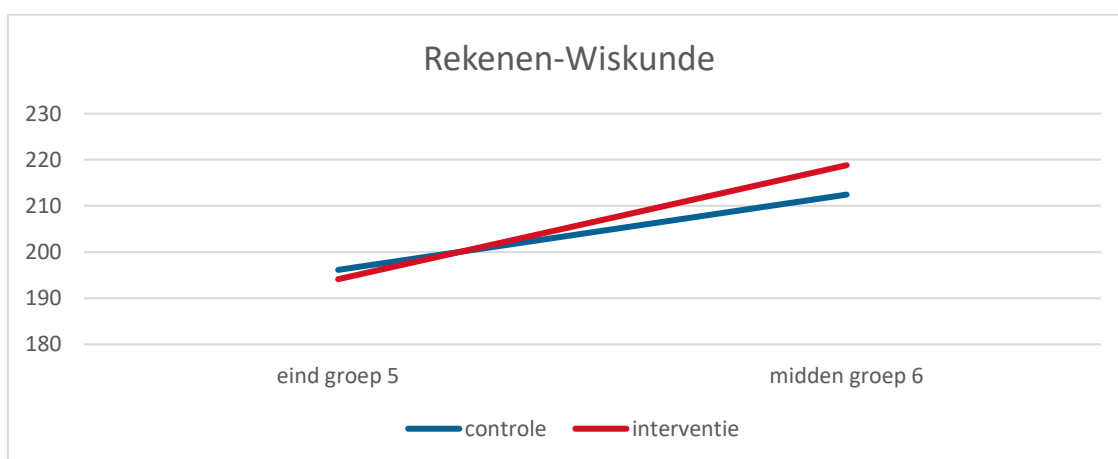
¹⁹ We hanteren dus feitelijk een *difference-in-differences* (DiD) aanpak. De belangrijkste reden hiervoor is dat we inhoudelijk geïnteresseerd zijn in de onderzoeksvraag of de gemiddelde leergroei van de leerlingen in de interventieconditie significant sterker is dan die van degenen in de controleconditie, ofwel of de leerachterstand van de leerlingen die aan de interventie hebben deelgenomen kleiner is geworden als gevolg van hun participatie daarin. In de literatuur bestaat veel discussie over deze aanpak (d.w.z. de analyse van verschil- of groeiscoringen) en wordt regelmatig gesuggereerd dat het beter is een regressiemodel te schatten waarbij de nameting afhankelijk wordt gesteld van de interventie en de voormeting als covariaat wordt opgenomen (covariantie-analyse). Deze kwestie staat ook wel bekend als *Lord's paradox* (Lord, 1967). Qua interpretatie wordt in het laatste geval echter een andere onderzoeksvraag beantwoord, namelijk of er een effect van de interventie is op het gemiddelde uitkomstverschil, geconditioneerd op de pre-score. Hoewel het een terechte vraag is of je in een gerandomiseerde studie een analyseaanpak van verandering (*change*) zou willen hanteren (zelfs het controleren voor de *baseline* zou in dit geval niet nodig zijn), blijft het onduidelijk welke aanpak de voorkeur verdient. Om vast te houden aan onze eigen onderzoeksvraag én tegelijkertijd tegemoet te komen aan de geopperde kritiek op de analyse van groeiscoringen (zoals het optreden van regressie naar het gemiddelde) schatten we in paragraaf 5.1.3 het effect van de interventie aan de hand van een regressiemodel waarbij de groeiscoring afhankelijk wordt gesteld van de interventie én de pre-score. Dit model is wiskundig gezien equivalent aan covariantie-analysebenadering.



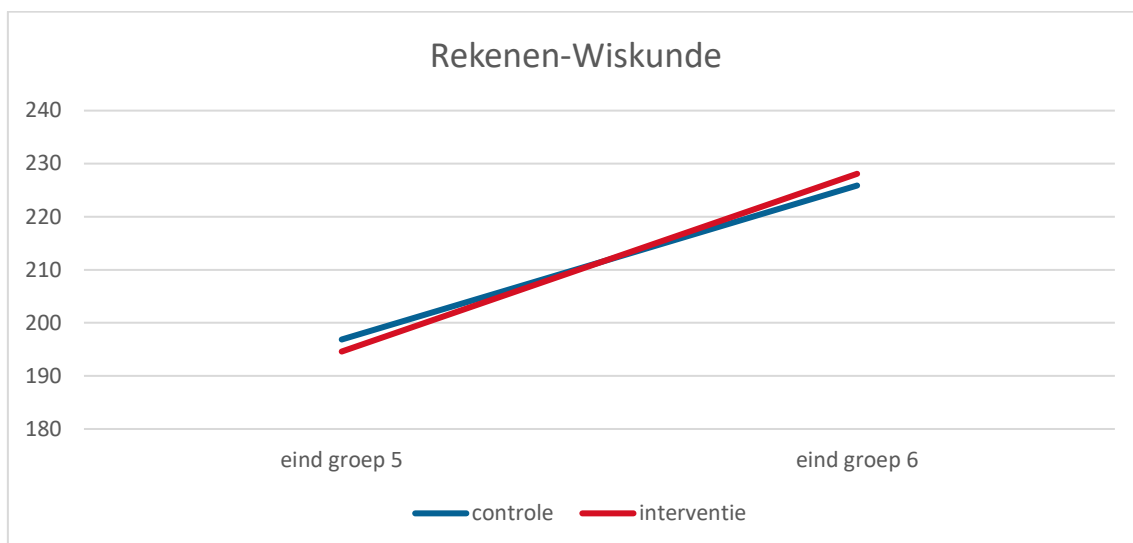
Figuur 5.1 Scoreverloop op LVS-toets Begrijpend lezen korte termijn, naar conditie (n=163)



Figuur 5.2 Scoreverloop op LVS-toets Begrijpend lezen middellange termijn, naar conditie (n=163)



Figuur 5.3 Scoreverloop op LVS-toets Rekenen-Wiskunde korte termijn, naar conditie (n=129)



Figuur 5.4 Scoreverloop op LVS-toets Rekenen-Wiskunde middellange termijn, naar conditie (n=112)

5.1.3 Multivariate analyse

In de voorgaande paragraaf zijn de groeiscoringen bivariaat geanalyseerd. In deze paragraaf worden dezelfde groeiscoringen aan de hand van multivariate analyse onderzocht. We gaan na of de leerlingen in de interventie- en controleconditie zich verschillend ontwikkeld hebben op de vaardigheden begrijpend lezen en rekenen-wiskunde, waarbij rekening wordt gehouden met relevante verschillen tussen leerlingen en scholen.

De resultaten die in deze paragraaf worden gepresenteerd, zijn gebaseerd op *fixed effects* regressiemodellen. Meer specifiek hebben we *least square dummy variable (LSDV)* modellen geschat. In deze modellen wordt het regressie-effect van de interventieconditie (experimenteel versus controle) op de groeiscore (korte of middellange termijn) bepaald waarbij – door dummies voor iedere school op te nemen – volledig rekening wordt gehouden met de verstoringende invloed van alle (ongemeten) verschillen tussen scholen.

Per afhankelijke variabele zijn twee verschillende modellen geschat. Allereerst betreft dit een model waarbij het effect van deelname aan de interventie (ongeacht type interventie – lezen of rekenen) is bepaald. Vervolgens gaat het om een model waarbij het effect van het specifieke type interventie is gepresenteerd. In beide modellen is de controlegroep als referentiecategorie gehanteerd.

Ook is de score op de voormeting (eind groep 5) in de modellen opgenomen om regressie naar het gemiddelde te voorkomen (vanwege bodem- en plafondeffecten). In tabel 5.2 zijn de resultaten van deze modellen samengevat, waarbij alleen de gestandaardiseerde effecten (bèta's) van de condities worden gepresenteerd. De volledige parameterschattingen zijn opgenomen in bijlage 2 (zie de tabellen B2.1 t/m B2.8).

Tabel 5.2 - Samenvatting resultaten fixed effects (LSDV) modellen

Groeiscore	Korte termijn Eind groep 5 – midden groep 6			Middellange termijn Eind groep 5 – eind groep 6		
	Bèta	sig	N	Bèta	sig	N
Begrijpend lezen			163			163
Interventie	-0,012			-0,052		
Interventie lezen	-0,007			-0,024		
Interventie rekenen	-0,016			-0,081		
	Bèta	sig	N	Bèta	sig	N
Rekenen-Wiskunde			129			112
Interventie	0,190	**		0,058		
Interventie lezen	0,048			-0,118		
Interventie rekenen	0,316	***		0,227	**	

*** p < 0,01; ** p < 0,05; * p < 0,1

NB. In alle modellen is gecontroleerd voor de score op de voormeting (eind groep 5)

In tabel 5.2 wordt duidelijk dat op het domein van begrijpend lezen de leerlingen in de verschillende condities een vergelijkbare leerwinst op de korte en middellange termijn doormaken. Er zijn immers geen significante verschillen in groei gevonden tussen leerlingen in de interventie- en controleconditie. Ook wanneer we onderscheid maken naar het type interventie (lezen of rekenen) dat de leerlingen hebben gevolgd, vinden we geen significant verschil in groei op de korte en middellange termijn. De intraklassecorrelatie (ICC) bedraagt 0,096 voor de korte termijn en 0,054 voor de middellange termijn. Dit betekent dat respectievelijk 9,6% van de variantie op de korte termijn en 5,4% van de variantie op de middellange termijn wordt verklaard door verschillen tussen scholen.

Voor de vaardigheid rekenen-wiskunde vinden we daarentegen wel significante verschillen in groei tussen leerlingen in de verschillende condities. Leerlingen in de interventieconditie maken op de korte termijn een significant sterkere groei in rekenvaardigheid door dan leerlingen in de controleconditie. Wanneer we vervolgens uitsplitsen naar het type interventie, dan blijkt dat dit verschil in groei tot stand komt door de leerlingen die de rekeninterventie hebben gevolgd. Zij vertonen een sterkere leerwinst op het domein rekenen-wiskunde, terwijl dit niet geldt voor leerlingen die aan de leesinterventie hebben deelgenomen. De intraklassecorrelatie (ICC) bedraagt 0,161 voor de korte termijn en 0,218 voor de middellange termijn. Dit betekent dat respectievelijk 16,1% van de variantie op de korte termijn en 21,8% van de variantie op de middellange termijn wordt verklaard door verschillen tussen scholen.

Ook op de middellange termijn is er een verschil in groei op het domein van rekenen-wiskunde tussen de leerlingen in de interventie- en controleconditie. Wederom geldt dat leerlingen die aan de rekeninterventie hebben deelgenomen, een sterkere leerwinst op dit domein kennen dan leerlingen die niet aan een interventie hebben deelgenomen.

Om de robuustheid van de gevonden regressie-effecten te checken, hebben we de hierboven gepresenteerde modellen nogmaals geschat, maar nu na statistische controle voor enkele leerlingkenmerken (geslacht, leeftijd in maanden, gesproken thuistaal en leerpotentiescore) om na te

gaan of de gevonden interventie-effecten (deels) vertekend worden door de invloed van deze leerlingkenmerken. Dit valt op voorhand overigens niet te verwachten gezien de random verdeling van leerlingen over de interventie- en controleconditie. Uit deze modellen komt dat inderdaad ook naar voren (zie de modellen 2 van bijlage 2 voor een weergave van de volledige resultaten). Alle gevonden interventie-effecten worden weliswaar iets kleiner, maar blijven significant na opname van controlevariabelen (model 2, zie bijlage 2).

Om de robuustheid van de resultaten verder na te gaan hebben we naast de fixed effects modellen ook *random intercept* modellen geschat met leerlingen genest binnen scholen. Hoewel de assumpties van beide typen modellen enigszins verschillen en beide hun eigen voor- en nadelen kennen, lijken ze sterk op elkaar. De resultaten van de random intercept modellen laten dat ook zien. Ook dan vinden we exact dezelfde interventie-effecten als in het geval van de fixed effects modellen.

Tot slot zijn groeicurve modellen geschat waarbij we de drie meetmomenten (eind groep 5 – midden groep 6 – eind groep 6) als *repeated measures* hebben geanalyseerd. De groeicurve modellen zijn als een multiniveau analyse opgezet waarbij meetmomenten zijn genest binnen leerlingen en leerlingen binnen scholen. In deze drieniveau analyse zijn de interventieconditie (inclusief het interventietype), het meetmoment en de (*cross-level*) interactie tussen beide variabelen opgenomen. Voor de schattingen zijn de leerlingen meegenomen met een geldige vaardigheidsscore op de voormeting (eind groep 5) en op ten minste één van de vervolgmetingen (midden groep 6 en/of eind groep 6). In bijlage 2 zijn de resultaten van de groeicurve modellen gepresenteerd (zie de tabellen B2.9 en B2.10). We kiezen ervoor deze groeicurve modellen slechts als aanvullende analyse te hanteren, omdat deze *random slope* modellen (net als random intercept modellen) veronderstellen dat de random effecten (variatie tussen groepen) normaal verdeeld zijn. Deze assumptie is discutabel bij N=8 scholen.

De resultaten van de groeicurve modellen laten zien dat ook deze wijze van analyseren nagenoeg dezelfde resultaten laat zien. In deze modellen worden namelijk ook geen verschillen in groei op het gebied van begrijpend lezen tussen leerlingen in de controle- en interventiecondities gevonden. In de groeicurve modellen die de ontwikkeling op de vaardigheid rekenen-wiskunde in kaart brengen vinden we – in overeenstemming met de hiervoor besproken analyses – wél een significant verschil in groei tussen de leerlingen in de interventieconditie (ongeacht type interventie – lezen of rekenen) en de leerlingen in de controleconditie op de korte termijn: de interactie *meting groep 6 * interventie* is positief en significant. Dit betekent dat leerlingen in de interventieconditie significant sterker groeien op de vaardigheid rekenen-wiskunde dan leerlingen in de controleconditie. Op middellange termijn wordt er ook in het groeicurve model geen significant verschil in groei op de vaardigheid rekenen-wiskunde gevonden tussen leerlingen in de interventie- en de controleconditie. Maken we onderscheid naar het type interventie dat de leerlingen in de interventieconditie hebben gevolgd, dan zien we dat het korte termijn effect van de interventie tot stand komt door de leerlingen die de rekeninterventie hebben gevolgd: zij laten op de korte termijn een significant sterkere groei zien dan leerlingen in de controleconditie. In tegenstelling tot de analyse in tabel 5.2 vinden we in het groeicurve model geen significant verschil in ontwikkeling op de middellange termijn tussen leerlingen die de rekeninterventie volgden en de leerlingen in de controleconditie. De richting van het (niet-significante) regressie-effect komt echter wel overeen met het gevonden effect in tabel 5.2. Dit verschil in bevindingen wordt mogelijk verklaard doordat de

geanalyseerde populatie in de groeicurve modellen iets afwijkt van die in de fixed effects regressiemodellen²⁰.

5.2 Invloed van implementatiekwaliteit en pedagogisch-didactische kwaliteit

In de voorgaande paragraaf hebben we het indicatieve effect van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' geanalyseerd.²¹ In deze paragraaf gaan we na in hoeverre dit effect afhankelijk is van de implementatiekwaliteit van de interventie en de pedagogisch-didactische kwaliteit van de interventielessen. We beantwoorden daarmee onderzoeksvragen 2 en 3:

2. *In hoeverre zijn de effecten van de interventie afhankelijk van de implementatiekwaliteit, c.q. de mate waarin het programma wordt uitgevoerd zoals bedoeld?*
3. *In hoeverre zijn de effecten van de interventie afhankelijk van de pedagogisch-didactische kwaliteit van het interactieproces tussen de (interventie)leerkrachten en de leerlingen?*

We onderzoeken de invloed van de implementatiekwaliteit en de pedagogisch-didactische kwaliteit op de omvang van de interventie-effecten door opnieuw fixed effects regressiemodellen te schatten. Ditmaal wordt de analyse enkel uitgevoerd op de groep van leerlingen die aan een interventie (ongeacht type) hebben deelgenomen ($n=87^{22}$). Omdat variatie in de implementatiekwaliteit en pedagogisch-didactische kwaliteit gemeten is op het niveau van het interventiegroepje en we *bias* door alle andere mogelijke verschillen tussen de interventiegroepjes willen uitsluiten, wordt in de modellen rekening hiermee gehouden door een dummyvariabele voor ieder interventiegroepje op te nemen. In afzonderlijke²³ modellen wordt de invloed van de implementatiekwaliteit en pedagogisch-didactische kwaliteit op de groei van leerlingen geschat. In deze modellen is ook weer de score op de voormeting (eind groep 5) opgenomen. In tabel 5.3 worden de resultaten van de geschatte fixed effects regressiemodellen samengevat door de gestandaardiseerde effecten (bèta's) te presenteren.

²⁰ In de fixed effects regressiemodellen wordt de groeiscore geanalyseerd waardoor leerlingen op beide meetmomenten een geldige score dienen te hebben om meegenomen te worden in de analyse. In de groeicurve modellen geldt deze voorwaarde niet: dit model mag ongebalanceerd zijn waarbij individuen niet in elk tijdvak even vaak of volledig zijn geobserveerd.

²¹ Zie de kanttekening in paragraaf 5.1 over de beperkte statistische power van deze analyses ten gevolge van de beperkte omvang van de onderzoeksgroep.

²² De N in de fixed effects regressiemodellen varieert vanwege het ontbreken van LVS-scores of scores op de onafhankelijke variabelen (zie ook hoofdstuk 4).

²³ Vanwege de geringe statistische power analyseren we de implementatiekwaliteit en pedagogisch-didactische kwaliteit in separate modellen.

Tabel 5.3 - Samenvatting resultaten fixed effects (LSDV) modellen

Groeiscore	Begrijpend lezen						Rekenen-Wiskunde					
	Eind 5 – midden 6 ^a			Eind 5 – eind 6 ^b			Eind 5 – midden 6 ^c			Eind 5 – eind 6 ^d		
	Bèta	sig	N	Bèta	sig	N	Bèta	sig	N	Bèta	sig	N
<i>Implementatiekwaliteit (gecentreerd gemiddelde)</i>	0,008		81	0,064		81	-0,175		65	0,018		60
<i>Pedagogisch-didactische kwaliteit (gecentreerd gemiddelde)</i>	0,896	***	86	0,699	***	85	0,610	**	70	0,497	*	62

^a ICC=0,116; ^b ICC=0,107; ^c ICC=0,127; ^d ICC=0,079

*** p < 0,01; ** p < 0,05; * p < 0,1

NB. In alle modellen is gecontroleerd voor de score op de voormeting (eind groep 5)

In tabel 5.3 wordt duidelijk dat de *implementatiekwaliteit* van de interventie niet van invloed is op de groei van leerlingen op begrijpend lezen en rekenen-wiskunde op de korte en middellange termijn. De *pedagogisch-didactische kwaliteit* van (het interactieproces in) de interventie heeft daarentegen wel een significante invloed op de groei die interventiedeelnemers doormaken bij begrijpend lezen en rekenen-wiskunde, op zowel de korte als de middellange termijn. Gesteld kan worden dat de groeiscore op beide domeinen op de korte en middellange termijn groter is naarmate de pedagogisch-didactische kwaliteit van het interactieproces tussen de interventie-leerkrachten (of coaches) en de leerlingen hoger is. Opname van relevante leerlingkenmerken in bovengenoemde modellen bleek niet mogelijk te zijn vanwege multicollineariteit als gevolg van het beperkte aantal vrijheidsgraden dat (nog) beschikbaar was na opname van de dummies voor ieder interventiegroepje.

Vergelijkbare effecten als bij de fixed effects regressiemodellen vinden we door het schatten van random intercept modellen, zij het dat deze alleen significant zijn in de modellen waar de invloed van de pedagogisch-didactische kwaliteit op de groei in rekenen-wiskunde wordt geschat. De standaardfouten van de parameterschattingen zijn bij de random intercept modellen groter dan bij de fixed effects modellen, waardoor de effecten minder snel significant zijn. Dit komt waarschijnlijk door het feit dat er (te) weinig variantie is binnen leerlingen tussen interventiegroepjes. De omvang hiervan varieert van drie tot acht leerlingen.

Tot slot zijn ook weer groeicurve modellen geschat. In deze modellen blijft enkel het effect van de pedagogisch-didactische kwaliteit op de korte termijn groei in rekenen-wiskunde overeind.

6 Conclusie en discussie

6.1 Uitgevoerd onderzoek

In de voorafgaande hoofdstukken is het onderzoek beschreven dat we hebben uitgevoerd naar de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' die in schooljaar 2023-2024 is geïmplementeerd bij leerlingen in groep 6 van Amsterdamse basisscholen met een hoge schoolweging. De interventie betreft een vorm van leertijduitbreiding: de deelnemers kregen tien weken lang twee keer per week in groepjes van maximaal acht leerlingen extra lessen in begrijpend lezen of rekenen-wiskunde aangeboden. De lessen vonden plaats buiten het klaslokaal, tijdens of na schooltijd, en werden verzorgd door een interne of externe coach die daarbij gebruikmaakte van het lesprogramma dat ontwikkeld is door Beuker Onderwijs.

Het in dit rapport beschreven onderzoek bestaat uit een kwantitatieve evaluatie volgens RCT-design van het effect van de interventie op de leergroei. De leergroei is vastgesteld op basis van de scores die de leerlingen eind groep 5, halverwege groep 6 en eind groep 6 hebben behaald op de Cito-LVS toetsen Begrijpend lezen en Rekenen-Wiskunde. Ook is onderzocht wat de invloed op de effectiviteit van de interventie is van de implementatiekwaliteit en pedagogisch-didactische kwaliteit van (het interactieproces tijdens) de interventie. De implementatiekwaliteit is vastgesteld met een logboek dat de coaches invulden over de uitgevoerde interventielessen, de pedagogisch-didactische kwaliteit van het interactieproces is gemeten aan de hand van lesobservaties uitgevoerd door onderzoeksassistenten.

Naast het onderzoek dat in dit rapport is beschreven, is een aanvullend onderzoek uitgevoerd, met extra controlegroepen, om het inzicht in het effect van de interventie verder te vergroten. De resultaten daarvan zijn beschreven in een addendum bij dit rapport.

Beperkingen onderzoek

Een belangrijke kanttekening bij het uitgevoerde onderzoek is dat de interventieontwikkelaar kort voor de start van de interventie op de scholen besloot de samenwerking op te zeggen met de organisatie die de externe coaches zou leveren. Het gevolg daarvan was dat van de negentien aangemelde interventiescholen er elf niet langer konden/wilden deelnemen aan de interventie en het interventieonderzoek, omdat er geen coach beschikbaar was om de interventielessen te verzorgen. Met het aantal deelnemende scholen nam ook het aantal deelnemende leerlingen af, wat (mogelijk) negatieve gevolgen heeft gehad voor (de implementatiekwaliteit van de interventie en) de statistische power van de analyse en ervoor zorgde dat de kans om een reëel programma-effect te detecteren sterk afnam. Deze beperking leidt ertoe dat de onderzoeksresultaten met de nodige voorzichtigheid moeten worden geïnterpreteerd. Ofschoon het onderzoek is opgezet als een RCT-studie om de effecten van de interventie te kunnen vaststellen, kunnen gezien de beperkte grootte van de onderzoeksgroep geen harde conclusies getrokken worden over de werking van de interventie.

Een tweede beperking van het onderzoek, zijdelings samenhangend met het voorafgaande, is dat de respons op de ouder- en leerkrachtvragenlijst dermate laag was dat één van de geformuleerde onderzoeksvragen, naar de invloed van de ouder- en leerkrachtbetrokkenheid op de effectiviteit van de interventie, niet kon worden beantwoord.

6.2 Beantwoording van de onderzoeksvragen

Hierna beantwoorden we de geformuleerde onderzoeksvragen, onder het bovengenoemde voorbehoud dat slechts voorzichtige conclusies kunnen worden getrokken gezien de lage aantallen deelnemers en de beperkte statistische power van de analyses. De uitspraken die we hieronder doen over (niet) gevonden effecten moeten in dat licht worden gezien.

1. *Wat is het effect van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' op de leerachterstand van de deelnemende leerlingen ten opzichte van andere leerlingen in dezelfde jaargroep?*

Uit de onderzoeksresultaten blijkt dat de interventiedeelnemers op de vaardigheid begrijpend lezen een vergelijkbare groei laten zien als de leerlingen in de controleconditie. Er zijn, met andere woorden, *geen verschillen in de groei in leesvaardigheid* van beide groepen gevonden; niet direct na de interventie (medio groep 6) en ook niet enkele maanden later (eind groep 6) is. Ook als we rekening houden met het type interventie (lezen of rekenen) dat de interventiedeelnemers hebben gevolgd, vinden we geen verschillen in leesgroei op de korte en middellange termijn.

Voor de vaardigheid rekenen-wiskunde zijn daarentegen *wel verschillen in groei* geconstateerd tussen de leerlingen in de verschillende condities. De interventiedeelnemers hebben op de korte termijn een sterkere groei in rekenvaardigheid doorgemaakt dan de leerlingen in de controleconditie. Dit verschil blijkt tot stand te komen door de leerlingen die de rekeninterventie hebben gevolgd. Zij vertonen een sterkere groei op het domein rekenen-wiskunde, terwijl dit niet geldt voor leerlingen die aan de leesinterventie hebben deelgenomen. Ook op de middellange termijn geldt dat leerlingen die aan de rekeninterventie hebben deelgenomen een sterkere leerwinst op dit domein kennen dan leerlingen die niet aan een interventie hebben deelgenomen.

2. *In hoeverre zijn de effecten van de interventie afhankelijk van de implementatiekwaliteit, c.q. de mate waarin het programma wordt uitgevoerd zoals bedoeld?*

Uit de resultaten van onze analyses blijkt dat de implementatiekwaliteit van de interventie – gemeten met de logboeken die de coaches zelf ingevuld hebben over de uitgevoerde interventielessen – *niet* van invloed is op de groei van de lees- of rekenvaardigheid van de deelnemers op de korte of de middellange termijn.

3. *In hoeverre zijn de effecten van de interventie afhankelijk van de pedagogisch-didactische kwaliteit van het interactieproces tussen de interventieleerkracht (coach) en de leerlingen?*

De pedagogisch-didactische kwaliteit van het interactieproces tijdens de interventie – gemeten aan de hand van lesobservaties – heeft volgens de onderzoeksresultaten *wel een invloed* op de groei die interventiedeelnemers doormaken bij begrijpend lezen en rekenen-wiskunde, zowel op de korte als de middellange termijn. De groeiscore op beide domeinen is op de korte en de middellange termijn groter naarmate de pedagogisch-didactische kwaliteit van het interactieproces tussen de coach en leerlingen hoger is.

4. *In hoeverre zijn de effecten van de interventie afhankelijk van de mate waarin de reguliere groepsleerkrachten en ouders bij de interventie worden betrokken?*

De onderzoeksgegevens waren ontoereikend om de invloed van de ouder- en leerkrachtbetrokkenheid op de effectiviteit van de interventie te kunnen vaststellen en deze vraag te beantwoorden. Dit is het gevolg van de hoge non-respons op de vragenlijsten voor de ouders en groepsleerkrachten van de interventiedeelnemers.

6.3 Reflectie en aanbevelingen

Interpretatie bevindingen

Op basis van onze RCT-studie concluderen we dat de deelname aan de extra lessen rekenen-wiskunde in het kader van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen' samengaat met een groei in rekenvaardigheid van de betreffende leerlingen in vergelijking tot de controlegroep, die enige maanden later nog steeds zichtbaar is. De leesvaardigheid van de deelnemers aan de rekeninterventie is niet sterker toegenomen dan die van de controlegroep; voor zover de leergroei kan worden toegeschreven aan de interventie is het effect daarvan dus domeinspecifiek. Een vergelijkbare relatie vinden we niet voor de deelnemers aan de extra lessen begrijpend lezen in het kader van de interventie; hun leerwinst bij lezen en rekenen is niet groter dan van leerlingen in de controleconditie.

Ofschoon het gevonden effect van de interventie dus beperkt is en bovendien (vanwege de kleine onderzoeksgroep) niet met grote zekerheid kan worden vastgesteld, kan deze – mede gezien het lage aantal deelnemende scholen en leerlingen in het interventieonderzoek – toch ook verrassend genoemd worden. Dat uitsluitend voor rekenen-wiskunde een indicatief effect van de interventie is gevonden, valt mogelijk te verklaren uit het feit dat dit meer een schoolse vaardigheid is – c.q. een vaardigheid die vooral op school wordt aangeleerd – terwijl begrijpend lezen een complexere vaardigheid is die ook deels thuis wordt geleerd. Immers, begrijpend lezen hangt sterk samen met andere vaardigheden, zoals de woordenschat, achtergrondkennis en leescultuur thuis (Heister-Swart, 2018). Ook uit ander onderzoek blijkt dat interventies op rekenvaardigheid vaak hogere effectgroottes laten zien dan leesgerelateerde interventies (zie bijvoorbeeld de meta-analyse van Wanzek et al., 2021). Uiteraard spelen daarbij ook andere aspecten een rol, zoals de kwaliteit van de interventie en van de uitvoerder, en de intensiteit, groepsgrootte en selectie van deelnemers (zie ook onze literatuurstudie zoals beschreven in hoofdstuk 2).

Een tweede bevinding in ons onderzoek is dat de kwaliteit van het pedagogisch-didactisch interactieproces tijdens de interventielessen van invloed is op de groei die interventiedeelnemers doormaken bij begrijpend lezen en rekenen-wiskunde, zowel op de korte als de middellange termijn. Een vergelijkbaar modererend effect vinden we niet voor de implementatiekwaliteit van de interventie. Voor de interpretatie van deze bevindingen zijn diverse relevante punten aan te wijzen. Ten eerste is in ons onderzoek de implementatiekwaliteit gemeten aan de hand van logboekdata die de coaches zelf hebben aangeleverd. In feite was er dus sprake van 'slagers die hun eigen vlees keuren' en wellicht moeten we concluderen dat met deze zelfbeoordelingen de werkelijke implementatiekwaliteit inclusief eventuele variatie daarin tussen de coaches niet goed in beeld kon worden gebracht. Zo blijken de externe coaches (die lesgeven aan negen van de negentien interventiegroepjes) gemiddeld genomen een significant lager cijfer te hebben gegeven aan de door hen uitgevoerde interventielessen dan de interne coaches, ofschoon eerder het

omgekeerde verwacht mag worden; de externe coaches hebben immers zelf het interventiemateriaal ontwikkeld. De pedagogisch-didactische kwaliteit is daarentegen door objectieve externe beoordelaars vastgesteld, en daarmee wellicht betrouwbaarder gemeten. Het zal dan ook niet verbazen dat de pedagogisch-didactische kwaliteit van het interactieproces tijdens de lessen van de externe coaches gemiddeld genomen gunstiger is beoordeeld dan tijdens de lessen van de interne coaches. Aanvullend merken we op dat – ten gevolge van het personeelsgebrek op de betreffende scholen – niet alle interne coaches gekwalificeerde leerkrachten waren, terwijl het materiaal van Beuker wel met die doelgroep voor ogen is ontwikkeld.

Anderzijds is het ook weer niet zo dat alleen de externe coaches hoog scoren op pedagogisch-didactisch gedrag; er zijn ook interne coaches die daaraan bijdragen. Dat brengt ons op een andere mogelijke reden waarom de pedagogisch-didactische kwaliteit wél, maar de implementatiekwaliteit niet van invloed is, namelijk dat het interventiemateriaal in principe weinig vrijheid biedt aan de coaches om deze op verschillende manieren te gebruiken. Volgens de ontwikkelaars zijn de lessen immers zodanig uitgeschreven dat de coaches zich inhoudelijk nauwelijks hoeven voor te bereiden. Het is dan wellicht logisch dat de pedagogisch-didactische aanpak van de coach – intern of extern – vooral ‘het verschil maakt’.

Of onze interpretaties houtsnijden moet in vervolgonderzoek verder worden uitgezocht, zeker gezien het lage aantal deelnemers aan de interventie en het onderzoek – een gevolg van het op het laatste moment afhaken van reeds geworven scholen vanwege onvoldoende beschikbare gekwalificeerde coaches. Vast staat in ieder geval dat de gevonden invloed van de pedagogisch-didactische kwaliteit op de effectiviteit van de interventie aansluit op de elders al vaker gerapporteerde bevinding dat de kwaliteit van de interventie-uitvoerders een zeer belangrijke randvoorwaarde is voor effectiviteit van een VLT-interventie. Hoewel we het niet op basis van onze eigen resultaten kunnen staven, valt niet uit te sluiten dat onder het concept ‘pedagogisch-didactische kwaliteit’ nog andere relevante aspecten schuilgaan, zoals de mate waarin de coaches bekend zijn met het interventiemateriaal of op school gefaciliteerd worden bij het aanbieden van de twintig interventielessen.²⁴

Sterke en zwakke punten

Sterke punten van ons interventieonderzoek zijn de volgende:

- Door de bestaande contacten van de interventieontwikkelaar met twee schoolbesturen in Amsterdam, verliep de werving van scholen eenvoudig. De schoolbesturen moedigden hun scholen aan deel te nemen aan de interventie en het onderzoek, en dat leidde tot voldoende aanmeldingen van scholen die qua leerlingenpopulatie zeer geschikt waren.
- De randomisatieprocedure verliep (over het algemeen) naar wens en had een succesvolle uitkomst: tussen de leerlingen die random toegewezen waren aan de experimentele- of de controleconditie blijken geen verschillen in startniveau en leerpotentie, en nauwelijks verschillen in achtergrondkenmerken te bestaan.
- Alle scholen werkten al jaren met de toetsen Begrijpend lezen en Rekenen-Wiskunde van Cito-LVS. Deze toetsen zijn goed gevalideerd en betrouwbaar; de leerwinst van de leerlingen kon zodoende ook betrouwbaar en valide worden bepaald.
- De pedagogisch-didactische kwaliteit van het interactieproces tijdens de interventielessen is gemeten aan de hand een observatieschema en de lesobservaties zijn uitgevoerd

²⁴ Het moment waarop de interventielessen worden gegeven (tijdens of na schooltijd), blijkt overigens niet significant van invloed te zijn op de vaardigheidsgroei van leerlingen.

door getrainde en objectieve onderzoeksassistenten. Vastgesteld is dat dit betrouwbare data heeft opgeleverd.

- Er is gebruikgemaakt van geavanceerde statistische methoden en verschillende analysemodellen (met ieder zijn eigen voor- en nadelen) die grofweg alle dezelfde uitkomsten gaven. Er zijn *fixed effects least squares dummy variable* regressiemodellen geschat. De *random intercept* modellen gaven dezelfde uitkomsten. Om de robuustheid van de resultaten nader te checken, zijn ook groeicurve modellen geschat waarbij de drie meetmomenten als *repeated measures* zijn geanalyseerd. Verder zijn er diverse analyses uitgevoerd ter verificatie van het succes van de randomisatie en om na te gaan in hoeverre de experimentele uitval random was.

Zwakkere punten aan ons interventieonderzoek zijn de volgende:

- Uiteindelijk is het onderzoek uitgevoerd met een veel lager aantal deelnemende scholen en leerlingen dan was voorzien, hetgeen van invloed was op de statistische power van de analyse en ertoe leidde dat de kans om een reëel programma-effect te detecteren sterk afnam. Analysebevindingen aangaande (niet) gevonden effecten moeten in dat licht worden gezien.
- Het digitale logboek, het instrument dat is ontwikkeld voor het meten van de implementatiekwaliteit, was bij nader inzien minder geschikt voor het doel. De te implementeren interventieactiviteiten verschilden nogal tussen de lessen, afhankelijk van de lesvariant en interventiefase, waardoor de geregistreeerde logboekdata zich lastig laten vergelijken. Bovendien vulden de coaches het logboek vermoedelijk niet na elke les in, terwijl dat wel de bedoeling was. Op basis van de resultaten kon de betrouwbaarheid en validiteit van het instrument niet goed worden bepaald. Een beter alternatief was wellicht een vragenlijst geweest over de implementatiegetrouwheid van de coaches, al zou dat wel meer 'hoog over' zijn geweest en eveneens neerkomen op zelfbeoordelingen.
- De respons op oudervragenlijst was te laag, mogelijk doordat veel van de betreffende ouders de Nederlandse taal onvoldoende machtig waren. Ook de leerkrachtvragenlijst werd onvoldoende ingevuld, ondanks herhaaldelijk rappelleren en de financiële compensatie voor de scholen. Onderzoeksvraag 4 kon daardoor niet beantwoord worden;
- Door de korte onderzoeksperiode konden met het onderzoek geen langetermijneffecten van de interventie in kaart gebracht worden.

Aanbevelingen

Vervolgonderzoek met een groter aantal deelnemende scholen en leerlingen is nodig om met meer zekerheid uitspraken te kunnen doen over de effectiviteit van de interventie 'Doelgericht aanbod verlengde leertijd in kleine groepen'. Tegelijkertijd is het opmerkelijk dat ondanks het kleine aantal deelnemers in ons onderzoek toch een mogelijk interventie-effect is gevonden. Het lijkt te suggereren dat als scholen deze interventie aanbieden aan leerlingen in groep 6 die daarvoor in aanmerking komen (hun LVS-scores verwijzen naar vaardigheidsniveau C, D of hoge E), dit gunstig uitpakt voor het verkleinen van de leerachterstand van deze leerlingen, in het bijzonder op het gebied van rekenen. Of dat ook geldt voor leerlingen in andere leerjaren is niet onderzocht. Ook is onbekend hoe lang de effectiviteit van de interventie in de praktijk aanhoudt. In ons onderzoek hebben we door de beperkte tijd die er beschikbaar was, niet verder kunnen kijken dan enkele maanden. Essentieel is in elk geval dat – mochten scholen kiezen voor deze interventie – interventie gekwalificeerde leerkrachten worden ingezet die pedagogisch-didactisch gezien goed onderlegd zijn. Gezien het huidige gebrek aan onderwijspersoneel, zeker op scholen met een complexe leerlingenpopulatie, is dit momenteel geen eenvoudige opgave.

Voor een eventueel vervolgonderzoek met RCT-design naar het effect van de interventie raden we af om de randomisatieprocedure zo in te richten dat leerlingen *binnen* een klas worden toegewezen aan de experimentele of controleconditie. Voor leerkrachten is het frustrerend dat het lot uitwijst wie van de in aanmerking komende klasgenoten wel en wie niet mag deelnemen aan de interventie; hetzelfde geldt vaak ook voor de leerlingen zelf en voor hun ouders. Beter zou het zijn parallelklassen of – liever nog – scholen toe te wijzen aan de verschillende condities. Daarnaast raden we aan een eventueel vervolgonderzoek pas te starten als zeker is dat er voldoende gekwalificeerde coaches beschikbaar zijn om de interventie aan te bieden aan een minimaal vereist aantal scholen en leerlingen, zodat het onderzoek kan worden verricht met meer statistische power en meer mogelijkheden voor het analyseren van andere effecten en condities. Een laatste aanbeveling is dat een betrouwbaar instrument voor het meten van de implementatiekwaliteit dient te worden ontwikkeld.

Literatuur

- Appelhof, P. (2009). *Een oriëntatie naar verlengde onderwijstijd; inrichting en effecten*. Utrecht: Oberon.
- Beekhoven, S., Hoogeveen, K., Kruiter, J. & Vander Heyden, K. (2017). *Schakelklassen, zomerscholen en onderwijstijdverlenging. Een inventariserend onderzoek onder gemeenten en scholen*. Utrecht: Sardes
- Bluemink, Ch., Jenniskens, T., Langen, A. van, Leest, B., Wolbers, M. (2022). *Onderpresteren in het Nederlandse basisonderwijs anno 2021*. Nijmegen: KBA Nijmegen.
- Centraal Plan Bureau (2018). *Bewezen (in)effectieve maatregelen tegen leerachterstanden in het primair onderwijs. Achtergronddocument*. Den Haag: Centraal Plan Bureau.
- Centraal Bureau voor de Statistiek. (2024, 17 oktober). *Onderzoek herijking risico-indicator onderwijsachterstanden – fase 1*. [<https://www.cbs.nl/nl-nl/longread/aanvullende-statistische-diensten/2024/onderzoek-herijking-risico-indicator-onderwijsachterstanden-fase-1>]
- Cook, P. J. et al. (2015). Not too late: Improving academic outcomes for disadvantaged youth. *Northwestern University Institute for Policy Research Working Paper*, (15-01).
- Driessen, G., Claassen, A., Smit, F. (2010). *Variatie in schooltijden. Een internationale literatuurstudie naar de effecten van verschillende invullingen van de schooldag, de schoolweek en het schooljaar*. Nijmegen: ITS.
- Educational Endowment Foundation (2018). *Teaching and Learning Toolkit*. [<https://educationendowmentfoundation.org.uk/>]
- Jenniskens, T., Langen, A. van, Stronkhorst, E., Veen, I. van der, Veen, A., Wolbers, M. (2022). Meting leerpotentie in groep 6. Technisch rapport. KBA Nijmegen/Kohnstamm Instituut.
- Haelermans, C., J. Ghysels, & M. Monfrance (2018). *Effectanalyse lente- en zomerscholen 2017. TIER Research Report*. Maastricht University, Maart 2018.
- Haelermans, C., van der Velden, R., Aarts, B., Bijlsma, I., Huijgen, T., Jacobs, M., van Vugt, L., & van Wetten, S. (2021). *Door eerste schoolsluiting minder leergroei, vooral bij begrijpend lezen*. ROA External Reports NCO Factsheet No. 1
- Haelermans, C., Van Vugt, L., Van Wetten, S., Baumann, S., Havermans, W., Hendrikse, A., Meijer, R. & Smeets, C. (2023). *Balans na drie jaar COVID-19: grootste vertraging leergroei voor rekenen-wiskunde, spelling en begrijpend lezen schommelen*. NCO Factsheet No. 2023-1.
- Heister-Swart, N.M. (2018): *Lexical quality effects in reading comprehension: A developmental perspective*, Nijmegen. Radboud Universiteit.
- Jez, S. J., & Wassmer, R. W. (2015). The impact of learning time on academic achievement. *Education and Urban Society*, 47(3), 284-306.
- Kennisrotonde (2019). *Draagt een verlengde schooldag bij aan de ontwikkeling van leerlingen?* [<https://www.kennisrotonde.nl/vraag-en-antwoord/effectiviteit-leertijdverlenging>]
- Kidron, Y., & Lindsay, J. (2014). *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia.
- Kline, P. (1999). *Handbook of Psychological Testing*. London: Routledge
- Kortekaas-Rijlaarsdam, A.F., Turkeli, R., de Vries, N., Ehren, M.C.M., Meeter, M. (2020). *Effectiviteit van inhaal- en ondersteuningsprogramma's om onderwijsachterstanden in te halen*. Amsterdam: VU.
- Kuiken, F. (2016). *Werkzame bestanddelen van onderwijstijdverlenging*. Amsterdam: UVA.

- Lord, F.M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 68, 304-305.
- Meyer, E., & Van Klaveren, C. (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands. *Economics of Education Review*, 36, 1-11.
- Mulder, L., Hoeven, A. van der, Vierke, H., Ledoux, G., Veen, I. van der, Oud, W., Daalen, M. van & Roeleveld, R. (2008). *Inrichting en effecten van schakelklassen. Resultaten van het evaluatieonderzoek schakelklassen in het schooljaar 2006/2007*. Nijmegen/Amsterdam: ITS, Radboud Universiteit /SCO-Kohnstamm Instituut.
- Mulder, L., Hoeven, A. van der, Vierke, H., Veen, I. van der & Elshof, D. (2009). *Inrichting en effecten van schakelklassen*. Nijmegen/Amsterdam: ITS Radboud Universiteit/SCO-Kohnstamm Instituut.
- Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985–2009). *Review of educational research*, 80(3), 401-436.
- Rijksoverheid. (2021). Leervertraging en zorgen over welzijn en corona. [<https://www.rijksoverheid.nl/actueel/nieuws/2021/10/28/leervertragingen-en-zorgen-over-welzijn-door-corona>]
- Slaap, C. & Kuiken, F (2013). *Evaluatieonderzoek Vakantieschool Taal 2012-2013*. Amsterdam: UVA.
- Slavin, R. E., Lake, C., & Groff, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, 79(2), 839-911.
- Smit, F., Driessen, G., & van Kuijk, J. (2015). Variatie in schooltijd en onderwijskwaliteit. Een internationale literatuurstudie naar effecten van verschillende invullingen van de onderwijstijd op de onderwijskwaliteit. *Nijmegen: ITS*.
- Vander Heyden, D. K. (2021). *Onderwijskansen voor het jonge kind*. [https://www.onderwijskennis.nl/sites/onderwijskennis/files/media-files/Thema_overzicht%20Jonge%20kind_DEF%20-%20hz.pdf]
- Van der Vegt, D. A. L. (2021). Ouderbetrokkenheid en onderwijskansen. *Onderwijskansen in het primair onderwijs*.
- Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2021). Targeted school-based interventions for improving reading and mathematics for students with, or at risk of, academic difficulties in Grades 7–12: A systematic review. *Review of Educational Research*, 91(2), 261–301. [<https://doi.org/10.3102/0034654320975583>]
- Wedel, K. (2021). Instruction time and student achievement: The moderating role of teacher qualifications. *Economics of Education Review*, 85, Article 102183. <https://doi.org/10.1016/j.econedurev.2021.102183>.

Bijlage 1 Ontwikkelde onderzoeksinstrumenten

Leerlinginformatieformulier 1 (excl. toelichting)

Invulformulier groep 6 2023-2024 Interventie verlengde leertijd									
Schoolnaam									
SCHOOL ID									
Naam groep 6									
Leerlingnummer	Geboortedatum		Geslacht (1=jongen, 2=meisje, 3=anders)	Thuis taal Nederlands (1=ja, 2=nee, 3=overig)	Cito LVS Begrijpend lezen juni 2023		Cito LVS Rekenen-wiskunde juni 2023		Reden uitsluiting 1 - 6 (zie toelichting)
	MAAND	JAAR			Vaardigheidsscore	Toetsversie	Vaardigheidsscore	Toetsversie	

Invulformulier juni 2024 Groep 6 Onderzoek Interventie verlengde leertijd																
Schoolnaam	SCHOOL ID		NAAM GROEP													
Leerlingnummer	Geboortedatum		MAAND	JAAR	Geslacht (1=jongens, 2=meisjes, 3=anders)	Cito LVS Begrijpend lezen jan/feb 2024		Cito LVS Begrijpend lezen mei/juni 2024		Cito LVS Rekenen-wiskunde jan/feb 2024		Cito LVS Rekenen-wiskunde mei/juni 2024		Bijzonderheden leesontwikkeling	Bijzonderheden rekenontwikkeling	Bijzonderheden algemeen
						Vaardigheidsscore	Toetsversie	Vaardigheidsscore	Toetsversie	Vaardigheidsscore	Toetsversie					

Logboek interventielessen (versie begrijpend lezen; vragen die in analyses zijn meegenomen)

Onderdelen	Scores
A. Lesonderdelen behandeld conform de handleiding: Introductie	
Terugblik	Nee=0 / Ja=1
Voorkennis activeren	Nee=0 / Ja=1
Lesdoelen	Nee=0 / Ja=1
Lesoverzicht	Nee=0 / Ja=1
B. Lesonderdelen behandeld conform de handleiding: Instructie en oefenen	
Woordenschatuitbreiding	Nee=0 / Ja=1
Uitleg: strategie	Nee=0 / Ja=1
Uitleg: vraag	Nee=0 / Ja=1
Oefenen	Nee=0 / Ja=1
C. Lesonderdelen behandeld conform de handleiding: Afronding	
Nabespreken	Nee=0 / Ja=1
Vooruitblik	Nee=0 / Ja=1
D. Lesdoelen	
Lesdoel strategie behaald	Geen=0 / Deel=1 / Alle=2D
Denkstappen verwoord	Geen=0 / Deel=1 / Alle=2D
E. Algemeen oordeel les	
Rapportcijfer uitvoering les	1 t/m 10

Logboek interventielessen (versie rekenen-wiskunde; vragen die in analyses zijn meegenomen)

Onderdelen	Scores
A. Lesonderdelen behandeld conform de handleiding: Introductie	
Terugblik	Nee=0 / Ja=1
Voorkennis activeren	Nee=0 / Ja=1
Lesdoelen	Nee=0 / Ja=1
Lesoverzicht	Nee=0 / Ja=1
B. Lesonderdelen behandeld conform de handleiding: Instructie en oefenen	
Woordenschat	Nee=0 / Ja=1
Uitleg: rekenonderwerp	Nee=0 / Ja=1
Oefenen: opdrachten	Nee=0 / Ja=1
Oefenen: woorden	Nee=0 / Ja=1
C. Lesonderdelen behandeld conform de handleiding: Afronding	
Nabespreken	Nee=0 / Ja=1
Vooruitblik	Nee=0 / Ja=1
D. Lesdoelen	
Lesdoel behaald	Geen=0 / Deel=1 / Alle=2D
Denkstappen verwoord	Geen=0 / Deel=1 / Alle=2D
E. Algemeen oordeel	
Rapportcijfer uitvoering les	1 t/m 10

Datum: _____ Schoolnaam: _____ Schoolnr. _____		Nr. interventiegroepje: _____	Vak: _____	Lesnummer: _____	A of B: _____
Coach: _____		Observator: _____	Eindtijd observatie: _____		
Onderdeel	De coach ...	0=nee 1=ja, matige kwaliteit 2=ja, goede kwaliteit 9=niet van toepassing	Aantekeningen/opmerkingen/toelichting		
Introductie	Blijkt terug op vorige interventies	0 – 1 – 2 – 9			
	Haalt voorkennis op	0 – 1 – 2 – 9			
	Verheldert het lesdoel (wat ga je leren, waarom?)	0 – 1 – 2 – 9			
	Vertelt hoe de les eruit zal zien (lesoverzicht)	0 – 1 – 2 – 9			
Instructie	Geeft duidelijke uitleg over nieuwe lesstof	0 – 1 – 2 – 9			
	Legt nieuwe/moeilijke woorden pakkend uit	0 – 1 – 2 – 9			
	Doet hardop voor hoe het moet ('modellering')	0 – 1 – 2 – 9			
	Checkt of leerlingen het begrijpen	0 – 1 – 2 – 9			
	Herhaalt uitleg indien nodig	0 – 1 – 2 – 9			
	Stelt interactieve vragen	0 – 1 – 2 – 9			
Verwerking	Differentieert op verschillende niveaus indien nodig	0 – 1 – 2 – 9			
	Laat leerlingen zelf oefenen aan de hand van opdrachten	0 – 1 – 2 – 9			
	Zorgt voor veel interactie en samenwerking tussen leerlingen	0 – 1 – 2 – 9			
	Geeft duidelijke uitleg voorafgaand aan een opdracht	0 – 1 – 2 – 9			
	Checkt of de leerlingen de opdracht begrijpen	0 – 1 – 2 – 9			
	Begeleidt/ondersteunt de leerlingen tijdens de opdrachten	0 – 1 – 2 – 9			
	Geeft feedback op de uitgevoerde opdrachten	0 – 1 – 2 – 9			

Onderdeel	De coach ...	0=nee 1=ja, matige kwaliteit 2=ja, goede kwaliteit 9=niet van toepassing	Aantekeningen/opmerkingen/toelichting
Pedagogisch aanpak en interactie	Signaleert probleemgedrag en stuurt dit tijdig bij	0 – 1 – 2 – 9	
	Stimuleert taakgericht gedrag	0 – 1 – 2 – 9	
	Heeft positieve interacties met leerlingen; plezierig klimaat	0 – 1 – 2 – 9	
	Zorgt dat leerlingen zich op hun gemak voelen	0 – 1 – 2 – 9	
	Anticipeert op de behoeften van leerlingen	0 – 1 – 2 – 9	
Terugblik en afsluiting	Evalueert het lesverloop tijdens de lesafsluiting	0 – 1 – 2 – 9	
	Komt terug op de lesdoelen	0 – 1 – 2 – 9	
	Blijkt vooruit op de volgende les	0 – 1 – 2 – 9	
	Heeft les logisch opgebouwd	0 – 1 – 2 – 9	
Betrokkenheid leerlingen bij de les		0=slecht 1=matig 2=goed 9=nvt	Aantekeningen/opmerkingen/toelichting
Leerling a		0 – 1 – 2 – 9	
Leerling b		0 – 1 – 2 – 9	
Leerling c		0 – 1 – 2 – 9	
Leerling d		0 – 1 – 2 – 9	
Leerling e		0 – 1 – 2 – 9	
Leerling f		0 – 1 – 2 – 9	
Leerling g		0 – 1 – 2 – 9	
Leerling h		0 – 1 – 2 – 9	

Vragenlijst voor ouders/verzorgers

Deze vragenlijst gaat over de extra lessen taal of rekenen die uw kind in groep 6 krijgt.

Vragen over de extra lessen

Zet bij elke vraag een rondje om het antwoord dat u wilt geven. Voorbeeld: Ja / Nee / Weet ik niet

Vindt u het prettig/fijn dat uw kind extra lessen krijgt?—————Ja / Nee / Weet ik niet

Bent u tevreden over het tijdstip (dag en tijd) van de extra lessen?—————Ja / Nee / Weet ik niet

Vertelt de school/leerkracht u wat uw kind leert bij de extra lessen?—————Ja / Nee / Weet ik niet

Vertelt de school/leerkracht u hoe het met uw kind gaat bij de extra lessen?——Ja / Nee / Weet ik niet

Praat u met uw kind over de extra lessen?—————Ja / Nee / Weet ik niet

Vindt uw kind de extra lessen leuk?—————Ja / Nee / Weet ik niet

Leert uw kind nieuwe dingen in de extra lessen?—————Ja / Nee / Weet ik niet

Werkt uw kind hard tijdens de extra lessen?—————Ja / Nee / Weet ik niet

Doet u thuis met uw kind de spelletjes uit de extra lessen?—————Ja / Nee / Weet ik niet

Gaat het op school/in de klas beter met uw kind door de extra lessen?—————Ja / Nee / Weet ik niet

Leest u thuis boekjes samen met uw kind?—————Ja / Nee / Weet ik niet

Doet u thuis rekenspelletjes samen met uw kind?—————Ja / Nee / Weet ik niet

Als u dat wilt, kunt u hieronder nog iets opschrijven over wat u vindt van de extra lessen:

Wie heeft deze vragenlijst ingevuld?

- ☐ Vader of verzorger
- ☐ Moeder of verzorgster
- ☐ Anders, namelijk

Dank u wel voor het invullen! Op de voorkant staat hoe u dit vragenlijstje kunt terugsturen.

Vragenlijst groepsleerkrachten (excl. toelichting)

A. De start van de extra lessen

- 1 Ik ben voldoende betrokken bij het besluit van mijn school om mee te doen aan de interventie/de extra lessen
- 2 Ik ben tevreden over de selectie van leerlingen uit mijn klas voor de extra lessen
- 3 Ik ben tevreden over de tijdstippen waarop de extra lessen plaatsvonden
- 4 Ik ben bij de start voldoende geïnformeerd over de inhoud en werkwijze van de extra lessen
- 5 Degene die de extra lessen gaf heeft van tevoren de specifieke onderwijsbehoeften van de deelnemende leerlingen uit mijn klas geïnventariseerd

Ruimte voor uw eventuele opmerkingen over de start van de interventie en de extra lessen

B. De kwaliteit en het verloop van de extra lessen

- 6 De inhoud van de extra lessen sloot goed aan bij de lesstof die in de klas werd behandeld
- 7 De extra lessen sloten goed aan bij de individuele onderwijsbehoeften van de deelnemende leerlingen uit mijn klas
- 8 Er werd in de extra lessen voldoende gedifferentieerd naar de verschillende niveaus
- 9 De deelnemende leerlingen uit mijn klas gingen graag naar de extra lessen
- 10 De didactische aanpak in de extra lessen sloot goed aan op de aanpak in de klas

Ruimte voor uw eventuele opmerkingen over de kwaliteit en het verloop van de extra lessen

C. De opbrengst van de extra lessen

- 11 De achterstand van de deelnemende leerlingen uit mijn klas is door de extra lessen kleiner geworden
- 12 De deelnemende leerlingen doen door de extra lessen beter mee met de lessen in de klas
- 13 De deelnemende leerlingen uit mijn klas hebben door de extra lessen meer zin om te leren dan voorheen
- 14 Ook andere leerlingen uit mijn klas zouden moeten kunnen deelnemen aan deze extra lessen
- 15 De oefeningen, opdrachten en/of spelletjes uit het lesboek van de extra lessen worden ook in de klas toegepast

Ruimte voor uw eventuele opmerkingen over de opbrengst van de extra lessen

Bijlage 2 Tabellen multivariate analyse

Tabellen bij paragraaf 5.1.3 Parameterschattingen (school) fixed effects (LSDV) modellen

Tabel B2.1 - Parameterschattingen LSDV-model, Begrijpend lezen korte termijn

Begrijpend lezen – korte termijn (eind groep 5 – midden groep 6 n=163)	Model 1		Model 2	
	Bèta		Bèta	
Score eind groep 5 (gecentreerd)	0,293	***	0,429	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie	-0,012		0,029	
<i>Controlekenmerken</i>				
Meisje			0,054	
Leeftijd in maanden			-0,104	
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,204	**
Leerpotentiescore			0,316	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.2 - Parameterschattingen LSDV-model, Begrijpend lezen korte termijn, naar interventietype

Begrijpend lezen – korte termijn (eind groep 5 – midden groep 6 n=163)	Model 1		Model 2	
	Bèta		Bèta	
Score eind groep 5 (gecentreerd)	0,294	***	0,430	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie lezen	-0,007		0,036	
Interventie rekenen	-0,016		0,005	
<i>Controlekenmerken</i>				
Meisje			0,054	
Leeftijd in maanden			-0,105	
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,204	**
Leerpotentiescore			0,317	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.3 - Parameterschattingen LSDV-model, Begrijpend lezen middellange termijn

Begrijpend lezen – middellange termijn (eind groep 5 – eind groep 6 n=163)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,237	***	0,360	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie	-0,052		-0,036	
<i>Controlekenmerken</i>				
Meisje			0,005	
Leeftijd in maanden			-0,028	
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,226	***
Leerpotentiescore			0,285	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.4 - Parameterschattingen LSDV-model, Begrijpend lezen middellange termijn, naar interventietype

Begrijpend lezen – middellange termijn (eind groep 5 – eind groep 6 n=163)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,241	***	0,363	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie lezen	-0,024		-0,008	
Interventie rekenen	-0,081		-0,079	
<i>Controlekenmerken</i>				
Meisje			0,007	
Leeftijd in maanden			-0,031	
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,224	**
Leerpotentiescore			0,289	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.5 Parameterschattingen LSDV-model, Rekenen-Wiskunde korte termijn

Rekenen-Wiskunde – korte termijn (eind groep 5 – midden groep 6 n=129)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,322	***	0,347	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie	0,190	**	0,163	**
<i>Controlekenmerken</i>				
Meisje			0,085	
Leeftijd in maanden			-0,090	
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,164	**
Leerpotentiescore			0,182	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.6 - Parameterschattingen LSDV-model, Rekenen-Wiskunde korte termijn, naar interventietype

Rekenen-Wiskunde – korte termijn (eind groep 5 – midden groep 6 n=129)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,338	***	0,356	***
<i>Conditie</i>				
Controle	ref.			
Interventie lezen	0,048		0,090	
Interventie rekenen	0,316	***	0,217	**
<i>Controlekenmerken</i>				
Meisje			0,079	
Leeftijd in maanden			-0,076	
<i>Gesproken taal thuis</i>				
Nederlands			ref	
Niet-Nederlands / overig / onbekend			0,169	**
Leerpotentiescore			0,170	**
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.7 Parameterschattingen LSDV-model, Rekenen-Wiskunde middellange termijn, naar interventie-type

Rekenen-Wiskunde – middellange termijn (eind groep 5 – eind groep 6 n=112)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,355	***	0,426	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie	0,058		0,078	
<i>Controlekenmerken</i>				
Meisje			0,088	
Leeftijd in maanden			-0,166	**
<i>Gesproken taal thuis</i>				
Nederlands			ref.	
Niet-Nederlands / overig / onbekend			0,237	***
Leerpotentiescore			0,317	***
<i>Incl. dummyvariabelen voor school</i>				

Tabel B2.8 Parameterschattingen LSDV-model, Rekenen-Wiskunde middellange termijn, naar interventie-type

Rekenen-Wiskunde – middellange termijn (eind groep 5 – eind groep 6 n=112)	Model 1		Model 2	
	Bèta		Bèta	
Intercept				
Score eind groep 5 (gecentreerd)	0,383	***	0,443	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie lezen	-0,118		-0,030	
Interventie rekenen	0,227	**	0,179	*
<i>Controlekenmerken</i>				
Meisje			0,089	
Leeftijd in maanden			-0,158	**
<i>Gesproken taal thuis</i>				
Nederlands				
Niet-Nederlands / overig / onbekend			0,246	***
Leerpotentiescore			0,305	***
<i>Incl. dummyvariabelen voor school</i>				

Tabellen bij paragraaf 5.1.3 Parameterschattingen groeicurve modellen

Tabel B2.9 Parameterschattingen groeicurve model, Begrijpend lezen en Rekenen-Wiskunde middellange termijn, naar conditie

	Begrijpend lezen		Rekenen-Wiskunde	
	B		B	
Intercept	138,187	***	197,151	***
<i>Meting</i>				
Eind groep 5	ref.		ref.	
Midden groep 6	19,544	***	16,722	
Eind groep 6	26,653	***	29,196	*
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie	-1,179		-3,481	
<i>Interactie tussen meting en groep</i>				
Midden groep 6 * interventie	0,637		5,388	*
Eind groep 6 * interventie	-1,294		1,470	
Variantie tussen scholen	5		58	
Variantie binnen scholen tussen leerlingen	284	***	443	***
Variantie binnen leerlingen tussen metingen	153	***	124	***
Variantie regressieparameter midden groep 6	24	#	1043	***
Variantie regressieparameter eind groep 6	0	#	0	#
N scholen	8		7	
N leerlingen	164		129	
N metingen	490		370	

***: significant bij $p < 0,01$; **: significant bij $p < 0,05$; *: significant bij $p < 0,1$

#: redundante variantieparameter

Tabel B2.10 Parameterschattingen groeicurve model, Begrijpend lezen en Rekenen-Wiskunde middellange termijn, naar interventietype

	Begrijpend lezen		Rekenen-Wiskunde	
	B		B	
Intercept	138,052	***	197,152	***
<i>Meting</i>				
Eind groep 5	ref.		ref.	
Midden groep 6	19,161	***	16,080	***
Eind groep 6	26,669	***	28,376	***
<i>Conditie</i>				
Controle	ref.		ref.	
Interventie lezen	-4,359		-9,807	*
Interventie rekenen	5,348		4,383	
<i>Interactie tussen meting en groep</i>				
Midden groep 6 * interventie lezen	0,917		5,392	
Midden groep 6 * interventie rekenen	0,982		6,976	*
Eind groep 6 * interventie lezen	-0,120		1,627	
Eind groep 6 * interventie rekenen	-4,285		2,717	
Variantie tussen scholen	4		42	
Variantie binnen scholen tussen leerlingen	284	**	424	***
Variantie binnen leerlingen tussen metingen	153	***	123	***
Variantie regressieparameter midden groep 6	15		40	
Variantie regressieparameter eind groep 6	3	#	1	#
N scholen	8		7	
N leerlingen	164		129	
N metingen	490		370	

***: significant bij $p < 0,01$; **: significant bij $p < 0,05$; *: significant bij $p < 0,1$

#: redundante variantieparameter

