CLST | Centre for Language and Speech Technology
Radboud University

HOGESCHOOL ROTTERDAM
Kenniscentrum Talentontwikkeling

zwijsen
Breng leren tot leven

NRO
NATIONAAL REGIEORGAAN ONDERWIJSONDERZOEK

Foto: Renate Reitler

Oefensoftware bij het leren lezen

Hoe zet je het in en wat levert het op?

Een onderzoek naar de leerlingsoftware bij de methode Veilig leren lezen

Marjoke Bakker, Erik van Schooten, Marjon Verstappen, Rosemarie Irausquin, Maartje Giessen, Yvonne Harmens, Mariska de Weerd, Helmer Strik

Eindrapport van het onderzoeksproject EDiLe “Effecten van een digitaal leermiddel bij het leren lezen”, gefinancierd door het Nationaal Regieorgaan Onderwijsonderzoek (NRO), projectnummer 405-15-507.

ONDERZOEKSRAPPORT

Oefensoftware bij het leren lezen

Hoe zet je het in en wat levert het op?

Een onderzoek naar de leerlingsoftware bij de methode Veilig leren lezen

Marjoke Bakker	CLST, Radboud Universiteit Nijmegen
Erik van Schooten	Kenniscentrum Talentontwikkeling, Hogeschool Rotterdam
Marjon Verstappen	Uitgeverij Zwijsen
Rosemarie Irausquin	Uitgeverij Zwijsen
Maartje Giessen	CLST, Radboud Universiteit Nijmegen
Yvonne Harmens	KBS De Vuurvlinder, Veenendaal
Mariska de Weerd	Wilhelminaschool, Amerongen
Helmer Strik	CLST, Radboud Universiteit Nijmegen

Dit onderzoeksrapport is het eindrapport van het onderzoeksproject EDiLe “Effecten van een digitaal leermiddel bij het leren lezen”. Het onderzoeksproject werd gefinancierd door het Nationaal Regieorgaan Onderwijsonderzoek (NRO), projectnummer 405-15-507.

Het onderzoek werd uitgevoerd door een consortium bestaande uit het Centre for Language and Speech Technology (CLST) van de Radboud Universiteit Nijmegen, Kenniscentrum Talentontwikkeling van de Hogeschool Rotterdam, Uitgeverij Zwijsen en zes basisscholen: KBS De Vuurvlinder te Veenendaal, Wilhelminaschool te Amerongen, Admiraal de Ruyterschool te Amsterdam, Geert Groteschool te Zwolle, De Ceder te Veenendaal, en Martinusschool te Millingen aan de Rijn. Naast de scholen uit het consortium hebben ruim 50 andere basisscholen aan het onderzoek deelgenomen. Wij danken alle deelnemende leerkrachten en leerlingen voor hun medewerking.

Meer informatie over het onderzoeksproject is te vinden op http://hstrik.ruhosting.nl/EDiLe en www.veiliglerenlezen.nl/effectiviteitleerlingsoftware.

Samenvatting

Bij het leren lezen kan oefensoftware veel voordelen bieden. Leerlingen kunnen individueel en op hun eigen niveau oefenen met de leerstof en krijgen daarbij directe feedback. Het is echter niet altijd makkelijk om oefensoftware op een goede manier in het onderwijs in te zetten, bijvoorbeeld vanwege een beperkt aantal computers op school, of vanwege organisatorische moeilijkheden bij het inpassen van het computerwerk in het lesprogramma. In het onderzoeksproject EDiLe (Effecten van een Digitaal leermiddel bij het leren Lezen) is onderzocht hoe oefensoftware het beste kan worden ingezet in het leesonderwijs in groep 3 en welke effecten de software heeft op de leeruitkomsten. Het onderzoek werd gefinancierd door het Nationaal Regieorgaan Onderwijsonderzoek (NRO).

In het onderzoek is gekeken naar de leerlingsoftware bij de methode Veilig leren lezen (VLL, kim-versie, zie Figuur 1). Met deze software kunnen leerlingen op een leuke manier extra oefenen met de leerstof uit de methode, waarbij de oefeningen worden afgestemd op het niveau van de leerling. De software kan naast op school ook thuis worden gebruikt.

Figuur 1. De leerlingsoftware bij VLL kim-versie. Links: startscherm, rechts: oefening uit het onderdeel *Lezen en spelling.*

Deel 1 – Oefensoftware bij het leren lezen: Hoe zet je het in?

Het eerste deel van het onderzoek richtte zich op de vraag: Hoe kan oefensoftware voor het leren lezen het beste in de onderwijspraktijk worden ingezet? Om deze vraag te kunnen beantwoorden werden een interviewstudie en een evaluatiestudie uitgevoerd. In de interviewstudie vroegen we zeven leerkrachten van groep 3 naar hun ervaringen en meningen ten aanzien van het gebruik van de leerlingsoftware bij VLL. In de evaluatiestudie werd met zes leerkrachten van groep 3 geëvalueerd hoe het werken met de leerlingsoftware gedurende een aantal maanden was verlopen en in hoeverre vooraf gemaakte plannen konden worden gerealiseerd. Op basis van deze studies werden een aantal aandachtspunten opgesteld voor het inzetten van de leerlingsoftware bij VLL. Sommige van deze punten zijn heel specifiek voor de leerlingsoftware bij VLL, andere zijn meer algemeen en breder toepasbaar. Enkele belangrijke punten zijn:

- *Voorbereiding:* Leerkrachten zijn vaak niet goed op de hoogte van wat er allemaal met de leerlingsoftware mogelijk is en benutten daarom niet alle mogelijkheden. We adviseren leerkrachten daarom om de software zelf uit te proberen en de handleiding te bestuderen.

- *Plan van aanpak:* Het is goed als leerkrachten een plan hebben over hoe zij de leerlingsoftware in hun onderwijs willen inzetten. Het gaat hierbij bijvoorbeeld om hoe vaak en op welke momenten de leerlingsoftware wordt ingezet. Om optimaal gebruik te maken van een beperkt aantal computers, is het goed om meer momenten te creëren waarop met de software gewerkt kan worden, bijvoorbeeld voor sommige leerlingen al vóór het werken in werkboekjes. Ook loont het om een goed systeem te vinden voor het bijhouden van leerlingbeurten, zodat beurtwisselingen soepel en zo veel mogelijk zonder tussenkomst van de leerkracht kunnen plaatsvinden.
- *Computerproblemen opvangen:* Helaas zijn er op scholen regelmatig problemen met computers of het internet. Om deze problemen zo goed mogelijk op te kunnen vangen, kan het helpen om de computers al aan te zetten voordat de kinderen op school komen, en om een plan te hebben van wat te doen als een computer niet werkt.
- *Thuissoftware:* Om leerlingen te stimuleren om met de thuissoftware te werken, is het goed om als leerkracht in de klas af en toe terug te komen op het werken met de thuissoftware, ook later in het schooljaar. Ook is het aan te raden ouders herhaaldelijk te informeren over de mogelijkheid van het thuis oefenen.

Adviezen voor de onderwijspraktijk

- Lees de handleiding van de leerlingsoftware en probeer de software zelf uit
- Maak een plan van aanpak voor het inzetten van de leerlingsoftware in het onderwijs: wanneer, hoe vaak, bijhouden leerlingbeurten
- Weet wat te doen als een computer niet werkt
- Informeer ouders over de thuissoftware en bespreek het werken met de thuissoftware af en toe in de klas

Deel 2 – Oefensoftware bij het leren lezen: Wat levert het op?

In het tweede deel van het onderzoek is gekeken naar de effectiviteit van de leerlingsoftware bij VLL op de leeruitkomsten van de leerlingen. Er is een grootschalige effectstudie uitgevoerd, waaraan 73 groepen 3 hebben deelgenomen van 57 verschillende basisscholen (een totaal van 1434 leerlingen). In het onderzoek werd een vergelijking gemaakt tussen drie verschillende onderzoeksgroepen:

- Scholen die de VLL kim-versie gebruikten met de *leerlingsoftware op school*
- Scholen die de VLL kim-versie gebruikten met de *leerlingsoftware op school en thuis*
- Scholen die de VLL kim-versie gebruikten *zonder de leerlingsoftware*

We hebben onderzocht of leerlingen op scholen met de leerlingsoftware meer vooruit zijn gegaan in hun leesprestaties dan leerlingen op scholen waar niet met de leerlingsoftware gewerkt werd. Ook hebben we gekeken naar de rol van de mate van softwaregebruik: zijn leerlingen die meer met de leerlingsoftware hebben gewerkt ook meer vooruit gegaan? Verder onderzochten we of de software effectiever is voor bepaalde typen leerlingen, zoals leerlingen met een laag beginniveau.

Onderzoeksmethode

Het onderzoek omvatte de eerste helft van het schooljaar 2015/2016, van augustus/september tot januari. In deze periode werden toetsresultaten van de deelnemende leerlingen verzameld, en er werden gegevens bijgehouden over de mate waarin de leerlingsoftware werd gebruikt.

Toetsen

Om te kunnen meten welke effecten de leerlingsoftware heeft op de leeruitkomsten van de leerlingen, hebben we gebruik gemaakt van toetsresultaten uit het Cito leerlingvolgsysteem en van resultaten op methodetoetsen van VLL. Naast de ontwikkeling van de *technische leesvaardigheid* hebben we ook gekeken naar de ontwikkeling van de gerelateerde vaardigheden *letterkennis* en *spelling*. We hebben gebruik gemaakt van de volgende toetsresultaten:

- *Technisch lezen:* VLL Woordleestoetsen, Cito Drie-Minuten-Toets (DMT) midden groep 3
- *Letterkennis:* VLL Letterkennistoetsen
- *Spelling:* VLL Spellingtoetsen, Cito Spelling 3.0 midden groep 3

Gegevens softwaregebruik

Om te kunnen onderzoeken wat de invloed is van de mate waarin de leerlingsoftware door leerlingen wordt gebruikt, hebben we gebruik gemaakt van automatisch verzamelde gegevens over het softwaregebruik. Het ging om gegevens over het aantal gemaakte oefeningen lezen en spelling in de leerlingsoftware. Deze gegevens waren beschikbaar voor zowel de schoolversie als de thuisversie van de software.

Resultaten

Met de verzamelde gegevens zijn verschillende analyses uitgevoerd. In sommige analyses werden er verschillen in leerwinst gevonden tussen de onderzoeksgroepen, in andere analyses waren er geen verschillen. Als er verschillen waren, waren die altijd in het voordeel van de scholen waar de leerlingsoftware gebruikt werd: op die scholen gingen leerlingen gemiddeld meer vooruit dan op scholen waar de leerlingsoftware niet gebruikt werd. De gevonden verschillen in de leerwinst waren echter niet zo groot, mogelijk vanwege het relatief beperkte gebruik van de leerlingsoftware op de deelnemende scholen (gemiddeld minder dan 10 minuten per week per leerling). We vonden over het algemeen geen verschillen in leerwinst tussen scholen waar alleen op school met de software werd gewerkt en scholen waar daarnaast ook gebruik werd gemaakt van de thuissoftware. Dit kan deels verklaard worden door het geringe gebruik van de thuissoftware.

Technisch lezen

Voor technisch lezen vonden we aanwijzingen dat het gebruik van de leerlingsoftware bijdraagt aan de leerresultaten. De leerwinst in technisch lezen was gemiddeld hoger op de scholen waar de leerlingsoftware werd gebruikt, dan op de scholen waar de leerlingsoftware niet werd gebruikt. Ook vonden we een relatie tussen de mate van gebruik van de leerlingsoftware en de leerwinst: Leerlingen die op school meer met de software hadden geoefend, lieten een hogere leerwinst in leesvaardigheid zien. Ook voor de thuissoftware vonden we dat leerlingen die meer hadden geoefend meer vooruit gingen in leesvaardigheid. Dit effect was voor de thuissoftware wel een stuk kleiner dan voor de software op school. Leerlingen hadden dus meer profijt van het op school oefenen met de software dan van het thuis oefenen met de software. Voor technisch lezen vonden we over het algemeen geen verschillende effecten voor verschillende typen leerlingen. De leerlingsoftware lijkt dus voor alle leerlingen ongeveer even effectief.

Letterkennis en spelling

Ook voor letterkennis en spelling vonden we enkele aanwijzingen dat leerlingen op scholen waar de software werd gebruikt gemiddeld meer vooruitgingen dan leerlingen op scholen waar de software niet werd gebruikt. Vermeld moet worden dat op de scholen waar de leerlingsoftware gebruikt werd, de leerlingen aan het begin van het schooljaar gemiddeld een lagere letterkennis hadden dan

op de scholen waar de software niet gebruikt werd. Hierdoor was er op de scholen waar de software werd gebruikt mogelijk meer ruimte voor vooruitgang voor de leerlingen. Bij letterkennis en spelling was het niet zo dat leerlingen die meer met de software hadden geoefend ook meer hadden geleerd. Voor letterkennis en spelling blijkt de toegevoegde waarde van de software dus minder duidelijk uit ons onderzoek. Wel vonden we dat leerlingen met een lager beginniveau of met laagopgeleide ouders (leerlinggewicht) meer profijt hadden van het thuis oefenen met de software.

Conclusies

De resultaten laten zien dat de leerlingsoftware een toegevoegde waarde biedt bij de ontwikkeling van de leesvaardigheid. Leerlingen die meer met de software oefenden hadden een hogere leerwinst in technisch lezen. Dit geldt vooral voor het op school oefenen met de software, maar in mindere mate ook voor het thuis oefenen. Voor de gerelateerde vaardigheden spelling en letterkennis vonden we ook enkele aanwijzingen voor een effect van de leerlingsoftware, maar hier is de toegevoegde waarde minder duidelijk aanwezig.

De gevonden verschillen in leerwinst tussen de onderzoeksgroepen waren niet zo groot, mogelijk vanwege het relatief beperkte gebruik van de leerlingsoftware op de deelnemende scholen. De positieve relatie tussen mate van softwaregebruik en leerwinst in technische leesvaardigheid laat zien dat er nog meer winst te behalen is als er op scholen meer met de software wordt gewerkt.

De gevonden effecten verschillen niet veel tussen verschillende typen leerlingen. Wel bleek dat leerlingen met een laag beginniveau of laagopgeleide ouders (leerlinggewicht) wat betreft letterkennis en spelling meer kunnen profiteren van de thuissoftware dan andere leerlingen.

Adviezen voor de onderwijspraktijk

Op basis van de resultaten van de effectstudie kunnen een aantal adviezen voor de onderwijspraktijk worden gegeven:

- Zet de leerlingsoftware in als structureel onderdeel van het onderwijsprogramma, het levert immers een bijdrage aan de ontwikkeling van de leesvaardigheid
- Zet de leerlingsoftware in voor alle leerlingen
- Gebruik de thuissoftware *niet in plaats van* de software op school, maar slechts als aanvulling
- Voor zwakke lezers en leerlingen met een leerlinggewicht kan het lonen om het gebruik van

Inhoudsopgave

Inleiding 8
Achtergrond 8
Het onderzoek 9
De leerlingsoftware bij Veilig leren lezen 9
Deel 1 – Oefensoftware bij het leren lezen: Hoe zet je het in? 12
Interviewstudie - Good practices voor het inzetten van de leerlingsoftware 12
Onderzoeksmethode 12
Resultaten 13
Procesevaluatie - Ervaringen van leerkrachten 14
Onderzoeksmethode 14
Resultaten 15
Gebruiksscenario's VLL leerlingsoftware 16
Deel 2 – Oefensoftware bij het leren lezen: Wat levert het op? 19
Onderzoeksmethode 19
Resultaten 32
Discussie en conclusies 35
Conclusies en adviezen 37
Adviezen voor de onderwijspraktijk 37
Literatuur 38
Bijlage bij deel 2: Details data-analyse en resultaten 39

Inleiding

Steeds vaker wordt er in het basisonderwijs gebruik gemaakt van oefensoftware. Zo ook bij het leren lezen in groep 3. Leerkrachten zien de potentie van deze software in de mogelijkheid die wordt geboden om individueel en op eigen niveau te oefenen, waarbij de leerling directe terugkoppeling krijgt. Tevens heeft software de mogelijkheid om visuele en auditieve informatie te koppelen, wat een belangrijke ondersteuning biedt bij het leren lezen. Het gebruik van oefensoftware in de onderwijspraktijk is echter niet altijd makkelijk. Het kan bijvoorbeeld organisatorisch lastig zijn om alle leerlingen aan de beurt te laten komen, of er kunnen problemen optreden met computers. Vragen uit de onderwijspraktijk zijn daarom: hoe kan oefensoftware het beste ingezet worden, en heeft het ook de gewenste effecten op de leerresultaten? In het onderzoeksproject EDiLe (Effecten van een Digitaal leermiddel bij het leren Lezen) is onderzoek gedaan naar deze vragen. In dit onderzoek is gekeken naar de leerlingsoftware bij de methode Veilig leren lezen (kim-versie), de meest gebruikte methode voor aanvankelijk lezen in Nederland.

Achtergrond

Bij het aanvankelijk leren lezen is het belangrijk dat leerlingen intensief oefenen en kilometers maken (bijv. Adams, 1990; Samuels & Flor, 1997). Om leerlingen gerichte oefening te bieden bij het leren lezen en het proces van technisch lezen te helpen automatiseren, kan gebruik gemaakt worden van oefensoftware. Eerdere studies lieten al de potentie zien van educatieve software bij het leren lezen (bijv. Irausquin, Drent, & Verhoeven, 2005; Savage, Abrami, Hipps, & Deault, 2009). Onder andere bij het oefenen van de auditief-visuele koppeling en het fonemisch bewustzijn (bijv. Adams, 1990; Ehri et al., 2001) kan software een belangrijke bijdrage leveren, door het bieden van oefeningen waarin klanken en tekens gelijktijdig worden aangeboden (bijv. Blok, Oostdam, Otter, & Overmaat, 2002). Educatieve software werkt vaak motiverend, kan leerlingen directe feedback geven waardoor ze zelfstandig kunnen oefenen, en biedt vaak differentiatiemogelijkheden, bijvoorbeeld in de vorm van adaptiviteit (zie bijv. Cheung & Slavin, 2012).

Oefensoftware in de onderwijspraktijk

Helaas is het op scholen niet altijd makkelijk om educatieve software op een goede manier in te zetten. Vaak is er bijvoorbeeld maar een beperkt aantal computers beschikbaar waarop gewerkt kan worden (zie bijv. Kennisnet, 2015). Ook kunnen leerkrachten problemen ervaren met de organisatie van het softwaregebruik in de klas (bijv. Macaruso & Hook, 2007).

Effectiviteit van oefensoftware bij het leren lezen

Ondanks de potentie van oefensoftware als ondersteuning bij het leren lezen, is het de vraag of zulke software in de dagelijkse schoolpraktijk daadwerkelijk tot hogere leeropbrengsten leidt. Een meta-analyse van Blok en collega's (2002) liet zien dat educatieve software over het algemeen een positief effect heeft bij het aanvankelijk leren lezen. Het overgrote deel van de geanalyseerde studies was echter kleinschalig en vond plaats in een gecontroleerde setting. Om te weten of de software ook tot de gewenste effecten leidt in de – vaak niet optimale – dagelijkse schoolpraktijk, zijn grootschalige praktijkgerichte studies nodig.

Recent zijn er in het buitenland enkele grootschalige praktijkgerichte onderzoeken uitgevoerd waarbij positieve effecten werden gevonden van digitale leermiddelen bij het beginnend lezen en spellen (bijv. Cassady & Smith, 2014; Schneider et al., 2016). In Nederland is er nog weinig onderzoek gedaan naar

de vraag of oefensoftware geïmplementeerd in de dagelijkse onderwijspraktijk ook een toegevoegde waarde heeft bij het bevorderen van de leesvaardigheid (zie bijvoorbeeld Kennisnet, 2009).

Verschillende effecten voor verschillende typen leerlingen

Uit eerder onderzoek blijkt dat effecten van educatieve software in het leesonderwijs verschillend kunnen zijn voor verschillende typen leerlingen. Zo laten Cheung en Slavin (2012) bijvoorbeeld zien dat zwakke of gemiddelde lezers over het algemeen meer profiteren van digitale leermaterialen dan sterke lezers. Ook Cassady en Smith (2014) vonden voor zwakke lezers een groter effect.

Intensiteit van softwaregebruik

Een ander belangrijk aspect bij de effectiviteit van educatieve software is de mate, of intensiteit, waarin de software gebruikt wordt. Cheung en Slavin (2012) bevelen daarom aan om in onderzoek naar effectiviteit van educatieve softwareprogramma's ook te kijken naar de intensiteit van het softwaregebruik.

Het onderzoek

Het onderzoeksproject EDiLe had als doel om meer te weten te komen over hoe oefensoftware voor het leren lezen het beste kan worden ingezet in het onderwijs en wat de effecten zijn van deze software op de leeruitkomsten. Het onderzoek bestond uit twee delen:

- *Deel 1 – Oefensoftware bij het leren lezen: Hoe zet je het in?* In dit deel werden een interviewstudie en een evaluatiestudie uitgevoerd om te achterhalen wat goede manieren zijn om oefensoftware in het leesonderwijs in te zetten, en wat voor problemen hierbij kunnen optreden.
- *Deel 2 – Oefensoftware bij het leren lezen: Wat levert het op?* In dit deel voerden we een grootschalige effectstudie uit naar de toegevoegde waarde van oefensoftware in het leesonderwijs. Hierbij keken we ook naar de rol van de intensiteit van het softwaregebruik, en we onderzochten of er verschillende effecten waren voor verschillende typen leerlingen.

Het onderzoek richtte zich op de leerlingsoftware bij de methode Veilig leren lezen (kim-versie) voor groep 3 (Uitgeverij Zwijsen, 2013). Veilig leren lezen (VLL) is de meest gebruikte leermethode voor aanvankelijk lezen in Nederland.

Het onderzoek werd uitgevoerd in het schooljaar 2015/2016. Dat was het tweede jaar dat de VLL leerlingsoftware op scholen werd gebruikt (de VLL kim-versie en de bijbehorende leerlingsoftware zijn in 2014 uitgebracht).

Onderzoeksvragen

In het onderzoek werden de volgende onderzoeksvragen beantwoord:

Hoe zet je het in?

1. *Wat zijn good practices voor het inzetten van leerlingsoftware als ondersteuning bij het aanvankelijk leesonderwijs in groep 3?*
2. *Wat zijn de ervaringen van leerkrachten met het inzetten van de software volgens de good practices?*

Wat levert het op?

3. *Wat is de toegevoegde waarde van de leerlingsoftware, wat betreft de leeruitkomsten op het gebied van technische leesvaardigheid en spelling?*
4. *Wat is de relatie tussen de mate waarin de leerlingsoftware door leerlingen wordt gebruikt en hun leeruitkomsten?*
5. *In welke mate verschillen deze effecten voor verschillende typen leerlingen?*

De leerlingsoftware bij Veilig leren lezen

Om een duidelijk beeld te krijgen van de software waarop in dit onderzoek wordt gefocust – de leerlingsoftware bij VLL – volgt hieronder een beschrijving van de software en wat achtergrondinformatie over de methode VLL zelf.

Veilig leren lezen

Opbouw in kernen

In de methode VLL (kim-versie) is de leerstof verdeeld in kernen. Aan het begin van het schooljaar wordt gestart met kern Start, daarna volgen kern 1 t/m kern 11 en tot slot de kern Afsluiting. De kernen Start t/m 6 zijn gericht op het leren van alle letters. In deze periode, die de eerste helft van het schooljaar omvat, worden alle 34 leesletters aangeleerd en worden eenlettergrepige, klankzuivere woorden aangeboden zonder medeklinkerclusters. Vanaf kern 7 worden moeilijkere woordtypen aangeleerd, zoals woorden met medeklinkerclusters en meerlettergrepige woorden. In ons onderzoek richtten we ons op de eerste fase van het leren lezen: de kernen Start t/m 6.

Differentiatie

In VLL worden drie differentiatieniveaus onderscheiden: maan, ster en zon. Kinderen met een normale leesontwikkeling (de grootste groep) volgen de maan-aanpak: zij volgen het reguliere methodeprogramma. Risicolezers krijgen verlengde instructie en uitbreiding van de leertijd in de ster-aanpak. Vlotte lezers krijgen extra verdieping in de zon-aanpak. Bij binnenkomst in groep 3 zijn leerlingen die de zon-aanpak volgen de leerlingen die al correct en vlot eenlettergrepige woorden zonder medeklinkerclusters kunnen lezen, een (bijna) volledige letterkennis hebben en een zelfstandige en gemotiveerde werkhouding laten zien.

De leerlingsoftware

De leerlingsoftware bij VLL kim-versie is een online programma waarin leerlingen zelfstandig en op hun eigen niveau kunnen oefenen met de leerstof. De software kan zowel op een computer/laptop als op een tablet worden gebruikt. Het programma heeft een aantrekkelijke vormgeving en bevat motiverende spelelementen zoals geluidseffecten en een beloningssysteem (leerlingen verdienen in de oefeningen punten, waarmee ze spullen kunnen kopen om in het beloningsspel hun boomhut te versieren). Het programma bevat drie onderdelen: *Lezen en spelling* (‘de speeltuin’), *Woordenschat* (‘de dierentuin’), en *Leesbevordering* (‘de bibliotheek’) (zie Figuur 2). In ons onderzoek richten we ons vooral op het onderdeel *Lezen en spelling*, omdat het hierin gaat om de technische lees- en spellingvaardigheid, de focus van ons onderzoek.

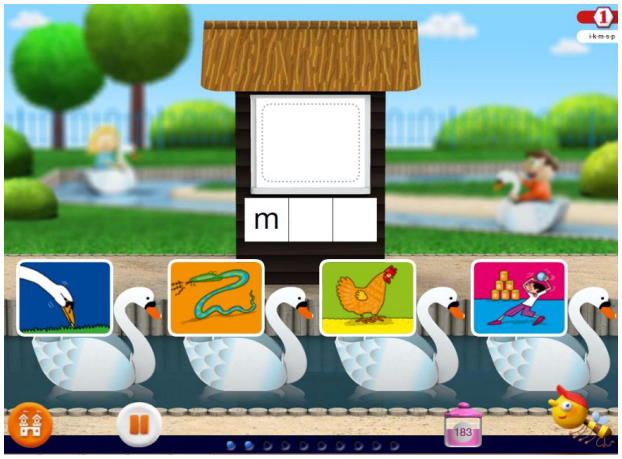
Figuur 2. Startscherm van de leerlingsoftware bij VLL kim-versie. Middenboven is de poort naar de speeltuin (*Lezen en spelling*), linksboven de poort naar de dierentuin (*Woordenschat*) en rechtsboven de poort naar de bibliotheek (*Leesbevordering*).

Lezen en spelling

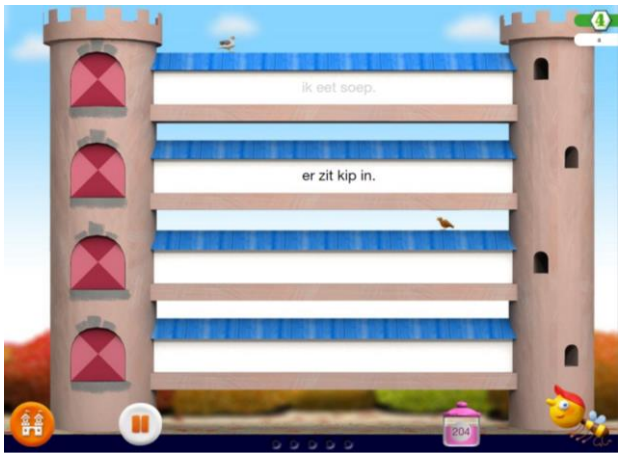
Het onderdeel *Lezen en spelling* (de speeltuin) is het grootste en meest gebruikte onderdeel van de leerlingsoftware. In dit onderdeel kunnen leerlingen werken aan 15 verschillende soorten oefeningen voor technisch lezen en spellen, gericht op auditieve analyse en synthese, de verbinding van fonemen en grafemen, het lezen van woorden, zinnen en teksten, en spelling (zie de voorbeeldoefeningen in Figuur 3, 4 en 5). De inhoud van de oefeningen (gebruikte letters, woordtypen, etc.) sluit nauw aan op wat er in de methode aan bod is gekomen.

Differentiatie

In het onderdeel *Lezen en spelling* is op verschillende manieren differentiatie ingebouwd. Ten eerste zijn de oefeningen die een leerling krijgt aangeboden afgestemd op de aanpak waarin de leerling werkt (maan, zon of ster).



Figuur 3. Oefening *Grafeemkennis* uit de leerlingsoftware. In deze oefenvorm moet de leerling het plaatje aanklikken van het woord dat met de gegeven letter begint (in dit geval 'mik'). Bij de plaatjes worden de woorden door de computer benoemd.



Figuur 4. Oefening *Tekst lezen, correct en vlot* uit de leerlingsoftware. In deze oefenvorm leest de computer zinnen voor en moet de leerling binnen een bepaalde tijd het fout voorgelezen woord aanklikken.

Figuur 5. Oefening *Spelling* uit de leerlingsoftware. In deze oefenvorm moet de leerling het woord invoeren (letters aanklikken) dat bij het plaatje hoort. Het woord wordt ook door de computer benoemd.

Leerlingen die werken met de ster-aanpak krijgen bijvoorbeeld meer tijd in oefeningen met een tijdslimiet, en leerlingen die werken met de zon-aanpak krijgen moeilijkere woordtypen aangeboden. Daarnaast is er adaptiviteit in het programma ingebouwd, waardoor leerlingen vooral oefeningen krijgen met letters die zij nog niet zo goed kennen. Een derde vorm van differentiatie wordt gebaseerd op de resultaten van de letterkennistoetsen uit de methode: als een leerling naast de aangeboden letters al meer letters kent, komen deze extra letters voor in keuze-oefeningen.

Schoolsoftware en thuissoftware

De VLL leerlingsoftware is in eerste instantie bedoeld voor gebruik op school. De software heeft echter ook een thuisversie, zodat de leerlingen ook thuis met de software kunnen oefenen. In de thuissoftware kunnen dezelfde soorten oefeningen gedaan worden als in de schoolsoftware.

Voor de thuissoftware moeten scholen naast de schoolsoftware een aparte licentie aanvragen. Er zijn scholen die alleen de schoolsoftware gebruiken, en er zijn scholen die daarnaast ook gebruik maken van de thuissoftware.

Gegevens softwaregebruik

In de leerlingsoftware wordt automatisch bijgehouden hoe de leerlingen met de software hebben gewerkt. Deze gegevens kunnen via Digiregie, het online registratiesysteem van VLL, door de leerkracht worden bekeken. De leerkracht kan bijvoorbeeld zien hoeveel tijd de leerlingen aan de software hebben besteed, hoeveel oefeningen ze gemaakt hebben, en hoe ze op die oefeningen gepresteerd hebben.

Deel 1

Oefensoftware bij het leren lezen:

Hoe zet je het in?

In dit deel geven we antwoord op de eerste twee onderzoeksvragen.

1. *Wat zijn good practices voor het inzetten van leerlingsoftware als ondersteuning bij het aanvankelijk leesonderwijs in groep 3?*
2. *Wat zijn de ervaringen van leerkrachten met het inzetten van de software volgens de good practices?*

Onderzoeksvraag 1 hebben we beantwoord door middel van een *interviewstudie*, waarin we leerkrachten hebben gevraagd naar hun eerdere ervaringen met en wensen ten aanzien van de leerlingsoftware bij VLL. Onderzoeksvraag 2 hebben we beantwoord door middel van een *procesevaluatie*, waarbij voorafgaand aan het werken met de leerlingsoftware een plan van aanpak werd opgesteld en achteraf werd geëvalueerd hoe het werken volgens dit plan is verlopen.

Interviewstudie - *Good practices voor het inzetten van de leerlingsoftware*

Onderzoeksmethode

Er werden semi-gestructureerde interviews uitgevoerd met in totaal zeven leerkrachten van vijf verschillende basisscholen. Met vier van deze leerkrachten werd het interview individueel afgenomen, telefonisch of in persoon. Bij de drie andere leerkrachten, van één school, had het interview de vorm van een groepsinterview/focusgroep. Het doel van de interviews was om een set van good practices samen te stellen voor het inzetten van de VLL leerlingsoftware.

Van de zeven deelnemende leerkrachten, allemaal vrouwen, hadden er vijf ervaring met het inzetten van de VLL leerlingsoftware in het leesonderwijs in groep 3. Een van de leerkrachten in het groepsinterview kende de leerlingsoftware nog niet, maar zou er het komende schooljaar mee gaan werken. Een laatste leerkracht had zich wel verdiept in de leerlingsoftware, maar had deze nog niet in de klas kunnen inzetten vanwege technische problemen.

In alle interviews werden dezelfde onderwerpen besproken, volgens een vooraf opgestelde interview-leidraad. Er werden open vragen gesteld, waarbij er op werd gelet of bepaalde punten aan bod kwamen. Punten die niet aan bod kwamen, werden expliciet door de interviewer aan de orde gebracht. De interviewvragen richtten zich op de meningen, ervaringen en verwachtingen van leerkrachten ten aanzien van het gebruik van digitaal leermateriaal in het algemeen en de VLL leerlingsoftware in het bijzonder, sterke en zwakke punten van de VLL leerlingsoftware, en mogelijke problemen bij het inzetten van de leerlingsoftware en mogelijke oplossingen daarvoor. Ook werd aan de leerkrachten gevraagd wat zij andere leerkrachten zouden adviseren.

Van de interviews werden audio-opnames gemaakt, van het groepsinterview ook een video-opname. De interviews werden vervolgens uitgeschreven. Voor het opstellen van de good practices werden van het groepsinterview en (vanwege tijdsdruk) van twee van de vier individuele interviews de uitgeschreven fragmenten onderverdeeld naar de onderwer-

pen waarop zij betrekking hadden, zoals de klasse-organisatie en planning bij het inzetten van de leerlingsoftware, de instellingen van de leerlingsoftware, en het gebruik van de thuissoftware. Vervolgens werden de overige twee individuele interviews doorlopen/beluisterd en werden eventuele aanvullingen op de informatie per onderwerp genoteerd. Tot slot werd de informatie per onderwerp vertaald naar good practices voor de onderwijspraktijk.

Resultaten

Uit de interviews bleek dat leerkrachten enthousiast zijn over de software: het ziet er leuk uit en leerlingen werken er graag mee. Er werd echter ook aangegeven dat leerkrachten vaak niet alle mogelijkheden van de software kennen of benutten, dat het moeilijk kan zijn om alle leerlingen aan de beurt te laten komen, en dat er soms technische problemen optreden met computers of internet.

De good practices die uit de interviewstudie naar voren kwamen hadden betrekking op veel verschillende aspecten van het gebruik van de leerlingsoftware. Sommige van de good practices zijn heel specifiek voor de leerlingsoftware bij VLL, andere zijn meer algemeen en breder toepasbaar. Hieronder volgt een samenvatting van de belangrijkste punten. Voor de complete lijst van good practices verwijzen we naar het document “Aandachtspunten & adviezen voor het gebruik van de leerlingsoftware bij de kim-versie van Veilig Leren Lezen”, te downloaden op www.veiliglerenlezen.nl/effectiviteitleerlingsoftware.

Voorbereiding

Een belangrijk punt dat in veel van de interviews terug kwam was de voorbereiding van de leerkracht. Het bleek dat leerkrachten niet altijd goed weten wat er allemaal met de leerlingsoftware mogelijk is, waardoor vaak niet alle mogelijkheden van de software worden benut. De richtlijn die hieruit volgt is dat leerkrachten, voordat zij met de software gaan werken, de handleiding bestuderen en de software zelf uitproberen.

Plan van aanpak

In de interviews kwam naar voren dat leerkrachten het belangrijk vinden een goed plan te hebben van hoe zij de leerlingsoftware in hun onderwijs gaan inzetten. Het gaat hierbij bijvoorbeeld om een plan van hoe vaak en op welke momenten de leerlingsoftware wordt ingezet, en een schema voor het bijhouden van de leerlingbeurten.

Leerkrachten vonden dat de VLL leerlingsoftware een structureel onderdeel zou moeten zijn van het onderwijs en dat de software voor alle leerlingen zou moeten worden ingezet. Een belangrijk aandachtspunt dat hieruit volgde was het creëren van voldoende momenten in het lesprogramma waarop met de software gewerkt kan worden.

Instellingen leerlingsoftware

Een punt dat verwant is met het eerstgenoemde punt over de voorbereiding, betreft het goed instellen van de leerlingsoftware. In de VLL leerlingsoftware kunnen leerkrachten door middel van instellingen de leerstof beter laten aansluiten bij de behandelde stof in de methode, en bij het niveau van de leerling (maan, zon, ster). Verder kan ingesteld worden hoeveel tijd leerlingen per keer met de software kunnen werken en welke onderdelen van de software beschikbaar zijn. Uit de interviews bleek dat leerkrachten niet altijd op de hoogte zijn van deze instelmogelijkheden en deze niet altijd benutten.

Instructie aan leerlingen

Leerkrachten gaven aan dat het goed werkt om de software klassikaal op het digibord te introduceren. Verder werd aangegeven dat het goed is om leerlingen te leren hoe zij zelfstandig de software kunnen aanzetten, gebruiken en afsluiten, zodat het werken met de software en het wisselen van leerlingbeurten soepel kan verlopen.

Monitoring

Verwant met het vorige punt is het advies om af en toe mee te kijken als de leerlingen met de software werken. Dit meekijken dient niet alleen om na te gaan of leerlingen serieus met de oefeningen bezig zijn, maar ook om de boodschap over te brengen dat het werken met de software een onderdeel is van het onderwijs.

Softwaregegevens in Digiregie bekijken

Uit de interviews kwam naar voren dat de meeste leerkrachten het nuttig vinden dat zij in Digiregie gegevens over het softwaregebruik van de leerlingen kunnen bekijken. Veel leerkrachten gaven echter aan dat het er vaak niet van komt om deze gegevens te raadplegen. Een advies dat hieruit voortkomt, is dat leerkrachten een plan maken van hoe vaak en op welke momenten zij de gegevens willen bekijken, en wat zij hiermee willen doen. De gegevens kunnen voor verschillende doeleinden worden gebruikt: terugkoppeling naar de leerlingen (bijv. complimenten), sturen van leerlingen in het gebruik van de leerlingsoftware, inzicht krijgen in wat leerlingen al kunnen, en de evaluatie van het leerproces aan het eind van een kern.

Opvangen van computerproblemen

Een aantal leerkrachten gaf aan regelmatig problemen te hebben met computers of het internet. Om deze problemen zo goed mogelijk op te vangen, werd een aantal tips gegeven. Zo werd opgemerkt dat het tijd spaart als de computers al aangezet worden voordat de les begint, zodat eventuele updates al geïnstalleerd kunnen worden en eventuele computerproblemen mogelijk al kunnen worden opgelost voordat de leerlingen op school zijn. Ook is het handig als leerkrachten een plan hebben van wat er gedaan moet worden als computers niet goed werken, bijvoorbeeld zo snel mogelijk even tijd maken om ernaar te kijken, en indien nodig een ICT-coördinator of instanties inschakelen. Verder werd aangegeven dat het goed is als leerlingen weten wat te doen bij een niet werkende computer, bijvoorbeeld weer aan hun tafeltje gaan zitten en aan iets anders gaan werken.

Thuissoftware

Wat betreft de thuissoftware gaven de leerkrachten aan dat het belangrijk is de ouders te informeren over de mogelijkheid thuis met de software te werken. Dit kan bijvoorbeeld door middel van ouderbrieven, of door op ouderavonden de software te tonen op het digibord. Verder werd aangeraden om in de klas af en toe terug te komen op het gebruik van de thuissoftware, om zo het thuis oefenen verder te stimuleren.

Procesevaluatie - *Ervaringen van leerkrachten*

Onderzoeksmethode

In de procesevaluatie werd onderzocht wat de ervaringen van leerkrachten zijn als zij een periode met de leerlingsoftware hebben gewerkt volgens een vooraf opgesteld plan van aanpak. Aan deze evaluatiestudie deden zes leerkrachten van groep 3 mee, allen vrouwen, van vijf scholen die ook deelnamen aan de effectstudie (zie Deel 2 van dit rapport). Een van deze leerkrachten had ook deelgenomen aan de interviewstudie.

De procesevaluatie bestond uit twee fasen:

- Eind september 2015: Inventariseren van plannen en verwachtingen
- Begin februari 2016: Evaluatie van het werken met de leerlingsoftware

Plannen en verwachtingen

In september 2015 (ongeveer een maand na de start van het schooljaar) vulden de zes deelnemende leerkrachten een uitgebreide vragenlijst in over hun plannen en verwachtingen ten aanzien van het werken met de leerlingsoftware.[1] De leerkrachten gaven in deze

[1] Een van de leerkrachten vulde de vragenlijst pas in december in, omdat op haar school pas later werd gestart met het gebruik van de leerlingsoftware.

vragenlijst aan hoe zij de leerlingsoftware in hun onderwijs wilden inzetten.[2] De punten op de vragenlijst waren gebaseerd op de hierboven beschreven good practices. Deze good practices werden ook, ter inspiratie, meegestuurd met de vragenlijst. We vroegen de leerkrachten, in lijn met de good practices, naar hun plannen wat betreft:

- Voorbereiding
- Hoe, hoe vaak, op welke momenten en voor welke leerlingen wordt de leerlingsoftware ingezet
- Instellingen van de leerlingsoftware
- Instructie aan de leerlingen
- Monitoring
- Softwaregegevens in Digiregie bekijken
- Gebruik van de thuissoftware

Daarnaast hebben we de leerkrachten gevraagd naar verwachte problemen bij het werken met de leerlingsoftware en naar mogelijke oplossingen hiervoor.

Na het invullen van de vragenlijst werden alle leerkrachten telefonisch geïnterviewd, waarbij hun gemaakte plan van aanpak werd besproken en eventuele onduidelijkheden werden verhelderd.

Evaluatie van het werken met de leerlingsoftware

In februari 2016, na afloop van de onderzoeksperiode van de effectstudie, werden er opnieuw telefonische interviews gehouden met de zes leerkrachten. In deze interviews werd, aan de hand van het eerder ingevulde plan van aanpak, geëvalueerd hoe het werken met de leerlingsoftware was verlopen, welke problemen hierbij waren opgetreden, en welke oplossingen hiervoor zijn bedacht.

Na de telefonische interviews met de individuele leerkrachten, vond er een discussiesessie plaats met vier van de deelnemende leerkrachten. In deze sessie konden de leerkrachten hun ervaringen met de leerlingsoftware met elkaar uitwisselen. De discussiesessie verliep volgens een aantal discussiepunten, gebaseerd op de telefonische evaluatie-interviews: er werd vooral gefocust op moeilijkheden bij het gebruik van de leerlingsoftware en op punten waarin leerkrachten verschilden in hun aanpak.

Verwerking van de gegevens

Op basis van de ingevulde vragenlijsten en audio-opnames van de telefonische interviews aan het begin en eind van de onderzoeksperiode, werd voor elke school een verslag gemaakt van het gemaakte plan van aanpak, in hoeverre dit plan kon worden uitgevoerd, en eventuele problemen die hierbij optraden. Vanuit de verslagen per school werd vervolgens, in combinatie met de uitgeschreven discussiesessie, een overzicht gemaakt van voorkomende moeilijkheden en aandachtspunten bij het werken met de leerlingsoftware.

Resultaten

Uit de procesevaluatie bleek dat de leerlingsoftware door verschillende leerkrachten op verschillende manieren wordt ingezet. De manier van inzetten is onder andere afhankelijk van de beschikbaarheid van computers. Ter illustratie bevat het kader op de volgende pagina voor vier van de deelnemende scholen een beschrijving van de manier waarop de leerlingsoftware werd ingezet.

Uit de procesevaluatie kwamen verschillende aandachtspunten naar voren. De punten kwamen grotendeels overeen met de bevindingen uit de interviewstudie. We bespreken de belangrijkste punten.

[2] De meeste leerkrachten maakten op het moment van het invullen van de vragenlijst al gebruik van de leerlingsoftware. Toch werd bij het invullen van de vragenlijst benadrukt dat het ging om plannen over hoe zij de software de komende periode zouden willen inzetten. Vaak kwamen deze plannen overeen met hoe zij de software op dat moment al gebruikten.

Gebruiksscenario's VLL leerlingsoftware

School 1

Groep 3 heeft 19 leerlingen. Er zijn drie computers in de klas. De leerlingen zijn in een tabel verdeeld over de drie computers. Als de kinderen 's ochtends binnenkomen (om 8:20 uur) gaan ze werken aan verwerkingsstof van de taal-/leesles van de dag ervoor. Bij drie leerlingen ligt dan een kaartje op tafel, dat aangeeft dat zij op de computer mogen. Zij maken één oefening in de leerlingsoftware en roepen dan de volgende leerling op de lijst. Om ca. 9 uur begint dan de kring en om ca. 9:20 uur de instructie voor taal/lezen. Vóór de start van de instructie zijn meestal alle kinderen al aan de beurt geweest met de leerlingsoftware. Als kinderen nog niet klaar zijn, wordt dit tijdens de instructie afgemaakt. Op maandag, woensdag en vrijdag wordt er met het onderdeel *Lezen en spelling* gewerkt, op dinsdag en donderdag met het onderdeel *Woordenschat*. Met het onderdeel *Leesbevordering* en het beloningsspel (de boomhut) mogen de leerlingen alleen thuis werken. De reden dat de leerkracht kiest voor één oefening per keer is dat elke leerling dan elke dag aan de beurt komt. Een nadeel is dat leerlingen niet aan alle oefeningen toe komen.

School 2

Groep 3 heeft 25 leerlingen. Er zijn twee computers in de klas (één daarvan loopt soms vast), en daarnaast zes computers op de gang die gedeeld worden met twee andere groepen. De leerlingsoftware staat ingesteld op sessies van 10 minuten. 's Ochtends bij binnenkomst gaan twee leerlingen van de ster-aanpak met de leerlingsoftware werken achter de computers in de klas; de andere leerlingen zijn dan aan het stillezen. Tijdens de instructie taal/lezen kunnen de leerlingen van de zon-aanpak, als zij dan geen instructie hebben (om de dag), alvast achter de computer. Na de instructie worden leerlingen van de maan-aanpak aangewezen om achter de computer te gaan; als zij klaar zijn wijst de leerkracht weer nieuwe leerlingen aan. Ze streept op een lijst af welke leerlingen zijn geweest, maar soms vergeet ze dit ook. Kinderen van de ster-aanpak gaan soms na de verlengde instructie nog even op de computer. Bij het werken met de leerlingsoftware geeft de leerkracht soms aan met welk onderdeel de leerlingen moeten gaan werken, soms mogen de leerlingen zelf kiezen. Leerlingen komen niet elke dag aan de beurt; de leerkracht is blij als alle leerlingen één of twee keer in de week aan de beurt zijn geweest. 's Middags komt het er vaak niet van om achter de computer te werken, omdat er dan veel dingen zijn met de hele groep.

School 3

Groep 3 heeft 31 leerlingen. Er zijn drie computers in de klas en drie computers in het (lege) lokaal ernaast. Het zijn oude computers die vaak vastlopen. De drie computers in de klas zijn verdeeld over groepen leerlingen (elke computer hoort bij een rij leerlingen). Kinderen van de zon-aanpak werken vaak op de computers in het lokaal ernaast, omdat zij al goed zelfstandig kunnen werken. De leerlingsoftware staat ingesteld op 10 minuten per sessie. Kinderen werken met de leerlingsoftware in de leesmomenten, die zijn vóór de ochtendpauze, na de ochtendpauze, en na de lunchpauze. In het eerste leesmoment komen leerlingen aan de beurt die klaar zijn met hun andere zelfstandig werk van taal/lezen. In het tweede en derde leesmoment kunnen kinderen van de ster-aanpak achter de computer. Met de andere leerlingen wordt dan hardop of stil gelezen, dat gaat voor de leerlingen uit de ster-aanpak vaak nog wat te snel. Op deze manier komt elk kind wel een keer per week aan de beurt, de leerlingen van de ster-aanpak vaker. Omdat het werken met de leerlingsoftware voor de leerlingen van de maan- en zon-aanpak klaarwerk is, komen van deze leerlingen de snellere leerlingen wat vaker aan de beurt. Voor de leerlingen van de ster-aanpak wordt bijgehouden wie geweest is; voor de leerlingen van de maan- en zon-aanpak niet, dit regelen de leerlingen zelf (de eerste die klaar is gaat op de computer, als die klaar is gaat er weer iemand anders uit de rij die bij die computer hoort; wel is er de regel dat leerlingen niet vaker dan 1x per dag mogen). Er wordt vooral gewerkt met het onderdeel *Lezen en spelling*.

School 4

Groep 3 heeft 24 leerlingen. Er zijn twee computers in de klas, en op dinsdag kan de groep 45 minuten gebruik maken van 15 computers in de hal. Op de dinsdag werken alle leerlingen 20 minuten met de leerlingsoftware. Daarnaast is er in het klaarwerk, als leerlingen klaar zijn met hun verwerking van de taal-/leesles, de mogelijkheid om gedurende 10 minuten met de software te werken. Met een lijst wordt bijgehouden wie aan de beurt is: elke dag zijn twee andere leerlingen aan de beurt. Als die leerlingen op die dag niet op tijd klaar zijn met hun verwerking taal/lezen, mogen ze op een ander moment achter de computer, bijvoorbeeld na rekenen. Er wordt alleen gebruik gemaakt van het onderdeel *Lezen en spelling*, omdat hier de focus op ligt in het onderwijs. Voor het gebruik van de thuissoftware vraagt de school de ouders een bijdrage van 5 euro. Ongeveer twee derde van de ouders heeft ervoor gekozen deze bijdrage te betalen, de andere leerlingen hebben geen toegang tot de thuissoftware.

Frequentie van softwaregebruik

Een aantal leerkrachten gaf aan dat zij de software minder vaak konden gebruiken dan zij vooraf gepland hadden. Redenen die werden genoemd waren problemen met computers en moeilijkheden om voldoende tijd beschikbaar te maken in het lesprogramma.

Bij alle deelnemende leerkrachten werd de frequentie van het softwaregebruik beperkt door het beperkte aantal computers dat beschikbaar was. De leerkrachten gaven aan dat zij, indien dat mogelijk zou zijn, de leerlingen meer met de software zouden willen laten werken. In het ideale geval zouden ze kiezen voor 10 à 15 minuten per leerling per dag.

Bijhouden van leerlingbeurten

Zoals ook al bleek uit de interviewstudie, is het voor leerkrachten soms lastig om er voor te zorgen dat alle leerlingen aan de beurt komen om met de leerlingsoftware te werken. Leerkrachten hebben vaak wel een systeem hiervoor – bijvoorbeeld een namenlijstje – maar niet alle leerkrachten zijn tevreden met het systeem dat zij gebruiken. Dit blijft dus een aandachtspunt. Twee leerkrachten gaven bijvoorbeeld aan dat zij steeds zelf aangeven welke leerlingen op een bepaald moment aan de beurt zijn. Beurtwisselingen zouden soepeler kunnen verlopen als dit zonder tussenkomst van de leerkracht zou gaan. Dit blijkt echter niet altijd makkelijk te organiseren, bijvoorbeeld in gevallen dat een leerling die bovenaan de lijst staat nog niet klaar is met zijn andere werk of bij problemen met computers. Leerkrachten doen er dus goed aan een flexibel systeem te vinden dat goed aansluit bij hun onderwijspraktijk.

Monitoring

De leerkrachten gaven aan dat zij in het begin van het schooljaar regelmatig even langs de computers liepen, om te kijken hoe het werken met de leerlingsoftware verliep en om eventuele computerproblemen op te lossen. Na verloop van tijd werd er minder aandacht besteed aan het meekijken, deels omdat het minder nodig was (minder computerproblemen, leerlingen hadden geen hulp meer nodig), en deels vanwege tijdgebrek. Een van de leerkrachten gaf aan dat zij liever vaker met de leerlingen mee zou willen kijken, omdat ze merkt dat dit motiverend werkt voor de leerlingen.

Softwaregegevens in Digiregie bekijken

De meeste van de deelnemende leerkrachten gaven aan weinig gebruik te maken van de gegevens in Digiregie over het softwaregebruik van de leerlingen. Leerkrachten kwamen vaak niet toe aan het bekijken van de registratiegegevens en wisten vaak ook niet goed wat ze met de informatie konden doen. Leerkrachten die wel naar de gegevens in Digiregie hadden gekeken, deden dit vooral uit nieuwsgierigheid en om eventuele opvallende dingen aan te kunnen kaarten bij leerlingen (bijvoorbeeld als een leerling heel weinig oefeningen had gemaakt of erg laag had gepresteerd in de oefeningen). Ook voor hen was het onduidelijk waarvoor ze de gegevens verder nog zouden kunnen gebruiken. Hier ligt een aandachtspunt voor eventuele verbetering van de software of de bijbehorende leerkrachtinstructies, zodat leerkrachten beter weten wat zij met de registratiegegevens kunnen.

Opvangen van computerproblemen

De leerkrachten waren het er over eens dat goed werkende apparatuur een voorwaarde is voor het goed kunnen gebruiken van de leerlingsoftware. Als computers traag werken of vastlopen, lukt het niet de leerlingsoftware goed in te zetten of alle leerlingen aan de beurt te laten komen. Mogelijke oplossingen die genoemd werden zijn het proberen van een andere internetbrowser en het opnieuw opstarten van het programma. Verbeteringen van de software zouden ook kunnen helpen.

Thuissoftware

Op drie van de scholen die deelnamen aan de procesevaluatie werd ook met de thuissoftware gewerkt. De leerkrachten van deze scholen gaven aan dat zij hier vooral in het begin

van het schooljaar aandacht aan besteedden. Ze hebben de ouders geïnformeerd en af en toe in de les het werken met de thuissoftware besproken. Later in het schooljaar werd de thuissoftware steeds minder door de leerkrachten ter sprake gebracht, en de leerkrachten merkten ook op dat er door de leerlingen na verloop van tijd minder met de thuissoftware gewerkt werd. Het werken met de thuissoftware lijkt dus zowel voor leerlingen als voor leerkrachten in de loop van het schooljaar te verwateren. Om leerlingen te stimuleren om ook later in het schooljaar met de thuissoftware te blijven werken, zou het kunnen helpen om hier als leerkracht aandacht aan te blijven besteden. Ook vonden de leerkrachten het een goed idee om de thuissoftware herhaaldelijk bij de ouders onder de aandacht te brengen, bijvoorbeeld door middel van ouderbrieven.

Deel 2

Oefensoftware bij het leren lezen:

Wat levert het op?

In dit deel geven we antwoord op de onderzoeksvragen:

3. *Wat is de toegevoegde waarde van de leerlingsoftware, wat betreft de leeruitkomsten op het gebied van technische leesvaardigheid en spelling?*
4. *Wat is de relatie tussen de mate waarin de leerlingsoftware door leerlingen wordt gebruikt en hun leeruitkomsten?*
5. *In welke mate verschillen deze effecten voor verschillende typen leerlingen?*

Om deze vragen te kunnen beantwoorden, hebben we een grootschalige effectstudie uitgevoerd.

Onderzoeksmethode

In de effectstudie werd een vergelijking gemaakt tussen scholen waar VLL werd gebruikt met daarbij de leerlingsoftware, en scholen waar VLL werd gebruikt zonder de leerlingsoftware. Verder werd er onderscheid gemaakt tussen scholen waar de leerlingsoftware op school werd gebruikt, en scholen die naast de schoolversie van de software ook gebruik maakten van de thuisversie van de software. In de studie werden dus drie onderzoeksgroepen, of condities, vergeleken:

- *SoftwareSchool:* VLL kim-versie met de leerlingsoftware op school
- *SoftwareSchoolThuis:* VLL kim-versie met de leerlingsoftware op school en thuis
- *Controlegroep:* VLL kim-versie zonder de leerlingsoftware

Het onderzoek besloeg de eerste helft van het schooljaar 2015/2016, van september tot januari.

We hebben gekozen voor een quasi-experimentele onderzoeksopzet. Dit betekent dat de indeling van de scholen in de onderzoeksgroepen gebeurde op basis van de keuze van de scholen zelf. We gingen dus op zoek naar scholen die in het schooljaar 2015/2016 gebruik zouden maken van VLL kim-versie, al dan niet met de leerlingsoftware en eventueel de thuissoftware. De reden dat we de keuze bij de scholen lieten was tweeledig. Ten eerste wilden we het effect onderzoeken van de leerlingsoftware in een zo natuurlijk mogelijke situatie, dat wil zeggen op scholen die deze software normaal gesproken ook gebruiken. Ten tweede vonden we het niet ethisch om scholen die de leerlingsoftware normaal gesproken wel zouden gebruiken, het gebruik hiervan voor een half jaar te ontzeggen.

Werving van scholen

Om scholen te vinden die wilden deelnemen aan het onderzoek, hebben we in juni 2015 een groot aantal scholen die gebruik maakten van de VLL kim-versie telefonisch of per e-mail benaderd. Veel scholen waren geïnteresseerd om mee te doen. Uit de geïnteresseerde scholen hebben we een selectie gemaakt, waarbij we ervoor hebben gezorgd dat de drie verschillende onderzoeksgroepen zo goed mogelijk vergelijkbaar waren. Hierbij hebben we gelet op verschillende schoolkenmerken, zoals schoolgrootte, percentage leerlingen met een leerlinggewicht en anderstalige leerlingen, en of de school in een stedelijke of meer landelij-

ke omgeving staat. Ook werd er gekeken naar of de school al ervaring had met het werken met de kim-versie van VLL. Om de groepen beter vergelijkbaar te maken hebben we gericht nog extra scholen benaderd die aan bepaalde kenmerken voldeden.

Een probleem bij de werving van de scholen was dat er maar weinig scholen bleken te zijn waar de leerlingsoftware niet gebruikt werd. Om het aantal scholen in deze groep aan te vullen, hebben we in september/oktober 2015 nog verder gezocht naar deelnemende scholen voor deze groep. Sommige scholen uit deze groep hebben daarom niet de gehele looptijd van het onderzoek deelgenomen. Ook konden we bij deze extra scholen niet meer goed rekening houden met de vergelijkbaarheid. We hebben daarom in onze analyses nogmaals gekeken naar de vergelijkbaarheid van de groepen scholen.

Deelnemende scholen

Uiteindelijk hebben 57 scholen deelgenomen aan het onderzoek, met in totaal 73 klassen groep 3 en in totaal 1434 leerlingen. Van de deelnemende scholen gaven er 19 aan de VLL kim versie te (gaan) gebruiken met de leerlingsoftware op school, 26 met de leerlingsoftware op school en thuis, en 12 zonder de leerlingsoftware (Controlegroep). Later bleek echter dat het daadwerkelijke gebruik (al dan niet) van de leerlingsoftware niet bij alle scholen overeenkwam met wat eerder was aangegeven. Er waren een paar scholen uit de Controlegroep die gedurende het onderzoek toch de leerlingsoftware hadden aangeschaft, er waren scholen die de software alleen op school zouden gebruiken maar later toch ook de thuisversie hadden aangeschaft, en er waren enkele scholen die van plan waren de leerlingsoftware te gebruiken, maar waar dit door technische problemen niet mogelijk was. De uiteindelijke indeling in de drie onderzoeksgroepen hebben wij gemaakt op basis van wat er daadwerkelijk is gebeurd op de scholen. Verder is één school, waar de leerlingsoftware niet gebruikt werd, voortijdig gestopt met het onderzoek. De 56 overgebleven scholen waren uiteindelijk als volgt verdeeld over de onderzoeksgroepen: 15 scholen in de groep SoftwareSchool, 32 scholen in de groep SoftwareSchoolThuis, en 9 scholen in de Controlegroep (zie Tabel 1).

Tabel 1. *Aantal scholen, klassen en leerlingen per onderzoeksgroep (conditie).*

Onderzoeksgroep	Aantal scholen	Aantal klassen	Aantal leerlingen
SoftwareSchool	15	18	366
SoftwareSchoolThuis	32	44	877
Controlegroep	9	10	175
Totaal	56	72	1418

Procedure

Het onderzoek liep van het begin van het schooljaar (augustus/september 2015) tot halverwege het schooljaar (januari 2016). In deze periode werd in de deelnemende klassen gewerkt met VLL kim-versie zoals dat normaal ook zou gebeuren, al dan niet met de leerlingsoftware.

De leerkrachten van de scholen waar de leerlingsoftware gebruikt werd, kregen aan het begin van het schooljaar de lijst met aandachtspunten en adviezen die was opgesteld naar aanleiding van de interviewstudie (beschreven in Deel 1 van dit rapport). Zij konden deze aandachtspunten en adviezen naar eigen inzicht inzetten.

Gedurende de onderzoeksperiode werden verschillende gegevens verzameld, die hierna verder worden toegelicht:

- Toetsresultaten van toetsen van VLL
- Toetsresultaten uit het Cito leerlingvolgsysteem
- Gegevens uit de leerlingsoftware van VLL
- Leerkracht-logboeken

De Cito toetsresultaten en de leerkracht-logboeken werden direct bij de deelnemende leerkrachten opgevraagd. De VLL toetsresultaten en de gegevens uit de leerlingsoftware konden door ons worden opgevraagd via Digiregie, het online registratiesysteem van VLL. Om de gegevens uit Digiregie te mogen gebruiken, hebben we alle deelnemende scholen om toestemming gevraagd. We hebben de scholen verzocht om ook de ouders van de leerlingen om toestemming te vragen, middels een brief waarop ouders konden reageren als zij bezwaar hadden tegen de deelname van hun kind. De scholen gaven ons toestemming voor het gebruik van de gegevens door ons een toestemmingsformulier te sturen met de namen, en andere gegevens, van de leerlingen waarvoor de toestemming werd verleend. Alleen voor deze leerlingen werden de gegevens uit Digiregie opgevraagd.

Toetsen

Om te kunnen meten welke effecten de leerlingsoftware heeft op de ontwikkeling van de leesvaardigheid van de leerlingen, hebben we verschillende toetsresultaten van de leerlingen verzameld. Naast de ontwikkeling van de *technische leesvaardigheid* hebben we ook gekeken naar de ontwikkeling van de gerelateerde vaardigheden *letterkennis* en *spelling*. Vanwege de grootschaligheid van ons onderzoek, hebben we ervoor gekozen gebruik te maken van toetsen die in het reguliere toetsprogramma van de scholen werden afgenomen.

Om goed te kunnen onderzoeken of de vooruitgang in het lezen wordt beïnvloed door het gebruik van de leerlingsoftware, zijn er toetsresultaten op verschillende tijdsmomenten nodig: een meting aan het begin van de onderzoeksperiode, aan het eind van de onderzoeksperiode, en liefst ook metingen gedurende de onderzoeksperiode. We hebben in ons onderzoek gebruik gemaakt van zowel toetsen die horen bij de methode VLL, als toetsen uit het Cito leerlingvolgsysteem (zie Tabel 2). De toetsen van VLL hebben als voordeel dat ze op meer momenten worden afgenomen; de ontwikkeling van de leerlingen kan dus nauwkeuriger worden gevolgd. De toetsen uit het Cito leerlingvolgsysteem hebben als voordeel dat het gestandaardiseerde en methode-onafhankelijke toetsen zijn.

Toetsen van Veilig leren lezen

Letterkennistoetsen. In de *VLL Letterkennistoetsen Alle letters* wordt gemeten welke letters de leerlingen al kennen. De toetsen worden individueel afgenomen. Leerlingen krijgen een kaart voorgelegd met letters erop, en worden gevraagd deze letters voor te lezen. De *VLL Letterkennistoetsen Alle letters* worden afgenomen aan het begin van het schooljaar (kern Start), en na kern 3 en kern 6. In alle drie

Tabel 2. *Overzicht verzamelde toetsgegevens.*

Toets	Afnameperiode						
	eind aug / begin sep	eind sep	okt	okt / nov	nov / dec	dec / jan	jan / feb
Toetsen Veilig leren lezen[a]							
Letterkennis: VLL Letterkennistoetsen Alle letters	LKstart			LK3			LK6
Technisch lezen: VLL Woordleestoetsen		WL1	WL2	WL3	WL4	WL5	WL6
Spelling: VLL Spellingtoetsen				SP3	SP4	SP5	SP6
Toetsen Cito leerlingvolgsysteem							
Technisch lezen: Drie-minuten-toets (DMT)							DMT M3
Spelling: Cito Spelling 3.0							Spelling 3.0 M3

[a]LK = Letterkennistoets, WL = Woordleestoets, SP = Spellingtoets. 'start' en de cijfers 1 t/m 6 verwijzen naar de kernen in VLL.

deze toetsen worden alle 34 leesletters[3] getoetst. De toetsen verschillen wel wat in de manier van afnemen. In kern Start krijgen de leerlingen een letterkaart voorgelegd met de vraag om alle letters die zij kennen te benoemen ('Inventarisatie Letterkennis'). De toetsen van kern 3 en 6 worden op een meer gestructureerde manier afgenomen: de leerlingen lezen alle letters één voor één voor. Bij kern 3 bestaat de Letterkennistoets uit twee toetsen: 'Letterkennis aangeboden letters' en 'Letterkennis nog niet aangeboden letters'. Bij kern 6 wordt de toets als één geheel afgenomen, omdat dan alle letters zijn aangeboden.

De toetsscore op de *Letterkennistoetsen Alle letters* is het aantal correct benoemde letters. Voor kern 3 berekenden we de totaalscore als de som van het aantal correct benoemde letters in de twee deeltoetsen.

Woordleestoetsen. De *VLL Woordleestoetsen* meten de technische leesvaardigheid middels het hardop lezen van losse woorden. Ook deze toetsen worden individueel afgenomen. Leerlingen krijgen een kaart met rijtjes woorden voorgelegd, met de vraag om binnen een minuut zo veel mogelijk woorden correct te lezen. De *VLL Woordleestoetsen* worden aan het eind van elke kern afgenomen, vanaf kern 1. De toetsen bevatten steeds alleen woordjes met letters die al aangeboden zijn. De verschillende *Woordleestoetsen* zijn dus niet identiek, maar afgezien van de letters die in de woorden voorkomen, lijken de *Woordleestoetsen* van de kernen 1 tot en met 6 erg op elkaar. In al deze toetsen komen alleen klankzuivere éénlettergrepige woorden voor zonder medeklinkerclusters.[4] De toetsscore op de *VLL Woordleestoetsen* is het aantal correct gelezen woorden in 1 minuut.[5]

Spellingtoetsen. De *VLL Spellingtoetsen* worden vanaf kern 3 afgenomen, steeds aan het eind van de kern. Deze toetsen worden klassikaal afgenomen in de vorm van een dictee. Leerlingen moeten losse woorden opschrijven, die steeds in de context van een zin worden voorgelezen. De op te schrijven woorden bestaan steeds alleen uit letters die al aangeboden zijn. In de toetsen van kern 3 t/m 6 zijn het steeds klankzuivere éénlettergrepige woorden zonder medeklinkerclusters (kern 3: 10 woorden, kern 4 en 5: 12 woorden, kern 6: 14 woorden). De score op de *Spellingtoetsen* is het aantal juist gespelde woorden.

Beheersingstoetsen vs. vaardigheidstoetsen. De VLL Spellingtoetsen zijn bedoeld als beheersingstoetsen: ze toetsen of de aangeboden leerstof door de leerlingen wordt beheerst. Leerlingen die de leerstof beheersen halen gemakkelijk de maximale score op de toets. De Spellingtoetsen vertonen dus een plafondeffect. Dit maakt de toetsen minder geschikt voor een analyse van de voortgang in spellingvaardigheid. We hebben van deze toetsen daarom alleen de eerste toets gebruikt (*VLL Spellingtoets kern 3*), als een maat voor het 'beginniveau' van spellingvaardigheid (in combinatie met een eindmeting uit het Cito leerlingvolgsysteem, zie hierna). Hierbij moet wel vermeld worden dat, door het genoemde plafondeffect, deze toets slecht onderscheid kan maken tussen de hoog presterende leerlingen.

---

[3] De leesletters zijn: *i, e, o, a, u, aa, ee, oo, uu, oe, ij, eu, ie, ou, au, ui, ei, k, m, s, p, r, v, n, t, b, d, z, h, w, j, l, g, f.*

[4] Bij leerlingen in de zon-aanpak (vlotte lezers) worden naast de reguliere woordleestoetsen aanvullende toetsen afgenomen met moeilijkere woorden (medeklinkerclusters). Omdat deze aanvullende toetsen slechts bij een klein deel van de leerlingen werden afgenomen, hebben we deze niet meegenomen in ons onderzoek.

[5] De toetsscore op de *Woordleestoetsen* is normaal gesproken begrensd: De maximale toetsscore is het totaal aantal woorden op de leeskaart (kern 1: 60 woorden, kern 2-6: 80 woorden). Om een nog iets nauwkeurigere maat te hebben van de leesvaardigheid, hebben wij de toetsscore herberekend als het aantal correct gelezen woorden gedeeld door de daadwerkelijk bestede leestijd. Op deze manier kan meer onderscheid gemaakt worden tussen snelle lezers die al voor het eind van de afnametijd alle woorden hebben gelezen.

De VLL Woordleestoetsen zijn vaardigheidstoetsen, geen beheersingstoetsen. Deze toetsen zijn zo samengesteld dat de maximale score maar door zeer weinig leerlingen wordt behaald. De Woordleestoetsen kunnen daarom goed onderscheid maken tussen leerlingen met verschillende vaardigheidsniveaus, en kunnen goed gebruikt worden om de groei in leesvaardigheid te meten.

Voor de *Letterkennistoetsen Alle letters* geldt dat de toetsen van kern Start en kern 3 kunnen worden gezien als vaardigheidstoetsen. Deze toetsen gaan verder dan de aangeboden leerstof: ze toetsen niet alleen aangeboden maar ook nog niet aangeboden letters. Verwacht kan worden dat de maximale score niet vaak wordt behaald, en de toets goed gebruikt kan worden om onderscheid te maken tussen leerlingen met verschillende niveaus. De *Letterkennistoets* van kern 6 is echter een beheersingstoets. Aan het eind van kern 6 zijn namelijk alle 34 leesletters aangeboden; de *Letterkennistoets Alle letters* toetst dan dus de beheersing van de aangeboden leerstof. Voor kern 6 kan dus verwacht worden dat veel leerlingen de maximale score op de toets halen, waardoor de toets alleen onderscheid maakt tussen zwak presterende leerlingen.

Toetsen uit het Cito leerlingvolgsysteem

Drie-Minuten-Toets (DMT). De *Drie-Minuten-Toets (DMT)* is een standaardtoets voor het meten van technische leesvaardigheid (Jongen & Krom, 2010; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). De toets wordt volgens het Cito leerlingvolgsysteem elk half jaar afgenomen, van groep 3 tot en met groep 8. In ons onderzoek hebben we gebruikgemaakt van de DMT-resultaten van het afnamemoment M3 (midden groep 3). Dit is het eerste afnamemoment van de *DMT*.

De *DMT* word individueel afgenomen en bestaat uit drie leeskaarten met losse woorden (150 woorden per kaart). Per kaart moeten de leerlingen binnen een minuut zo veel mogelijk woorden correct voorlezen. Hiermee is de toets vergelijkbaar met de *VLL Woordleestoetsen*. Leeskaart 1 van de *DMT* bevat klankzuivere eenlettergrepige woorden zonder medeklinkerclusters, leeskaart 2 bevat eenlettergrepige woorden met medeklinkerclusters, en leeskaart 3 bevat woorden met meer dan één lettergreep. Meestal worden halverwege groep 3 alleen kaart 1 en 2 afgenomen, maar het komt ook voor dat alle drie de kaarten worden afgenomen of dat alleen kaart 1 wordt afgenomen.

De ruwe score per woordkaart is gelijk aan het aantal juist gelezen woorden binnen de afnametijd van 1 minuut. De scores van meerdere kaarten samen kunnen via omrekentabellen worden omgezet in een vaardigheidsscore (Jongen & Krom, 2010), zodat scores op verschillende combinaties van woordkaarten met elkaar kunnen worden vergeleken (bijv. kaart 1 en 2 vs. kaart 1, 2 en 3). Omdat niet op alle scholen in ons onderzoek dezelfde combinatie van woordkaarten werd afgenomen, hebben we gebruik gemaakt van deze vaardigheidsscores. Zo konden we de *DMT*-resultaten van alle scholen meenemen in onze analyses.

Spelling 3.0. De toets *Spelling 3.0* (Cito, 2014) toetst de spellingvaardigheid van de leerlingen in de vorm van een dictee. In ons onderzoek hebben we gebruik gemaakt van de toets op afnamemoment M3 (midden groep 3). Deze toets bestaat uit een dictee met losse woorden (40 woorden, verdeeld over 2 deeltoetsen). De toets kan klassikaal worden afgenomen, waarbij de dicteewoorden door de leerkracht worden voorgelezen, of digitaal, waarbij de computer de woorden voorleest. De leerlingen krijgen van elk op te schrijven of in te typen woord ook een afbeelding te zien.

De ruwe score op de toets *Spelling 3.0 M3* is het aantal juist gespelde woorden. Deze score kan, net als bij de *DMT*, via een omrekentabel worden omgezet in een vaardigheidsscore (Cito, 2014). Aangezien voor sommige van de

deelnemende scholen alleen deze vaardigheidsscores beschikbaar waren, en niet de ruwe scores, hebben we in onze analyses gebruik gemaakt van de vaardigheidsscores.

Beginmetingen, eindmetingen, tussenmetingen

In groep 3 wordt gestart met het leren lezen. Aan het begin van het schooljaar kunnen de meeste leerlingen dus nog niet lezen of spellen. Een beginmeting van het leesniveau aan het begin van groep 3 is daarom lastig. De enige toets die in aanmerking kwam als een echte beginmeting, dat wil zeggen aan het begin van de onderzoeksperiode, was de *VLL Letterkennistoets Alle Letters kern Start*. Deze toets toetst de kennis van de letters, een voorwaarde voor lees- en spellingvaardigheid, en kan daarom worden gebruikt als een maat van het beginniveau van de leerlingen.

Letterkennis is weliswaar een voorwaarde voor lees- en spellingvaardigheid, maar letterkennistoetsen toetsen natuurlijk niet dezelfde vaardigheid als lees- en spellingtoetsen. Voor het onderzoeken van leeropbrengsten op lees- en spellingvaardigheid hebben we daarom ook analyses uitgevoerd met als beginmeting de eerste echte toets van leesvaardigheid resp. spellingvaardigheid: de *VLL Woordleestoets kern 1* voor leesvaardigheid en de *VLL Spellingtoets kern 3* voor spellingvaardigheid. Deze beginmetingen hebben echter het nadeel dat ze niet aan het begin van de onderzoeksperiode zijn afgenomen. Ten tijde van de toetsafname is het werken met de leerlingsoftware dus al een tijdje bezig. Met deze “beginmetingen” wordt dus gekeken naar de voortuitgang van de leerlingen gedurende een deel van de onderzoeksperiode, vanaf de afname van de toets.

De toetsen uit het Cito leerlingvolgsysteem werden pas aan het eind van de onderzoeksperiode afgenomen en konden dus alleen als eindmeting in het onderzoek meegenomen worden. Het bijbehorende beginniveau werd gemeten met de bovengenoemde beginmetingen: de *VLL Letterkennistoets Alle letters kern Start*, en daarnaast voor leesvaardigheid de *VLL Woordleestoets kern 1* en voor spellingvaardigheid de *VLL Spellingtoets kern 3*.[6]

Omdat de toetsen van VLL meerdere keren werden afgenomen, kon hiermee nauwkeuriger naar de ontwikkeling van de leerlingen worden gekeken. Voor zowel de *Letterkennistoetsen* als de *Woordleestoetsen* onderzochten we de ontwikkeling van de prestaties op de verschillende toetsen.[7] Dit werd gedaan met groeimodellen (meer hierover volgt later), met de eerste toets van die vaardigheid als voormeting, de laatste toets als nameting, en de overige toets(en) als tussenmeting(en).

Verzamelen van de toetsresultaten

Zoals eerder aangegeven, werden de toetsresultaten van de VLL toetsen door ons verzameld via Digiregie. We vroegen de deelnemende scholen om de toetsresultaten in Digiregie in te voeren, zodat we toegang hadden tot deze resultaten. De toetsresultaten van de toetsen van het Cito leerlingvolgsysteem werden direct bij de scholen opgevraagd.

Gegevens softwaregebruik

Om te kunnen onderzoeken of er een relatie is tussen de mate van gebruik van de leerlingsoftware en de leeropbrengsten, hadden we gegevens nodig over de mate, of intensiteit, van het softwaregebruik. Deze gegevens werden automatisch geregistreerd in Digiregie.

---

[6] We wilden ook de resultaten van de toetsen *Taal voor Kleuters* groep 2 uit het Cito leerlingvolgsysteem meenemen als beginmeting. Deze toetsen waren afgenomen in januari en/of juni 2015. In onze analyses bleek echter dat deze toetsen niet geschikt waren om als beginmeting te gebruiken, omdat de scores geen hoge relatie hadden met de scores op de eindmetingen. Waarschijnlijk komt dit doordat de toetsen *Taal voor Kleuters* naast de voorlopers van het technisch lezen ook andere aspecten van taalvaardigheid meten (zoals woordenschat), en doordat deze toetsen een aantal maanden vóór het begin van de onderzoeksperiode waren afgenomen.

[7] Voor de *VLL Spellingtoetsen* deden we dit niet, omdat deze toetsen vanwege de plafondeffecten op elke toets niet geschikt waren voor het onderzoeken van groei.

Elke keer dat een leerling met de software werkte, werd onder andere bijgehouden hoeveel tijd de leerling aan de software besteedde en hoeveel oefeningen hij of zij maakte in de verschillende onderdelen van de leerlingsoftware. In ons onderzoek richtten we ons op de gegevens over het gebruik van het softwareonderdeel *Lezen en spelling* (de speeltuin), vanwege onze focus op leesvaardigheid.

In Tabel 3 is weergegeven hoeveel tijd de leerlingen in ons onderzoek gemiddeld per maand hebben besteed aan oefeningen in het onderdeel *Lezen en spelling*. Te zien is dat er in beide Software-groepen (SoftwareSchool en SoftwareSchoolThuis) *op school* gemiddeld ongeveer evenveel tijd is besteed aan de oefeningen in de software, gemiddeld 30 minuten per maand.[8] Deze gemiddelde tijdsbesteding is vrij laag: per week hebben leerlingen op school gemiddeld minder dan 10 minuten aan de software besteed. Dit steekt mager af tegen het in de evaluatiestudie genoemde ideaal van 10 à 15 minuten per dag. In lijn met wat naar voren kwam in de evaluatiestudie, is de mate van gebruik van de software dus niet optimaal geweest. Mogelijke verklaringen worden in Deel 1 van dit rapport gegeven en betreffen onder andere de beperkte beschikbaarheid van computers, en problemen met de organisatie van het softwaregebruik in de klas.

Het thuisgebruik van de software (in de SoftwareSchoolThuis-groep) was gemiddeld nog een stuk minder dan het gebruik op school: gemiddeld werkten leerlingen per maand 10 minuten thuis aan oefeningen *Lezen en spelling*. Dit kan deels te maken hebben met voorkomende problemen bij het werken met de thuissoftware: het bleek dat sommige leerlingen thuis geen toegang hadden tot de software. Verder is te zien dat het thuis werken met de software afnam over de tijd, wat ook al werd opgemerkt in de procesevaluatie.

Tabel 4 laat zien hoeveel oefeningen de leerlingen gemiddeld gemaakt hebben in het onderdeel *Lezen en spelling*. Dit waren op school gemiddeld zo'n 15 oefeningen per maand; thuis werden er gemiddeld zo'n 5 oefeningen per maand gemaakt.

Om de gegevens over de intensiteit van softwaregebruik in onze analyses mee te kunnen nemen, hebben we deze gegevens omgezet in een aantal variabelen, intensiteitsvariabelen genoemd. We hebben ons hierbij gericht op het aantal gemaakte oefeningen, omdat we dit beschouwen als de beste indicatie van de hoeveelheid leerstof die leerlingen in de leerlingsoftware aangeboden hebben gekregen.. De intensiteitsvariabelen werden apart berekend voor het gebruik van de schoolsoftware en de thuissoftware. Ze werden steeds per leerling berekend als het totaal aantal gemaakte software-oefeningen *Lezen en spelling* in een bepaalde periode. Voor elke analyse werd de periode waarover de intensiteitsvariabelen werden berekend, aangepast aan de afnamedata van de betreffende toetsen: we keken steeds naar het aantal gemaakte oefeningen voorafgaand aan een bepaalde toets (zie voor details de bijlage).

Leerkracht-logboeken

Aan de leerkrachten van de deelnemende klassen werd gevraagd om in de onderzoeksperiode vier keer een week lang een logboek bij te houden. De weken waarin de logboeken moesten worden ingevuld waren gelijkmatig verspreid over de onderzoeksperiode. De ingevulde gegevens dienden als achtergrondvariabelen in de analyses. Het ging om de hieronder genoemde gegevens.

Bestede tijd aan taal-/leesonderwijs

We vroegen de leerkrachten in de logboeken naar hoeveel tijd er in de klas werd besteed aan taal-/leesonderwijs. De reden dat we naar deze tijdsbesteding vroegen, is om hiervoor te

[8] De totaal bestede tijd aan alle oefeningen in de software, inclusief de andere onderdelen *Woordenschat* en *Leesbevordering*, was gemiddeld zo'n 40 minuten per maand. Het grootste deel van de totale softwaretijd werd dus besteed aan het onderdeel *Lezen en spelling*.

Tabel 3. *Bestede tijd (uren:minuten) aan oefeningen in het software-onderdeel Lezen en spelling, per maand per leerling.*

	SoftwareSchool-groep (*n* = 366) *schoolsoftware* (bestede tijd per leerling)				SoftwareSchoolThuis-groep (*n* = 877) *schoolsoftware* (bestede tijd per leerling)				SoftwareSchoolThuis-groep (*n* = 877) *thuissoftware* (bestede tijd per leerling)			
Maand	*M*	(*SD*)	min	max	*M*	(*SD*)	min	max	*M*	(*SD*)	min	max
augustus/september	0:29	(0:42)	0	2:51	0:31	(0:25)	0	3:31	0:11	(0:26)	0	3:51
oktober	0:29	(0:38)	0	2:53	0:31	(0:29)	0	3:18	0:15	(0:33)	0	6:16
november	0:40	(0:54)	0	4:21	0:36	(0:29)	0	4:18	0:11	(0:36)	0	12:47
december	0:24	(0:31)	0	2.25	0:19	(0:22)	0	4:00	0:07	(0:19)	0	3:52
januari	0:30	(0:38)	0	3.08	0:25	(0:23)	0	3:48	0:06	(0:19)	0	3:19
Totaal	2:32	(3:10)	0	15:03	2:22	(1:40)	0	8:49	0:50	(1:37)	0	23:39
Gemiddelde per maand	0:30				0:28				0:10			

Noot. M = gemiddelde, *SD* = standaarddeviatie, min = minimum, max = maximum.

Tabel 4. *Aantal gemaakte oefeningen in het software-onderdeel Lezen en spelling, per maand per leerling.*

	SoftwareSchool-groep (*n* = 366) *schoolsoftware* (aantal oefeningen per leerling)				SoftwareSchoolThuis-groep (*n* = 877) *schoolsoftware* (aantal oefeningen per leerling)				SoftwareSchoolThuis-groep (*n* = 877) *thuissoftware* (aantal oefeningen per leerling)			
Maand	*M*	(*SD*)	min	max	*M*	(*SD*)	min	max	*M*	(*SD*)	min	max
augustus/september	15.2	(20.5)	0	86	16.5	(14.7)	0	173	5.7	(12.6)	0	100
oktober	15.0	(19.3)	0	88	17.1	(16.2)	0	103	7.5	(17.4)	0	189
november	20.3	(26.7)	0	134	18.1	(14.4)	0	128	5.3	(14.2)	0	203
december	11.8	(15.6)	0	85	9.4	(10.5)	0	114	2.9	(8.4)	0	108
januari	16.4	(22.6)	0	114	12.9	(11.9)	0	84	2.9	(9.6)	0	136
Totaal	78.7	(97.6)	0	454	73.9	(52.5)	0	322	24.4	(45.2)	0	505
Gemiddelde per maand	15.7				14.8				4.9			

Noot. M = gemiddelde, *SD* = standaarddeviatie, min = minimum, max = maximum.

kunnen corrigeren in onze analyses. Het zou namelijk kunnen dat op scholen waar de leerlingsoftware werd gebruikt, het softwaregebruik als extra leestijd werd ingeroosterd, waardoor de totale tijd besteed aan taal/lezen op die scholen hoger zou zijn dan op scholen die de leerlingsoftware niet gebruikten. Een verschil in leerwinst tussen de Software-groepen en de Controlegroep zou dan al kunnen zijn veroorzaakt door het verschil in tijdsbesteding, en niet zozeer door het gebruik van de leerlingsoftware. Om deze alternatieve verklaring te kunnen uitsluiten, is het belangrijk om de totaal bestede tijd aan taal-/leesonderwijs als controlevariabele mee te nemen in de analyses.

In de logboekweken werd door de leerkrachten elke dag ingevuld hoeveel minuten er in de klas waren besteed aan taalonderwijs (Nederlands), en hoeveel minuten hiervan aan lezen/spelling. Dit werd apart gevraagd voor de verschillende aanpakken (maan, zon, ster), om het voor de leerkrachten zo eenduidig mogelijk te maken welke bestede tijd werd bedoeld. Voor de analyses hebben we per klas gemiddelde waarden voor de tijdsbesteding berekend op basis van de logboekgegevens. Deze waarden zijn gemiddeld over de 20 logboekdagen (vier logboekweken met vijf dagen per logboek; eventuele vrije dagen werden niet in het gemiddelde meegenomen). We keken hierbij alleen naar de tijds-

besteding voor de maan-aanpak, omdat dit de grootste groep leerlingen betreft. Zo werden per klas twee waarden voor gemiddelde tijdsbesteding berekend: gemiddelde bestede tijd per dag aan taalonderwijs (inclusief bijv. woordenschat), en gemiddelde bestede tijd per dag aan onderwijs in lezen en spelling (een onderdeel van taalonderwijs).

Gebruik van andere software voor taal/ lezen

In ieder logboek werd de leerkrachten gevraagd of zij in hun taal-/leesonderwijs gebruik maakten van andere oefensoftware voor taal/lezen dan de leerlingsoftware bij VLL. Ook hier ging het om een controlevariabele: het gebruik van andere software voor taal/lezen (zowel in de Controlegroep als in de Software-groepen) zou invloed kunnen hebben op de gemeten leerwinst. We wilden in onze analyses corrigeren voor deze mogelijke invloed.

In elk logboek werd de leerkrachten gevraagd om een schatting te maken van hoe vaak er in de klas door de leerlingen gebruik werd gemaakt van andere software voor taal/lezen, in aantal minuten per leerling per week. Op basis van de gegevens in de vier logboeken werd voor elke klas een gemiddelde tijdsbesteding aan andere taal-/leessoftware berekend.

Leerlingen maan-, zon- en ster-aanpak

In elk logboek vulden de leerkrachten in welke leerlingen die week in de zon-aanpak hadden gewerkt en welke leerlingen in de ster-aanpak. Hierbij gingen we ervan uit dat de niet-genoemde leerlingen de maan-aanpak volgden (de standaardaanpak). We gebruikten de gegevens van het eerste logboek als indicatie voor het beginniveau: maan = gemiddeld, zon = hoog, ster = laag.[9] Dit beginniveau werd gebruikt om te onderzoeken of leerlingen met een verschillend beginniveau ook een verschillend effect ondervinden van de leerlingsoftware.

Aanvullende leerling-, klas- en schoolgegevens

Naast de toetsresultaten, softwaregegevens, en leerkrachtlogboeken, werden nog enkele andere gegevens verzameld van de deelnemende leerlingen, klassen en scholen. Deze gegevens werden gebruikt als achtergrondvariabelen ('covariaten') in de analyses (hierover later meer). Ook werd een aantal leerlinggegevens gebruikt om te kunnen onderzoeken of de leerlingsoftware verschillende effecten had voor verschillende typen leerlingen.

Leerlinggegevens

Via de eerder genoemde toestemmingsformulieren die de scholen aan het begin van het schooljaar invulden, verzamelden we de volgende leerlinggegevens: geslacht, leeftijd (in maanden nauwkeurig), leerlinggewicht, thuistaal, en of leerlingen in Nederland of in het buitenland waren geboren. Verder vroegen we de deelnemende leerkrachten welke leerlingen zittenblijvers waren.

Het leerlinggewicht werd gebruikt als indicator voor het opleidingsniveau van de ouders. Bij leerlingen zonder leerlinggewicht (de grootste groep), heeft minstens een van de ouders een opleidingsniveau van vmbo gemengde of theoretische leerweg of hoger (havo, vwo, mbo, hbo, wo). Bij een leerlinggewicht van 0.3 hebben beide ouders maximaal praktijkonderwijs of vmbo basis- of kaderberoepsgerichte leerweg gedaan. Bij een leerlinggewicht van 1.2 heeft een van de ouders maximaal basisonderwijs afgerond en de andere ouder maximaal praktijkonderwijs of

[9] Voor een aantal scholen was de begin-indeling in maan-, zon- en ster-aanpak onbekend (onder andere voor een aantal scholen uit de Controlegroep die later met het onderzoek waren gestart). Voor de leerlingen van deze scholen gebruikten we, indien beschikbaar, de in Digiregie aangegeven indeling voor kern 1.

vmbo basis- of kaderberoepsgerichte leerweg (CFI, 2006).

Klasgegevens

Via twee online vragenlijsten voor deelnemende leerkrachten verzamelden we aanvullende klasgegevens, zoals de klasgrootte (aantal leerlingen), of het een combinatieklas was, en de leeftijd en het aantal jaren onderwijservaring van de leerkracht. Ook werd gevraagd of leerkrachten het vorige schooljaar al gewerkt hadden met VLL kim-versie.

Schoolgegevens

Aanvullende gegevens over de deelnemende scholen werden verzameld via internet. Via DUO[10] werden gegevens verzameld over de denominatie van de scholen en over het schoolpercentage leerlingen met een leerlinggewicht (dit is een maat voor het percentage achterstandsleerlingen). Via gegevens van het CBS[11] werd, op basis van het adres van de scholen, de mate van stedelijkheid en het gemiddelde inkomen achterhaald van de buurt waarin de school staat.

Initiële verschillen tussen de onderzoeksgroepen

Voorafgaand aan onze analyses hebben we gekeken naar de achtergrondkenmerken van de scholen, klassen en leerlingen in de drie onderzoeksgroepen. Om de vergelijkbaarheid tussen de drie onderzoeksgroepen (condities) te controleren, hebben we onderzocht of er verschillen waren tussen de onderzoeksgroepen op achtergrondkenmerken. De gegevens voor de drie onderzoeksgroepen en de vergelijkingen tussen de groepen, zijn weergegeven in Tabel 5 (leerlingkenmerken), Tabel 6 (klaskenmerken) en Tabel 7 (schoolkenmerken).

Voor de meeste achtergrondvariabelen vonden we geen significant verschil tussen de drie onderzoeksgroepen. Er waren bijvoorbeeld geen verschillen tussen de onderzoeksgroepen in gemiddelde leeftijd van de leerlingen, percentage leerlingen met een leerlinggewicht, klasgrootte, of stedelijkheid van de buurt. Ook vonden we geen verschil tussen de onderzoeksgroepen in de gemiddelde bestede tijd aan taal/lezen. In de Controlegroep werd dus gemiddeld ongeveer evenveel tijd besteed aan taal/lezen als in beide Softwaregroepen.

Op enkele punten vonden we wel een significant verschil tussen de onderzoeksgroepen. Een eerste verschil vonden we voor het percentage anderstalige leerlingen. In de SoftwareSchoolThuis-groep en de Controlegroep zaten meer anderstalige leerlingen dan in de SoftwareSchool-groep (zie Tabel 5). Daarnaast zaten in de Controlegroep ook meer Friestalige leerlingen. Voor deze verschillen werd waar nodig gecorrigeerd in de analyses.

Een tweede verschil tussen de onderzoeksgroepen betrof de score op de beginmeting, de *VLL Letterkennistoets Alle Letters kern Start* (zie Tabel 5). Op deze beginmeting scoorden de leerlingen in de Controlegroep gemiddeld hoger dan de leerlingen in de Software-groepen. Hoe dit verschil is veroorzaakt, is niet duidelijk. Het kan in ieder geval niet zijn veroorzaakt door het al dan niet gebruiken van de leerlingsoftware, omdat de *Letterkennistoets kern Start* al aan het begin van het schooljaar werd afgenomen. Het gevonden verschil werd meegenomen in de analyses en bij de interpretatie ervan. Gerelateerd aan het verschil in score op de *Letterkennistoets kern Start* vonden we ook een verschil tussen de onderzoeksgroepen in het beginniveau maan/zon/ster (zie Tabel 5). Zoals te verwachten is op basis van de hogere scores op de *Letterkennistoets kern Start*, waren er in de Controlegroep aanzienlijk meer leerlingen die startten in de zon-aanpak (hoog beginniveau), dan in de twee Softwaregroepen. Daarnaast waren er in de Controlegroep ook wat meer leerlingen die startten in de ster-aanpak (laag beginniveau) dan in de

[10] https://duo.nl/open_onderwijsdata/databestanden/
[11] http://www.cbsinuwbuurt.nl

Tabel 5. *Verschillen tussen de onderzoekscondities in leerlingkenmerken.*

		Geslacht	Leerlinggewicht[a]		Leeftijd	Zittenblijver[b]	Anderstalig[c] (behalve Fries)	Friestalig[c]	Buitenlands	Beginniveau[d] (maan/zon/ster)		Letterkennis kern Start[e]
Onderzoeksconditie	*n*	% meisje	% 0.3	% 1.2	*M* (*SD*)	% zitten-blijver	% anderstalig	% Friestalig	% niet in NL geboren	%zon	%ster	*M* (*SD*)
Controlegroep	175	44.0	6.9	2.9	6.34 (0.40)	4.0	9.1	5.7	2.3	14.5	22.1	22.2 (8.0)
SoftwareSchool	366	48.4	6.9	1.1	6.29 (0.39)	3.0	3.6	2.5	1.1	5.7	15.8	18.3 (8.1)
SoftwareSchoolThuis	877	49.4	5.5	3.6	6.32 (0.42)	4.4	7.9	0.5	1.8	6.8	13.0	17.4 (8.3)
Totaal	1418	48.4	6	2.9	6.32 (0.41)	4.0	6.9	1.6	1.7	7.2	14.6	18.3 (8.4)
Verschil tussen condities		$X^2(2) = 1.69$	$X^2(4) = 6.91$		$F(2,1415) = 1.30$	$X^2(2) = 1.37$	$X^2(2) = 9.07$**	$F(2,335.9) = 6.96$**[f]	$F(2,1415) = 0.63$[g]	$X^2(4) = 22.19$***		$F(2,1057) = 19.20$***
Effectgrootte		$\eta^2 = .001$	Cramer's V = .049		$\eta^2 = .002$	$\eta^2 = .001$	$\eta^2 = .080$	$\eta^2 = .139$	$\eta^2 = .001$	Cramer's V = .091		$\eta^2 = .035$

[a]*n* = 1416. [b]*n* = 1417. [c]*n* = 1415. [d]*n* = 1354. [e]*n* = 1060. [f]Welch F test i.v.m. niet voldoen aan aannames chi-kwadraat-toets (te lage verwachte celfrequentie) en significante afwijking van homogeniteit van varianties. [g]ANOVA F test i.v.m. niet voldoen aan aannames chi-kwadraat-toets.
** $p < .01$. *** $p < .001$.

Tabel 6. *Verschillen tussen de onderzoekscondities in klaskenmerken.*

		Klasgrootte (aantal leerlingen)	Combinatie-klas	Leeftijd leer-kracht[a]	Ervaring leer-kracht[a] (jaren)	Ervaring met VLL kim-versie	Tijdsbesteding leesonderwijs (uren per dag)	Tijdsbesteding taalonderwijs (uren per dag)	Gebruik andere software (minuten per week)
Onderzoeksconditie	*n*	*M* (*SD*)	% combiklas	*M* (*SD*)	*M* (*SD*)	% met ervaring	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Controlegroep	10	21.2 (5.7)	40.0	42.6 (14.2)	17.5 (11.3)	50.0	1.24 (0.21)	1.43 (0.22)	11.3 (21.8)
SoftwareSchool	18	23.3 (5.9)	33.3	40.4 (11.3)	17.4 (10.0)	61.1	1.25 (0.27)	1.49 (0.34)	5.2 (12.6)
SoftwareSchoolThuis	44	22.9 (4.2)	20.5	41.7 (10.5)	17.0 (9.5)	75.0	1.28 (0.23)	1.51 (0.24)	2.2 (5.0)
Totaal	72	22.8 (4.8)	26.4	41.5 (11.1)	17.1 (9.7)	68.1	1.27 (0.23)	1.49 (0.27)	4.2 (11.1)
Verschil tussen condities		$F(2,19.8) = 0.47$[b]	$F(2,20.6) = 0.93$[c]	$F(2,67) = 0.14$	$F(2,67) = 0.02$	$F(2,20.9) = 1.24$[c]	$F(2,69) = 0.18$	$F(2,69) = 0.33$	$F(2,16.6) = 1.24$[b]
Effectgrootte		$\eta^2 = .019$	$\eta^2 = .031$	$\eta^2 = .004$	$\eta^2 = .000$	$\eta^2 = .040$	$\eta^2 = .005$	$\eta^2 = .010$	$\eta^2 = .079$

Noot. Geen van de verschillen tussen condities waren significant ($p > .05$).
[a]*n* = 70 klassen. [b]Welch F test i.v.m. significante afwijking van homogeniteit van varianties. [c]Welch F test i.v.m. niet voldoen aan aannames chi-kwadraat-toets en significante afwijking van homogeniteit van varianties.

Tabel 7. *Verschillen tussen de onderzoekscondities in schoolkenmerken.*

Onderzoeksconditie	*n*	Schoolgrootte (aantal leerlingen) *M* (*SD*)	Schoolpercentage leerlingen met leerlingewicht[a] *M* (*SD*)	Stedelijkheid buurt (schaal van 1 tot 5) *M* (*SD*)	Gemiddeld inkomen buurt (in k€ per inwoner) *M* (*SD*)	Schoolregio #Noord, #Midden, #Zuid	Denominatie / Onderwijsvisie[b] #OPB, #PC, #RK, #ABZ, #GEV, #OPB/Dalton
Controlegroep	9	200.6 (145.9)	7.7 (8.8)	2.56 (1.67)	22.5 (4.3)	3, 6, 0	4, 3, 1, 0, 0, 1
SoftwareSchool	15	220.6 (145.0)	9.1 (5.1)	2.53 (1.30)	20.8 (2.7)	7, 3, 5	2, 8, 3, 0, 1, 1
SoftwareSchoolThuis	32	228.4 (117.5)	9.4 (9.7)	3.06 (1.50)	21.3 (2.9)	12, 13, 7	6, 11, 12, 1, 1, 1
Totaal	56	221.8 (127.8)	9.0 (8.4)	2.84 (1.47)	21.4 (3.1)	22, 22, 12	12, 22, 16, 1, 2, 3
Verschil tussen condities		$F(2,53) = 0.16$	$F(2,52) = 0.14$	$F(2,53) = 0.85$	$F(2,53) = 0.88$	-[c]	-[c]
Effectgrootte		$\eta^2 = .006$	$\eta^2 = .005$	$\eta^2 = .031$	$\eta^2 = .032$		

Noot. Geen van de verschillen tussen condities waren significant ($p > .05$). # = aantal.

[a]*n* = 55 scholen. [b]OPB = Openbaar, PC = Protestants-Christelijk, RK = Rooms-Katholiek, ABZ = Algemeen Bijzonder, GEV = Gereformeerd Vrijgemaakt. [c]Verschil tussen condities kon niet worden getoetst, omdat niet aan de aannames van chi-kwadraat toets was voldaan (te lage verwachte celfrequentie).

Software-groepen. Het lijkt er dus op dat er in de Controlegroep bij aanvang van het schooljaar wat meer spreiding was in het leesniveau.

Data-analyse

In onze analyses hebben we gekeken naar de leerwinst van de leerlingen in de periode van het onderzoek. We hebben gekeken naar de leerwinst op drie gebieden

- Letterkennis
- Technisch lezen
- Spelling

We hebben hierbij gebruik gemaakt van de resultaten op de eerder genoemde toetsen van VLL en uit het Cito leerlingvolgsysteem.

Omdat de deelnemende leerlingen waren gegroepeerd in klassen en scholen, hebben we gebruik gemaakt van 'multilevel' regressieanalyses. Met dergelijke analyses kunnen variabelen op verschillende niveaus (leerling-, klas-, en schoolniveau) in de analyses worden meegenomen.

Hieronder volgt een korte uitleg van de verschillende analyses die we hebben gedaan. Technische details over de analyses zijn terug te vinden in de bijlage.

Analyses letterkennis

Voor de analyses van de leerwinst in letterkennis gebruikten we *de VLL Letterkennistoetsen Alle Letters*. Vanwege het eerder genoemde plafondeffect op de *Letterkennistoets kern 6*, keken we alleen naar de leerwinst van *kern Start* tot *kern 3*. Dit deden we met een groeimodel.

Analyses technisch lezen

Voor de leerwinst in technisch lezen gebruikten we drie verschillende soorten analyses:

- Een analyse waarin we keken naar de leerwinst over meerdere meetmomenten, namelijk de *VLL Woordleestoetsen* van *kern 1* t/m *kern 6* (groeimodel).
- Een analyse met als eindmeting de *DMT M3* en als beginmeting de *VLL Letterkennistoets Alle Letters kern Start*.
- Een analyse met als eindmeting de *DMT M3* en als beginmeting de *VLL Woordleestoets kern 1*.

Analyses spelling

Voor de leerwinst in spelling gebruikten we twee verschillende soorten analyses:

- Een analyse met als eindmeting de Citotoets *Spelling 3.0 M3* en als beginmeting de *VLL Letterkennistoets Alle Letters kern Start.*
- Een analyse met als eindmeting de Citotoets *Spelling 3.0 M3* en als beginmeting de *VLL Spellingtoets kern 3* (de eerst afgenomen spellingtoets).

Correctie voor invloed van andere factoren

Naast het al dan niet gebruiken van de leerlingsoftware zijn er natuurlijk veel andere factoren die invloed kunnen hebben op de leerwinst van leerlingen in scholen. Voorbeelden zijn klas- en schoolkenmerken zoals het percentage achterstandsleerlingen en de tijd besteed aan taal-/leesonderwijs, en leerlingkenmerken zoals anderstaligheid en geslacht. Voor de invloed van dergelijke factoren werd in de analyses gecorrigeerd. Omdat er tussen scholen verschillen waren in de momenten waarop toetsen werden afgenomen, werd ook hiervoor in de analyses gecorrigeerd. Op deze manier kunnen we er zekerder van zijn dat overgebleven verschillen tussen de onderzoeksgroepen inderdaad veroorzaakt zijn door het al dan niet gebruiken van de leerlingsoftware.

Beantwoorden onderzoeksvragen

Binnen elk van de bovengenoemde analyses werd gekeken naar de antwoorden op de drie onderzoeksvragen:

3. *Wat is de toegevoegde waarde van de leerlingsoftware, wat betreft de leeruitkom-*

sten op het gebied van technische leesvaardigheid en spelling?

4. *Wat is de relatie tussen de mate waarin de leerlingsoftware door leerlingen wordt gebruikt en hun leeruitkomsten?*
5. *In welke mate verschillen deze effecten voor verschillende typen leerlingen?*

Verschil tussen de onderzoekscondities. Onderzoeksvraag 3 werd beantwoord door te kijken naar het verschil tussen de drie onderzoekscondities (SoftwareSchool, SoftwareSchoolThuis, en Controlegroep) in de leerwinst op de verschillende vaardigheden. We onderzochten of leerlingen op scholen waar de leerlingsoftware werd gebruikt meer leerden dan leerlingen op scholen waar de leerlingsoftware niet werd gebruikt.

Effect van intensiteit van softwaregebruik. Om onderzoeksvraag 4 te kunnen beantwoorden, keken we naar het effect van de intensiteit van softwaregebruik op de leerwinst. We gebruikten hiervoor de gegevens uit Digiregie over het aantal gemaakte oefeningen in het softwareonderdeel *Lezen en spelling* in de periode voor de afname van de betreffende toets. We keken apart naar de intensiteit van softwaregebruik op school en de intensiteit van softwaregebruik thuis.

Verschillende effecten voor verschillende leerlingen. Voor onderzoeksvraag 5 onderzochten we of er verschillen waren tussen typen leerlingen in het effect van de onderzoeksconditie. Oftewel: heeft de leerlingsoftware voor bepaalde typen leerlingen meer effect dan voor andere leerlingen? Ook onderzochten we of er verschillen waren tussen typen leerlingen in het effect van de intensiteit van softwaregebruik. Bij deze analyses werd gekeken naar de volgende leerlingkenmerken:

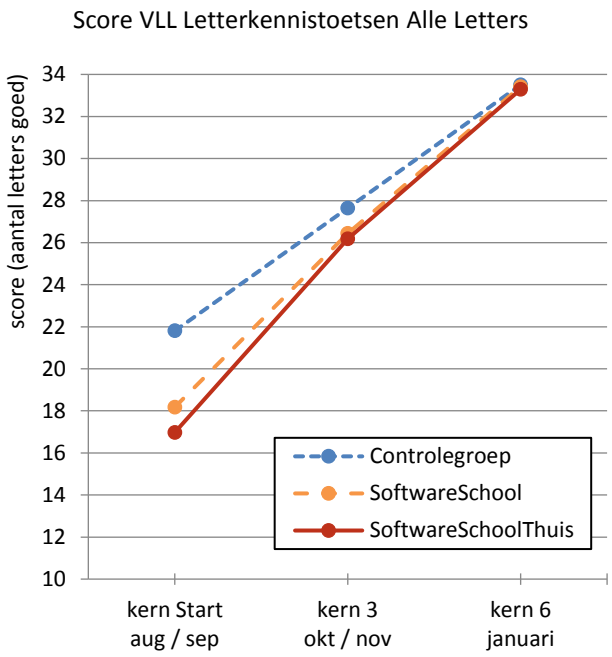
- Geslacht: jongens vs. meisjes
- Leerlinggewicht: leerlingen met leerlinggewicht (laag-opgeleide ouders) vs. leerlingen zonder leerlinggewicht
- Anderstaligheid: anderstalige vs. Nederlandstalige leerlingen
- Beginniveau: maan (gemiddeld) vs. zon (hoog) vs. ster (laag). In de analyses met een eindmeting en een beginmeting keken we ook naar de score op de beginmeting als beginniveau.

Resultaten

Letterkennis

De letterkennis van de leerlingen werd gemeten met de *VLL Letterkennistoetsen Alle Letters*, afgenomen in *kern Start* (augustus/september, aan het begin van de onderzoeksperiode), *kern 3* (oktober/november, halverwege de onderzoeksperiode) en *kern 6* (januari, aan het eind van de onderzoeksperiode). De gemiddelde scores op deze *Letterkennistoetsen* in de drie verschillende onderzoeksgroepen zijn weergegeven in Figuur 6.

In de grafiek is duidelijk te zien dat in alle drie de onderzoekgroepen de gemiddelde score op de laatste toets (kern 6) dicht bij de maximale score van 34 ligt. Met andere woorden, bij de laatste toets trad er een plafondeffect op. Dit was te verwachten, gezien het feit dat de *Letterkennistoets Alle Letters* in kern 6 een beheersingstoets is: op dat moment zijn alle letters aangeboden in de methode. Verder toont de grafiek het eerder genoemde verschil tussen de onderzoeksgroepen in de score op de beginmeting: in de Controlegroep hadden de leerlingen bij de start van het schooljaar een hogere letterkennis dan in de groepen SoftwareSchool en SoftwareSchoolThuis. Vanwege deze twee bevindingen – het beginverschil tussen de onderzoeksgroepen en het plafond effect op de laatste meting – hebben we ervoor gekozen in onze analyse voor letterkennis alleen te kijken naar de groei tussen de eerste twee metingen: de *Letterkennistoets kern Start* en de *Letterkennistoets kern 3*. Aangezien de *Letterkennistoets kern 3* in geen van de onderzoeksgroepen een plafondeffect vertoont,



Figuur 6. Gemiddelde scores *VLL Letterkennistoetsen Alle Letters* per onderzoeksgroep.

kunnen we ervan uitgaan dat er in de periode tussen kern Start en kern 3 in alle onderzoeksgroepen ruimte was voor groei. De analyses van de groei tussen deze twee metingen werden uitgevoerd door middel van een groeimodel (zie voor details de bijlage).

Verschil tussen de onderzoekscondities

Uit de analyses kwam naar voren dat de leerwinst op de letterkennis van kern Start tot kern 3 hoger was op scholen waar de leerlingsoftware gebruikt werd dan op scholen waar deze niet gebruikt werd. Deze hogere leerwinst is ook te zien in Figuur 6. Dit is een aanwijzing dat de leerlingsoftware een positief effect heeft op de letterkennis. Het gaat om een klein effect.

Er werd geen verschil in leerwinst gevonden tussen de SoftwareSchool-groep en de SoftwareSchoolThuis-groep. Er was dus geen aanwijzing dat het thuis oefenen met de software nog een extra bijdrage leverde naast het oefenen op school. Dit kan deels verklaard worden doordat er gemiddeld niet zo veel met de thuissoftware werd gewerkt. Een gemiddeld effect over alle leerlingen is dan moeilijk te vinden.

Effect van intensiteit van softwaregebruik

We vonden geen significante relatie tussen de intensiteit van softwaregebruik (het aantal gemaakte oefeningen *Lezen en spelling* in de software) en de leerwinst in letterkennis. Een invloed van de mate van softwaregebruik op de groei in letterkennis kon dus niet worden aangetoond.

Verschillende effecten voor verschillende leerlingen

We vonden geen verschillen in effecten voor jongens vs. meisjes, voor leerlingen met vs. zonder leerlinggewicht, of voor anderstalige vs. Nederlandstalige leerlingen. Wel vonden we een marginaal significante interactie tussen beginniveau (maan/zon/ster) en de intensiteit van softwaregebruik van de thuissoftware. Bij nadere inspectie bleek dat leerlingen in de ster-aanpak – de leerlingen met het laagste beginniveau – significant meer profiteerden per gemaakte oefening in de thuissoftware dan leerlingen in de maan- of zon-aanpak. Voor leerlingen in de ster-aanpak leverde het thuis werken met de software een positieve bijdrage aan de letterkennis, terwijl dit voor de leerlingen in de maan- en zon-aanpak niet zo was.

Technisch lezen

Zoals aangegeven werden voor het onderzoeken van effecten op de leerwinst in technische leesvaardigheid drie verschillende analysereeksen gebruikt. De resultaten van deze analyses worden hieronder samengevat. Details zijn te vinden in de bijlage.

Verschil tussen de onderzoekscondities

In een van de drie analysereeksen vonden we een significant verschil in leerwinst tussen de onderzoeksgroepen. In deze analyse vonden we in zowel de SoftwareSchool-groep als de SoftwareSchoolThuis-groep gemiddeld een hogere leerwinst dan in de Controlegroep. Dit is een aanwijzing dat de leerlingsoftware een positief effect heeft op de technische leesvaardigheid. Ook hier ging het om een klein effect. In de andere twee analyses met techni-

sche leesvaardigheid vonden we geen significante verschillen in leerwinst tussen de onderzoeksgroepen.

Effect van intensiteit van softwaregebruik

In alle drie de analysereeksen vonden we een duidelijk effect van het aantal gemaakte oefeningen in de leerlingsoftware op school. Leerlingen die op school meer met de software hadden gewerkt, lieten op de toetsen een hogere leerwinst in technische leesvaardigheid zien. In een van de analyses (het groeimodel met de *VLL Woordleestoetsen*) vonden we ook een effect van het aantal gemaakte oefeningen in de thuissoftware: Hier gold dat ook leerlingen die thuis meer met de software hadden geoefend, een hogere leerwinst hadden.

Het groeimodel geeft ook enkele aanwijzingen voor de grootte van het effect van het oefenen in de software. Voor de schoolsoftware vonden we dat voor elke 10 gemaakte oefeningen in het onderdeel *Lezen en spelling* de score op de *Woordleestoetsen* (aantal woorden per minuut) met gemiddeld 0,42 omhoog ging. Ter vergelijking: we vonden een gemiddelde leerwinst per week van 0,86 woorden op de *Woordleestoetsen*. Met 10 oefeningen in de leerlingsoftware werd dus een extra leerwinst behaald die vergelijkbaar is met wat leerlingen gemiddeld in een halve week leerden. Voor de thuissoftware was het effect een stuk kleiner: voor elke 10 gemaakte oefeningen in de thuissoftware ging de score op de *Woordleestoetsen* met gemiddeld 0,15 omhoog. Dit laat zien dat leerlingen over het algemeen meer leerden van de het oefenen met de software op school, dan van het oefenen met de thuissoftware.

Om de relatie tussen het aantal gemaakte oefeningen en de leerwinst goed te kunnen duiden, hebben we gecontroleerd of sterkere en zwakkere lezers ongeveer evenveel oefeningen in de software hadden gemaakt. Dit bleek niet zo te zijn: Sterkere lezers maakten op school gemiddeld meer oefeningen. Er waren echter geen verschillen tussen sterkere en zwakkere lezers in de op school bestede tijd aan de software (sterkere lezers maakten dus meer oefeningen in dezelfde hoeveelheid tijd). We hebben daarom ook gekeken naar het effect van de tijdsbesteding aan de leerlingsoftware (in het onderdeel *Lezen en spelling*) op de leerwinst. Ook hier vonden we een significant effect: leerlingen die op school meer tijd aan de software hadden besteed, hadden een hogere leerwinst in technische leesvaardigheid. Dit is een duidelijke aanwijzing dat het werken met de leerlingsoftware inderdaad bijdraagt aan de leesvaardigheid.

Verschillende effecten voor verschillende leerlingen

Voor de effecten van de intensiteit van softwaregebruik vonden we geen verschillen tussen verschillende typen leerlingen. Alle leerlingen lijken dus evenveel te profiteren van het oefenen met de software.

De verschillen tussen de onderzoekscondities waren in sommige gevallen wel groter voor sommige typen leerlingen dan voor andere. Zo vonden we dat in de SoftwareSchool-groep, en in mindere mate in de SoftwareSchoolThuisgroep, leerlingen in de ster-aanpak (laag beginniveau) een lagere leerwinst hadden dan in de Controlegroep. Deze verschillen in leerwinst voor leerlingen in de ster-aanpak kunnen echter niet duidelijk gerelateerd worden aan het gebruik van de leerlingsoftware, aangezien we de verschillen niet terugzagen bij de effecten van intensiteit van softwaregebruik. Nader onderzoek is nodig om het gevonden verschil goed te kunnen duiden.

Spelling

Voor het onderzoeken van effecten op de leerwinst in spelling werden twee verschillende analysereeksen gebruikt, beide met de Cito-toets *Spelling 3.0 M3* als eindmeting. De resultaten van deze analyses worden hieronder samengevat. Details zijn te vinden in de bijlage.

Verschil tussen de onderzoekscondities

In één van de twee analysereeksen (de analyse met als beginmeting de *VLL Letterkennistoets kern Start*) vonden we een klein significant verschil in leerwinst tussen de onderzoekscondities, in het voordeel van de Software-groepen. Leerlingen op scholen waar de leerlingsoftware werd gebruikt, hadden gemiddeld een hogere leerwinst in spellingvaardigheid. In de andere analyse (de analyse met als beginmeting de *VLL Spellingtoets kern 3*) vonden we geen verschil tussen de onderzoeksgroepen. Dit kan te maken hebben met het feit dat deze analyse een kortere periode omvatte (vanaf oktober/november in plaats van vanaf het begin van het schooljaar).

Effect van intensiteit van softwaregebruik

In geen van beide analyses vonden we een relatie tussen de intensiteit van het softwaregebruik op school en de leerwinst. We vonden dus geen aanwijzingen dat leerlingen die meer met de schoolsoftware hadden geoefend een hogere leerwinst hadden wat betreft spelling. Voor de intensiteit van gebruik van de thuissoftware vonden we, verrassend genoeg, een negatief effect. Leerlingen die thuis meer oefeningen in de software hadden gemaakt, gingen gemiddeld minder vooruit op spellingvaardigheid. Bij nadere inspectie bleek echter dat juist de zwakkere leerlingen (leerlingen met een lager beginniveau op de *Letterkennistoets kern Start*) meer met de thuissoftware hadden gewerkt, wat mogelijk deze negatieve relatie deels kan hebben veroorzaakt. Ook moet opgemerkt worden dat de intensiteitsvariabelen berekend waren als het totaal aantal gemaakte oefeningen in het onderdeel *Lezen en spelling* in de leerlingsoftware. In eerste helft van het schooljaar, waar ons onderzoek zich op richtte, worden in dit onderdeel vooral leesoefeningen aangeboden. Mogelijkerwijs hebben leerlingen thuis niet of weinig met specifieke spellingoefeningen gewerkt. Nader onderzoek zou moeten uitwijzen of het oefenen met specifiek die oefeningen die op spelling gericht zijn wel leidt tot een hogere spellingvaardigheid.

Verschillende effecten voor verschillende leerlingen

Het gemiddelde negatieve effect wat we vonden voor het oefenen in de thuissoftware bleek niet voor alle leerlingen te gelden. De analyses gaven aanwijzingen dat leerlingen met een leerlinggewicht (leerlingen met laagopgeleide ouders) en leerlingen met een lage score op de eerste spellingtoets (*VLL Spellingtoets kern 3*) voor spelling wel een positief effect ondervonden van het oefenen met de thuissoftware.

Discussie en conclusies

De resultaten geven duidelijke aanwijzingen voor een positief effect van de leerlingsoftware op de *technische leesvaardigheid*. Er was zowel een klein significant verschil in leerwinst tussen de onderzoekscondities in het voordeel van de Software-groepen (in een van de analyses), als een significant positief effect van de intensiteit van softwaregebruik op de leerwinst. Leerlingen profiteerden het meeste van het oefenen met de schoolsoftware, maar ook het oefenen met de thuissoftware leek een positieve bijdrage te leveren. Er waren geen verschillen tussen typen leerlingen in de mate waarin zij profiteerden van het oefenen met de software. De leerlingsoftware lijkt dus, wat betreft technische leesvaardigheid, voor alle leerlingen ongeveer even effectief.

Ook voor *letterkennis* en *spelling* vonden we enkele aanwijzingen voor een klein positief effect van de leerlingsoftware: leerlingen in de Software-groepen hadden een hogere leerwinst dan in de Controlegroep. Deze bevindingen werden echter niet ondersteund door een effect van de intensiteit van softwaregebruik. Voor letterkennis en spelling kon het effect van de leerlingsoftware dus niet duidelijk worden aangetoond. Bij spelling vonden we voor het oefenen met de thuissoftware zelfs een negatieve relatie met de leerwinst: leerlingen

die meer met de thuissoftware hadden geoefend gingen minder vooruit op spellingvaardigheid. Mogelijk hebben leerlingen in de thuissoftware niet gewerkt aan specifieke spellingoefeningen; dit moet verder onderzocht worden. Leerlingen met een laag beginniveau en leerlingen met een leerlinggewicht (laagopgeleide ouders) bleken wel baat te hebben bij het oefenen met de thuissoftware. Ook voor letterkennis vonden we dat leerlingen met een laag beginniveau profiteerden van de thuissoftware.

Het feit dat de gevonden verschillen in leerwinst tussen de onderzoeksgroepen niet zo groot zijn, is mogelijk gerelateerd aan het relatief beperkte gebruik van de leerlingsoftware op de deelnemende scholen (gemiddeld minder dan 10 minuten per week per leerling). De positieve relatie tussen de intensiteit van softwaregebruik en de leerwinst in leesvaardigheid laat zien dat er nog meer winst te behalen is als er op scholen meer met de software wordt gewerkt. De aandachtspunten die zijn voortgekomen uit Deel 1 van ons onderzoek kunnen er bij helpen om dit te realiseren.

Beperkingen van het onderzoek

Ondanks dat we zo goed mogelijk hebben gecorrigeerd voor het verschil tussen de onderzoeksgroepen in beginniveau in letterkennis, kunnen we niet uitsluiten dat het hogere beginniveau in de Controlegroep gedeeltelijk mee heeft gespeeld in de gevonden verschillen in leerwinst tussen de onderzoeksgroepen. Dit zou met name bij de leerwinst in letterkennis het geval kunnen zijn (leerlingen met een hogere beginscore hebben minder ruimte voor verbetering van de letterkennis, omdat zij minder letters nog niet kennen). Om de invloed van verschillen in beginniveau volledig uit te sluiten, zou een gerandomiseerd onderzoek nodig zijn waarin scholen of klassen willekeurig aan onderzoekscondities worden toegewezen. Vanwege ethische en praktische bezwaren hebben we daar in dit onderzoek niet voor gekozen.

Een andere beperking die genoemd moet worden is dat de gebruikte begin- en eindmetingen niet altijd goed op elkaar aansloten. Een reden was dat we ons vanwege de grootschaligheid van het onderzoek hebben beperkt tot toetsen die in het normale toetsprogramma van de scholen werden afgenomen. Vooral voor spelling was er geen erg geschikte beginmeting voorhanden (de *Letterkennistoets kern Start* was niet heel hoog gerelateerd met de eindmeting *Spelling 3.0 M3*; de *VLL Spellingtoetsen* vertoonden een plafondeffect).

Verder bleek dat er verschillen waren tussen zwakkere en sterkere lezers in het aantal gemaakte oefeningen in de software. We hebben deels rekening gehouden met deze verschillen door naast het aantal gemaakte oefeningen ook te kijken naar de tijd besteed aan de software. In nadere analyses zouden deze verschillen duidelijker geanalyseerd kunnen worden.

Conclusies en adviezen

Ons onderzoek naar oefensoftware bij het leren lezen bestond uit twee delen: *Hoe zet je het in?* en *Wat levert het op?* Hier volgen kort de conclusies van beide deelonderzoeken.

Deel 1 – Oefensoftware bij het leren lezen: Hoe zet je het in?

Uit de interviewstudie en de procesevaluatie bleek dat het inzetten van oefensoftware in het leesonderwijs over het algemeen niet vanzelf gaat. Het blijkt bijvoorbeeld vaak lastig om alle leerlingen aan de beurt te laten komen, onder andere door beperkte beschikbaarheid van computers. Ook weten leerkrachten vaak niet goed wat er allemaal met de software kan, waardoor mogelijkheden onbenut blijven. Leerkrachten doen er goed aan een plan van aanpak te maken voor het inzetten van de software, waarin concreet wordt gemaakt hoe vaak en op welke momenten de software zal worden gebruikt en hoe de leerlingbeurten worden bijgehouden.

Deel 2 – Oefensoftware bij het leren lezen: Wat levert het op?

De resultaten van de effectstudie laten zien dat de VLL leerlingsoftware een toegevoegde waarde biedt bij de ontwikkeling van de leesvaardigheid. Dit geldt vooral voor de schoolsoftware, maar in mindere mate ook voor de thuissoftware. Zowel zwakkere als sterkere lezers kunnen van de software profiteren.

Voor spelling en letterkennis vonden we ook enkele aanwijzingen voor een effect van de leerlingsoftware, maar hier was de toegevoegde waarde minder duidelijk aanwezig. Wel bleek dat leerlingen met een laag beginniveau of met laagopgeleide ouders wat betreft letterkennis en spelling baat hebben bij het thuis oefenen met de software.

We hebben we met ons onderzoek aangetoond dat oefensoftware een meerwaarde kan hebben in het leesonderwijs in groep 3.

Adviezen voor de onderwijspraktijk

Uit ons onderzoek kunnen een aantal adviezen voor de onderwijspraktijk worden afgeleid voor het gebruik van de leerlingsoftware bij VLL. Hieronder volgen de belangrijkste adviezen:

- Zet de leerlingsoftware in als structureel onderdeel van het onderwijsprogramma, het levert immers een bijdrage aan de leesvaardigheid
- Zet de leerlingsoftware in voor alle leerlingen
- Lees de handleiding van de leerlingsoftware en probeer de software zelf uit
- Maak een plan van aanpak voor het inzetten van de leerlingsoftware in het onderwijs: wanneer, hoe vaak, bijhouden leerlingbeurten
- Gebruik de thuissoftware *niet in plaats van* de schoolsoftware, maar slechts als aanvulling
- Informeer ouders over de thuissoftware en bespreek het werken met de thuissoftware af en toe in de klas
- Voor zwakke lezers en leerlingen met een leerlinggewicht kan het lonen om het gebruik van de thuissoftware wat meer te stimuleren

Literatuur

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72,* 101-130.

Cassady, J. C., & Smith, L.L. (2014). The impact of a structured integrated learning system on first-grade students' reading gains. *Reading & Writing Quarterly, 21*, 361-376.

CFI. (2006). *Nieuwe gewichtenregeling basisonderwijs.* Zoetermeer: CFI.

Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review, 7,* 198-215.

Cito. (2014). *Handleiding: Cito Volgsysteem Primair en speciaal onderwijs Spelling 3.0 Groep 3.* Arnhem: Cito.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading research quarterly, 36,* 250-287.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications.* Second Edition. New York: Routledge.

Irausquin, R. S., Drent, J., & Verhoeven, L. (2005). Benefits of computer-presented speed training for poor readers. *Annals of Dyslexia, 55,* 246-265.

Jongen, I., & Krom, R. (2010). *Drie-Minuten-Toets (DMT): Handleiding.* Arnhem: Cito.

Kennisnet. (2009). *Kwaliteit van digitaal leermateriaal: taalonderwijs op de basisschool.* Zoetermeer: Stichting Kennisnet.

Kennisnet. (2015). *Vier in balans-monitor 2015.* Zoetermeer: Stichting Kennisnet.

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *Wetenschappelijke verantwoording DMT en AVI.* Arnhem: Cito.

Macaruso, P., & Hook, P. E. (2007). Computer assisted instruction: Successful only with proper implementation. *Perspectives on Language and Literacy, 33,* 43-46.

Rasbash, J., Charlton, C., Browne, W.J., Healy, M., & Cameron, B. (2009). *MLwiN Version 2.1.* Bristol, UK: Centre for Multilevel Modelling, University of Bristol.

Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly, 13,* 107-121.

Savage, R. S., Abrami, P., Hipps, G., & Deault, L. (2009). A randomized controlled trial study of the ABRACADABRA reading intervention program in grade 1. *Journal of Educational Psychology, 101*(3), 590-604.

Schneider, D., Chambers, A., Mather, N., Bauschatz, R., Bauer, M., & Doan, L. (2016). The effects of an ICT-based reading intervention on students' achievement in Grade two. *Reading Psychology, 37*(5), 793-831.

Uitgeverij Zwijsen. (2013). *Veilig leren lezen: Gebruikswijzer en verantwoording.* Tilburg: Uitgeverij Zwijsen.

BIJLAGE BIJ DEEL 2

Details data-analyse en resultaten

Data-analyse

Om recht te doen aan de geclusterde datastructuur (leerlingen binnen klassen binnen scholen) hebben we gebruik gemaakt van multilevel regressieanalyses. We hebben hiervoor gebruik gemaakt van de software MLwiN (versie 2.36, zie Rasbash et al., 2009).

In alle analyses hebben we gekeken naar de leerwinst, waarbij onderzoeksconditie als predictor in het model werd opgenomen.

Analysemodellen

Om de leeropbrengsten op de eindmetingen *DMT M3* en *Spelling 3.0 M3* uit het Cito leerlingvolgsysteem te onderzoeken, hebben we gebruik gemaakt van multilevel ANCOVA regressieanalyses met de score op de eindmeting als afhankelijke variabele, en de score op de beginmeting als eerste covariaat. Voor elke Cito eindmeting hebben we, zoals eerder aangegeven, twee sets van analyses uitgevoerd, één met als beginmeting de *VLL Letterkennistoets Alle Letters kern Start* (meting van het niveau aan het begin van het schooljaar), en één met als beginmeting de eerste toets van de betreffende vaardigheid (voor de *DMT* was dat de *VLL Woordleestoets kern 1*, voor *Spelling 3.0 M3* was dat de *VLL Spellingtoets kern 3*). Naast de beginmeting werd ook de tijd tussen begin- en eindmeting meegenomen als covariaat (de afnamedata verschilden tussen scholen).

Met de toetsen van VLL, die meerdere keren werden afgenomen, konden we de ontwikkeling van de leerlingen gedetailleerder onderzoeken. Dit deden we voor de *VLL Letterkennistoetsen* en de *VLL Woordleestoetsen* (niet voor de *VLL Spellingtoetsen*, omdat deze op elke toets een plafondeffect vertoonden). We gebruikten hiervoor multilevel groeimodellen, met de herhaalde metingen als het laagste niveau, en leerling als tweede niveau, aangevuld met klas- en/of schoolniveau als dat nodig bleek. In deze groeimodellen was de herhaald gemeten toetsscore de afhankelijke variabele. Een tijdvarabele werd als covariaat meegenomen, gedefinieerd als het aantal weken tussen de afnamedatum van een herhaalde toets en de afnamedatum van de eerste toets in de reeks. De tijdvariabele voor de eerste toets werd op 0 gezet (vgl. Hox, 2010).

In alle analysemodellen werden covariaten op leerling-, klas- en schoolniveau meegenomen indien deze leidden tot een verbetering van de modelfit. In de groeimodellen werden indien significant ook interacties tussen de covariaten en de tijdvariabele meegenomen in het model.[12] Covariaten werden meegenomen om twee redenen. De eerste reden betreft het corrigeren voor mogelijke andere factoren dan het softwaregebruik die de leerwinst van leerlingen kunnen beïnvloeden. De tweede reden is dat covariaten een deel van de errorvariantie kunnen verklaren en zo de power van de analyses kunnen vergroten.

Elke analysereeks werd uitgevoerd volgens een stappenplan. Eerst werd bepaald welke niveaus (leerlingniveau, klasniveau, schoolniveau) moesten worden meegenomen, en of er random slopes op leerling-, klas- en/of schoolniveau in het model moesten worden meegenomen (bij de groeimodellen ging het om random slopes van de tijdvariabele, bij de ANCOVA regressieanalyses keken we naar random slopes van de beginmeting).[13] Het uitein-

[12] In een groeimodel kunnen significante hoofdeffecten van covariaten worden geïnterpreteerd als verklaringen van verschillen tussen leerlingen in beginniveau. Significante interacties tussen covariaten en de tijdvariabele geven aan dat de groeisnelheid van leerlingen op de gemeten vaardigheid verschilt afhankelijk van de waarde van het covariaat.

[13] Bij het toetsen van de in het model mee te nemen niveaus (random intercepts), werden modellen met en zonder de variantieterm op dat niveau met elkaar vergeleken. Hierbij werd het verschil in modelfit (-2loglikelihood) getoetst met een chi-kwadraat-toets met een significantieniveau van 10%, omdat varianties alleen een positieve waarde kunnen hebben (vergelijkbaar met een éénzijdige toets, zie Hox, 2010). Bij het toetsen van random slopes werd eerst alleen de variantie van de slope aan het model toegevoegd en getoetst met een significantieniveau van 10% (volgens dezelfde redenering als hierboven). Indien de variantie van de slope leidde tot een verbetering van de modelfit, werd vervolgens de covariantie tussen slope en intercept toegevoegd en getoetst, ditmaal

delijke model, met alleen de niveaus en random slopes die significant bijdroegen aan de modelfit, werd gebruikt als het nulmodel van de analysereeks. Vervolgens werden aan dit nulmodel covariaten op leerling-, klas- en/of schoolniveau één voor één toegevoegd en getoetst. Significant bevonden covariaten werden in het analysemodel opgenomen.

Analyse effect van conditie

Het verschil tussen de onderzoeksgroepen in leerwinst werd geanalyseerd door de variabele *conditie* als predictor toe te voegen aan het model met de significante covariaten. De *conditie*-variabele was gecodeerd als twee dummy variabelen: *SoftwareSchool* en *SoftwareSchoolThuis* (de Controlegroep was de referentie-categorie). Significantie werd getoetst door middel van een chi-kwadraattoets van het verschil in -2loglikelihood-waarden van het model met en zonder de twee dummy variabelen. Ook werd de significantie van de afzonderlijke dummy's getoetst door middel van een Wald test (regressiecoëfficiënt gedeeld door de standaarderror).

In de ANCOVA regressieanalyses werd conditie als hoofdeffect toegevoegd. In deze modellen kan een hoofdeffect van *conditie* geïnterpreteerd worden als het verschil tussen de drie onderzoeksgroepen in score op de eindmeting, gecorrigeerd voor score op de beginmeting. Anders gezegd weerspiegelt het hoofdeffect van *conditie* in deze modellen het effect van conditie op de leerwinst.

In de groeimodellen werd naast het hoofdeffect van *conditie* ook de interactie tussen *conditie* en de tijdvariabele opgenomen. In deze modellen weerspiegelt het hoofdeffect van *conditie* de mate waarin de drie onderzoeksgroepen verschillen in beginniveau (score op de eerste meting). De interactie van *conditie* met de tijdvariabele kan geïnterpreteerd worden als het effect van onderzoeksconditie op de groei in de gemeten vaardigheid, oftewel de leerwinst. In de groeimodellen is een effect van onderzoeksconditie op leerwinst dus af te lezen aan een significante interactie tussen *conditie* en de tijdvariabele.

met een significantieniveau van 5% (covarianties kunnen wel negatief zijn).

Analyse effect van intensiteit van softwaregebruik

Om het effect van de intensiteit van het softwaregebruik te toetsen, werden de intensiteitsvariabelen aan het model toegevoegd. Deze variabelen werden toegevoegd *naast* de *conditie*-variabele; we toetsten dus of er binnen scholen waar de software gebruikt werd een relatie was tussen de intensiteit van softwaregebruik en de leerwinst.[14] Steeds werd eerst de variabele getoetst die betrekking had op het gebruik van de software op school, en werd daarna de variabele toegevoegd die betrekking had op het thuisgebruik van de software.

Voor de ANCOVA regressiemodellen werden de intensiteitsvariabelen zo geconstrueerd dat zij betrekking hadden op het softwaregebruik in de periode tussen begin- en eindmeting. De variabele *aantaloefeningenL&S-school* werd bijvoorbeeld berekend als het totaal aantal oefeningen in het onderdeel *Lezen en spelling* in de schoolsoftware tussen begin- en eindmeting.

Voor de groeimodellen werden de intensiteitsvariabelen geconstrueerd als herhaalde, cumulatieve metingen. De intensiteitsvariabelen hadden betrekking op het softwaregebruik tussen het begin van het schooljaar en een bepaalde toetsafname. Een voorbeeld: Voor de vierde toets in de reeks van herhaalde metingen was de waarde van de variabele *aantaloefeningenL&S-thuis* gelijk aan het totaal aantal gemaakte oefeningen *Lezen en spelling* in de thuissoftware tussen het begin van het schooljaar en de afnamedatum van deze vierde toets. De reden dat we hebben kozen voor deze cumulatieve benadering is dat we verwachtten dat intensiteit van softwaregebruik een cumulatief effect heeft: meer softwaregebruik kan effect hebben op zowel de prestatie op de eerstvolgende toets, als op de prestatie op latere toetsen. Opgemerkt moet worden dat, door deze constructie van de intensiteitsvariabelen, deze variabelen gecorreleerd zijn met de tijdvariabele in de groeimodellen. Dit betekent dat de intensiteitsvariabelen alleen

[14] Door de intensiteitsvariabelen naast conditie aan het model toe te voegen, was er minder impact van de scheve verdeling van de intensiteitsvariabelen, die voor alle leerlingen in de controlegroep de waarde 0 hadden.

geïnterpreteerd kunnen worden als ook de tijdvariabele in het model is opgenomen.

Analyses van moderatoreffecten

Om te onderzoeken of de leerlingsoftware meer effect heeft voor bepaalde typen leerlingen, werden moderatoreffecten van leerlingkenmerken geanalyseerd. Hierbij werd gekeken naar de leerlingkenmerken geslacht, leerlinggewicht, anderstaligheid, en beginniveau (maan/zon/ster; in de ANCOVA regressiemodellen ook de score op de beginmeting). Er werd getoetst of er significante interacties waren tussen de leerlingkenmerken en conditie (in de groeimodellen keken we naar driewegsinteracties van leerlingkenmerk, conditie, en de tijdvariabele). Ook werd gekeken naar interacties tussen de leerlingkenmerken en de intensiteitsvariabelen. In de groeimodellen werd hierbij ook de interactie tussen het leerlingkenmerk en de tijdvariabele in het model opgenomen, vanwege de genoemde correlatie tussen de cumulatieve intensiteitsvariabelen en de tijdvariabele.

Ontbrekende gegevens

Niet van alle leerlingen waren alle gegevens beschikbaar. Op sommige scholen waren niet alle toetsen afgenomen, of niet ingevoerd in Digiregie, waardoor er voor sommige leerlingen toetsresultaten ontbraken. Ook kwam het voor dat sommige leerlingen op een school een toets niet hadden gemaakt, bijvoorbeeld door afwezigheid. Verder waren er enkele ontbrekende gegevens in een aantal klaskenmerken (doordat sommige leerkrachten een van de online vragenlijsten niet hadden ingevuld), en in een aantal schoolkenmerken (doordat voor sommige scholen de DUO-gegevens ontbraken).

In elk van de analyses werden die leerlingen meegenomen waarvan zowel de toetsresultaten op de betreffende toetsen, als de waarden op de in het model opgenomen covariaten, beschikbaar waren. In de groeimodel-analyses konden ook leerlingen worden meegenomen die ontbrekende toetsscores hadden op een of meerdere toetsen uit de reeks, indien zij tenminste wel een score hadden voor de eerste toets in de reeks.

Resultaten

Letterkennis

Vanwege een plafondeffect op de laatste *VLL Letterkennistoets* (kern 6), hebben we in onze analyses alleen gekeken naar de leerwinst tussen de *Letterkennistoets kern Start* en de *Letterkennistoets kern 3*. Om deze leerwinst te kunnen vergelijken voor de drie onderzoeksgroepen, hebben we een groeimodel-analyse uitgevoerd met score op de letterkennistoets als herhaald gemeten afhankelijke variabele, en een tijdvariabele als covariaat. Andere significant bevonden covariaten, en interacties tussen covariaten en de tijdvariabele, werden meegenomen in de analyses (zie Tabel 8, analysemodel 1). Vervolgens werd aan het model de *conditie*-variabele toegevoegd, bestaande uit de twee dummy's *SoftwareSchool* en *SoftwareSchoolThuis* (verschil tussen de onderzoeksgroepen op de beginmeting). Daarna werd voor beide *conditie*-dummy's de interactie met de tijdvariabele toegevoegd (verschil tussen de onderzoeksgroepen in de groei in letterkennis), en vervolgens de intensiteitsvariabelen *aantaloefeningenL&S-school* en *aantaloefeningenL&S-thuis* (invloed van het aantal gemaakte oefeningen in het softwareonderdeel *Lezen en spelling* op de behaalde scores op de letterkennistoetsen). In deze analysereeks werden de gegevens van 46 scholen, 59 klassen en 1066 leerlingen meegenomen (een aantal scholen, klassen en leerlingen konden niet in de analyses worden meegenomen omdat hun gegevens van de *Letterkennistoets kern Start* niet beschikbaar waren, of omdat zij ontbrekende gegevens hadden voor een van de meegenomen covariaten). De resultaten van de groeimodel-analyses zijn weergegeven in Tabel 8. De resultaten worden hieronder besproken. Voor effectgroottes (R^2 en R) verwijzen we naar de tabellen.

Effect van conditie

In analysemodel 2 zien we opnieuw het eerder genoemde verschil in beginniveau: het beginniveau van de leerlingen was in de Software-groepen lager dan in de Controlegroep (marginaal significant). Het toevoegen van de interactie tussen *conditie* en de tijdvariabele in analysemodel 3 leidde niet tot een significante verbetering van de modelfit (p = .106), maar voor beide dummy variabelen (*SoftwareSchool* en *SoftwareSchoolThuis*) waren de regressiecoëfficiënten van de interactie met de

tijdvariabele wel significant (*SoftwareSchool*tijd*: p = .043, *SoftwareSchoolThuis*tijd*: p = .015, eenzijdig). Als we in plaats van de twee *conditie*-dummy's één gecombineerde dummy *Software* gebruiken (waarde 1 voor de Software-groepen, waarde 0 voor Controlegroep), vinden we bij het toevoegen van de interactie *Software*tijd* wel een significante verbetering van de modelfit ($X^2(1)$ = 4.37, p =.037).[15] Er was geen significant verschil in leerwinst tussen de twee Software-groepen.

Effect van intensiteit van softwaregebruik

Het toevoegen van de intensiteitsvariabelen in analysemodel 4 en 5 leidde niet tot een significante verbetering van de modelfit (p > .10). Ook de Wald tests van de regressiecoëfficiënten van de intensiteitsvariabelen waren niet significant (p > .10, eenzijdig).

Moderatoreffecten

In analysemodel 3, 4 en 5 toetsten we interacties tussen enerzijds leerlingkenmerken en anderzijds de variabelen *conditie*tijd, aantaloefeningenL&S-school* en *aantaloefeningenL&S-thuis* (deze interacties zijn niet in de tabel opgenomen). Voor *conditie*tijd* en *aantaloefeningenL&S-school* vonden we geen significante moderatoreffecten (p > .10). We vonden wel een marginaal significante interactie tussen beginniveau (maan/zon/ster) en het aantal gemaakte oefeningen in de thuissoftware ($X^2(2)$ = 5.01, p = .082). Bij nadere inspectie bleek dat leerlingen met een laag beginniveau (leerlingen in de ster-aanpak) significant meer profiteerden per gemaakte oefening in de thuissoftware dan leerlingen met een gemiddeld of hoog beginniveau (maan of zon) (B = 0.26, SE = 0.13, p = 0.05). Hoewel er gemiddeld genomen geen significant effect was van de intensiteit van softwaregebruik, was dat er voor de thuissoftware wel voor leerlingen met een laag beginniveau (ster-aanpak).

Technisch lezen - Groeimodel Woordleestoetsen

Het groeimodel met de *VLL Woordleestoetsen* van kern 1 t/m kern 6 werd op dezelfde manier opgezet als het hiervoor genoemde groeimodel met de *VLL Letterkennistoetsen*. De afhankelijke variabele was de score op de herhaald gemeten *Woordleestoetsen*; als covariaat was een tijdvariabele opgenomen, en daarnaast een aantal andere significante covariaten (zie Tabel 9, analysemodel 1). Ook hier werden aan het model achtereenvolgens de *conditie*-variabele, de interactie tussen *conditie* en de tijdvariabele, en de intensiteitsvariabelen toegevoegd. In deze analysereeks werden de gegevens van 55 scholen, 71 klassen en 1321 leerlingen meegenomen. De resultaten van de groeimodelanalyses met de *Woordleestoetsen* worden weergegeven in Tabel 9.

Effect van conditie

Analysemodel 2 laat zien dat er geen significante verschillen tussen de onderzoeksgroepen waren in het beginniveau. Het beginverschil wat we zagen bij de *Letterkennistoetsen*, vonden we dus niet voor de *Woordleestoetsen*. In analysemodel 3 bleek dat er geen significant verschil tussen de onderzoeksgroepen was in leerwinst (p > .10).

Effect van intensiteit van softwaregebruik

Bij het toevoegen van de intensiteitsvariabele *aantaloefeningenL&S-school* in analysemodel 4 vonden we een duidelijke verbetering van de modelfit (p < .001). Leerlingen die op school meer met de software hadden gewerkt, lieten op de toetsen een hogere leesvaardigheid zien. De regressiecoëfficient van 0.42 laat zien dat per 10 gemaakte oefeningen in de schoolsoftware de score op de woordleestoetsen met gemiddeld 0,42 woorden per minuut omhoog ging. Ter vergelijking: de coëfficiënt van de tijdvariabele laat zien dat de wekelijkse leerwinst op de woordleestoetsen gemiddeld 0,86 woorden per minuut was. Met 10 oefeningen in de leerlingsoftware kon dus een extra leerwinst worden behaald die vergelijkbaar is met wat leerlingen in een halve week leerden.

In analysemodel 4 zien we ook dat de interactie tussen de *conditie*-dummy *SoftwareSchoolThuis* en de tijdvariabele significant negatief werd wanneer de intensiteitsvariabele werd toegevoegd aan het model. Dit geeft aan dat leerlingen in de SoftwareSchoolThuis-groep die niet met de leerlingsoftware hadden geoefend een lagere leerwinst hadden dan de leerlingen in de Controlegroep. Voor leerlingen die wel met de software hadden gewerkt, werd leerwinst al gauw hoger (bij 10 gemaakte oefeningen werd het negatieve effect al teniet gedaan). Met andere woorden: leerlingen leerden alleen

[15] Een model met een enkele dummy heeft meer power vanwege het hogere aantal vrijheidsgraden.

van de software als ze er daadwerkelijk mee werkten.

Om de relatie tussen het aantal gemaakte oefeningen en de leerwinst goed te kunnen duiden, hebben we gecontroleerd of sterkere en zwakkere lezers ongeveer evenveel oefeningen in de software hadden gemaakt. Dit bleek niet zo te zijn: Sterkere lezers maakten op school gemiddeld meer oefeningen. Er waren echter geen verschillen tussen sterkere en zwakkere lezers in de op school bestede tijd aan de software ($p < .10$). We hebben daarom ook gekeken naar het effect van een intensiteitsvariabele gebaseerd op de aan de software bestede tijd (tijdsbesteding in uren aan oefeningen in het software-onderdeel *Lezen en spelling*) op de leerwinst (niet in de tabel). Ook voor deze intensiteitsvariabele vonden we een significant positief effect ($B = 1.35$, $SE = 0.21$, $p < .001$).

In analysemodel 5 (Tabel 9) leidde het toevoegen van de intensiteitsvariabele *aantaloefeningenL&S-thuis* tot een marginaal significante verbetering van de modelfit ($p = .063$). De Wald test van de regressiecoëfficiënt laat wel een significant effect zien ($p = .031$, éénzijdig). Dit geeft aan dat ook leerlingen die thuis meer oefeningen in de software hadden gemaakt, hoger scoorden op de leesvaardigheid. Opvallend is dat het effect hier wel een stuk kleiner is dan voor de schoolsoftware. De regressiecoëfficiënt van 0.14 geeft aan dat voor elke 10 gemaakte oefeningen in de thuissoftware de score op de toetsen gemiddeld met 0,14 omhoog gaat.

Moderatoreffecten

We vonden een significante driewegs-interactie tussen *beginniveau* (maan/zon/ster), *conditie* en *tijd* ($X^2(4) = 11.05$, $p = .026$). Bij nadere inspectie bleek dat leerlingen met een laag beginniveau (ster-aanpak) in de SoftwareSchool-groep een significant lagere leerwinst hadden dan in de Controlegroep of de SoftwareSchoolThuis-groep ($B = -0.50$, $SE = 0.21$, $p = .015$). We vonden echter geen significant interactie tussen beginniveau en de intensiteitsvariabelen, dus het gevonden verschil in leerwinst is niet duidelijk gerelateerd aan het gebruik van de software. Ook was het niet zo dat leerlingen uit de ster-aanpak minder met de software hadden gewerkt. Nader onderzoek zou moeten uitwijzen waardoor het gevonden verschil verklaard kan worden.

Technisch lezen – ANCOVA regressieanalyses met DMT M3 als eindmeting

Effecten van de leerlingsoftware op de technische leesvaardigheid gemeten met de *DMT M3* hebben we, zoals eerder aangegeven, onderzocht met multilevel ANCOVA regressieanalyses met de score op de *DMT M3* als afhankelijke variabele en de score op de beginmeting als eerste covariaat. We hebben analyses uitgevoerd met als beginmeting de *VLL Letterkennistoets kern Start* (afgenomen aan het begin van het schooljaar) en analyses met als beginmeting de *VLL Woordleestoets kern 1* (afgenomen eind september/begin oktober). Naast de score op de beginmeting werd de tijd tussen begin- en eindmeting (in dagen) als covariaat in de analyses meegenomen, en daarnaast andere significant bevonden covariaten.

Eindmeting DMT M3 en beginmeting Letterkennistoets kern Start

De resultaten van de multilevel analyses met de score op de *DMT M3* als afhankelijke variabele en de score op de *VLL Letterkennistoets kern Start* als eerste covariaat zijn weergegeven in Tabel 10. In deze analyses werden de gegevens van 45 scholen, 58 klassen, en 1033 leerlingen meegenomen.

Effect van conditie

In een ANCOVA model kan het effect van de *conditie*-variabele op de leerwinst (score op de eindmeting gecorrigeerd voor score op de beginmeting) onderzocht worden door deze variabele als hoofdeffect toe te voegen aan het model. Anders dan bij groeimodellen zijn hier geen interacties met een tijdvariabele nodig. In analysemodel 2 zien we dat het toevoegen van *conditie* niet leidde tot een significante verbetering van de modelfit. Voor zowel SoftwareSchool als SoftwareSchoolThuis was de regressiecoëfficient echter wel significant (SoftwareSchool: $p = .021$, SoftwareSchoolThuis: $p = .037$, eenzijdig). Als we in plaats van de twee *conditie*-dummy's één gecombineerde dummy *Software* gebruiken vinden we wel een marginaal significante verbetering van de modelfit ($X^2(1) = 3.76$, $p = .053$). Gecombineerd met de significante regressiecoëfficiënten zijn er dus goede aanwijzingen voor een effect van conditie, waarbij de leerwinst in beide Software-groepen hoger was dan in de Controlegroep.

Effect van intensiteit van softwaregebruik

Net als in het groeimodel met de woordleestoetsen, vonden we ook in dit ANCOVA model een significant effect van de intensiteitsvariabele *aantaloefeningenL&S-school*: in analysemodel 3 leidde het toevoegen van deze variabele tot een significante verbetering van de modelfit (p = .017). Ook hier lieten leerlingen die meer met de software hadden gewerkt een hogere leesvaardigheid zien. De regressiecoëfficiënt van 0.16 is in dit model moeilijker te interpreteren dan in het groeimodel, omdat de begin- en eindmeting niet op dezelfde schaal gemeten waren.

Voor de intensiteitsvariabele *aantaloefeningen-L&S-thuis* (analysemodel 4) vonden we geen significant effect (p > .10). Er waren in dit analysemodel dus geen aanwijzingen voor een relatie tussen het aantal gemaakte oefeningen in de thuissoftware en de leerwinst.

Moderatoreffecten

Moderatoreffecten waren niet significant.

Eindmeting DMT M3 en beginmeting Woordleestoets kern 1

Tabel 11 toont de resultaten van de analyses met de *DMT M3* als eindmeting (afhankelijke variabele) en de *VLL Woordleestoets kern 1* als beginmeting. In deze analyses werden de gegevens van 55 scholen, 71 klassen, en 1302 leerlingen meegenomen.

Effect van conditie

In analysemodel 2 zien we dat het toevoegen van *conditie* geen significante verbetering van de modelfit opleverde (p > .10). Ook de Wald-tests van de regressiecoëfficiënten waren niet significant (p > .10, eenzijdig). In dit model vonden we dus geen aanwijzingen voor een effect van de onderzoeksconditie.

Effect van intensiteit van softwaregebruik

Het toevoegen van de intensiteitsvariabele *aantaloefeningenL&S-school* in analysemodel 3 leidde tot een marginaal significante verbetering van de modelfit (p = .081). De regressiecoëfficiënt was significant (p = .039, eenzijdig). Dus ook in dit model vonden we een aanwijzing dat leerlingen die op school meer met de software hadden geoefend een hogere leerwinst hadden. Intensiteit van gebruik van de thuissoftware had in dit model geen significant effect op de leerwinst (p > .10, analysemodel 4).

Moderatoreffecten

We vonden een marginaal significante interactie tussen *anderstaligheid* en *conditie* ($X^2(2) = 4.64$, p = .098). Bij nadere inspectie bleek het hier te gaan om een verschil tussen de SoftwareSchool-groep en de SoftwareSchoolThuis-groep. In de SoftwareSchoolThuis-groep hadden anderstalige leerlingen een significant lagere leerwinst dan in de SoftwareSchool-groep (B = -5.41, SE = 2.56, p = .035).

Verder vonden we een marginaal significante interactie tussen *beginniveau* (maan/zon/ster) en *conditie* ($X^2(4) = 8.58$, p = .073). In de Softwaregroepen hadden leerlingen met een laag beginniveau (ster-aanpak) een lagere leerwinst dan in de Controlegroep (*SoftwareSchool*ster*: B = -4.79, SE = 1.94, p = .014; *SoftwareSchoolThuis*ster:* B = -3.24, SE = 1.82, p = .075). Dit moderatoreffect is vergelijkbaar met wat we vonden in het groeimodel met de Woordleestoetsen.

Andere moderatoreffecten werden niet gevonden.

Spelling – ANCOVA regressieanalyses met Spelling 3.0 M3 als eindmeting

Om te onderzoeken of de leerlingsoftware naast effecten op leesvaardigheid ook effecten heeft op de gerelateerde vaardigheid spelling, hebben we multilevel ANCOVA regressieanalyses uitgevoerd met de score op de Cito-toets *Spelling 3.0 M3* als afhankelijke variabele. We hebben analyses uitgevoerd met als beginmeting de *VLL Letterkennistoets kern Start* (afgenomen aan het begin van het schooljaar) en analyses met als beginmeting de *VLL Spellingtoets kern 3* (de eerste spellingtoets, afgenomen in oktober/november). Naast de score op de beginmeting werden de tijd tussen begin- en eindmeting (in dagen), en andere significante bevonden covariaten, in de analyses meegenomen.

Eindmeting Spelling 3.0 M3 en beginmeting Letterkennistoets kern Start

De resultaten van de analyses met de score op de toets *Spelling 3.0 M3* als afhankelijke variabele en de score op de *VLL Letterkennistoets kern Start* als eerste covariaat zijn weergegeven in Tabel 12. In deze analyses werden de gegevens van 41 scholen, 53 klassen, en 929 leerlingen meegenomen.

Effect van conditie

Het toevoegen van *conditie* in analysemodel 2 leidde niet tot een significante verbetering van de modelfit (p > .10). De regressiecoëfficiënt van de

dummy-variabele SoftwareSchoolThuis was echter wel significant (p = .022, eenzijdig): in de SoftwareSchoolThuis-groep scoorden leerlingen significant hoger op spellingvaardigheid dan in de Controlegroep. De regressiecoëfficiënt van SoftwareSchool was marginaal significant (p = .081, eenzijdig); ook hier leek de spellingvaardigheid hoger te zijn dan in de Controlegroep. Wanneer we in plaats van de twee *conditie*-dummy's één gecombineerde dummy *Software* gebruikten, vonden we wel een marginaal significante verbetering van de modelfit ($X^2(1)$ = 3.67, p =.055) en een significante regressiecoëfficiënt voor de dummy *Software* (B = 14.35, SE = 7.32, p = 0.050). Deze resultaten wijzen op een positief effect van de leerlingsoftware op de spellingvaardigheid.

Effect van intensiteit van softwaregebruik

We vonden geen significant effect van de intensiteitsvariabele *aantaloefeningenL&S-school* (analysemodel 3, p > .10). We vonden dus geen aanwijzingen dat leerlingen die meer met de schoolsoftware hadden geoefend hoger scoorden op spelling. Voor de intensiteitsvariabele *aantaloefeningenL&S-thuis* vonden we een significant negatief effect van de intensiteitsvariabele (p = .011). Dus leerlingen die thuis meer oefeningen in de software hadden gemaakt, scoorden lager op spelling. Bij nadere inspectie bleek dat juist de zwakkere leerlingen (leerlingen met een lager beginniveau op de Letterkennistoets kern Start) meer met de thuissoftware hadden gewerkt, wat deze negatieve relatie deels kan hebben veroorzaakt. Verder moet vermeld worden dat de intensiteitsvariabelen berekend waren als het totaal aantal gemaakte oefeningen in het onderdeel *Lezen en spelling* in de leerlingsoftware. In de eerste helft van het schooljaar, waar ons onderzoek zich op richtte, worden in dit onderdeel vooral leesoefeningen aangeboden. Nadere onderzoek zou kunnen uitwijzen of het oefenen met specifiek die oefeningen die op spelling gericht zijn wel leidt tot een hogere spellingvaardigheid.

Moderatoreffecten

We vonden een marginaal significante interactie tussen *leerlinggewicht* (dummy-variabele leergewicht vs. geen leerlinggewicht) en de intensiteitsvariabele *aantaloefeningenL&S-thuis* ($X^2(1)$ = 3.79, p = .052). Leerlingen met een leerlinggewicht, oftewel leerlingen met laagopgeleide ouders, leken meer van het oefenen met de thuissoftware te profiteren (B = 9.87, SE = 5.05, p = .051). Hoewel we gemiddeld genomen een negatief effect vonden van de intensiteitsvariabele *aantaloefeningenL&S-thuis*, was dit effect voor leerlingen met een leerlinggewicht positief. Voor deze leerlingen gold dus dat zij wel degelijk meer vooruit gingen op spelling als zij meer oefeningen hadden gemaakt in de thuissoftware.

Eindmeting Spelling 3.0 M3 en beginmeting VLL Spellingtoets kern 3

Tabel 13 toont de resultaten van de analyses met de score op de toets *Spelling 3.0 M3* als afhankelijke variabele en de score op de *VLL Spellingtoets kern 3* als eerste covariaat. In deze analyses werden de gegevens meegenomen van 48 scholen, 62 klassen, en 1185 leerlingen.

Effect van conditie

Het toevoegen van *conditie* in analysemodel 2 leidde niet tot een significante verbetering van de modelfit (p > .10). Ook de Wald-tests van de regressiecoëfficiënten waren niet significant (p > .10, eenzijdig). In dit model vonden we dus geen effect van conditie.

Effect van intensiteit van softwaregebruik

Net als in het hiervoor besproken model met de *Letterkennistoets kern Start* als voormeting, vonden we ook in dit model geen significant effect van de intensiteitsvariabele *aantaloefeningenL&S-school* (p > .10) en een negatief effect van de intensiteitsvariabele *aantaloefeningenL&S-thuis* (p = .020). Er was dus geen significante relatie tussen het aantal op school gemaakte oefeningen in de software en de leerwinst in spellingvaardigheid, terwijl het aantal thuis gemaakte softwareoefeningen negatief gerelateerd was met de leerwinst in spelling.

Moderatoreffecten

We vonden een marginaal significante interactie tussen *leerlinggewicht* (dummy-variabele leergewicht vs. geen leerlinggewicht) en *conditie* ($X^2(2)$ = 5.31, p = .070). In de SoftwareSchool-groep hadden leerlingen met een leerlinggewicht (leerlingen met laagopgeleide ouders) een significant hogere leerwinst dan in de Controlegroep (B = 36.59, SE = 16.10, p = .023).

Verder vonden we een significante interactie tussen de score op de beginmeting (*VLL Spellingtoets*

kern 3) en de intensiteitsvariabele *aantaloefeningenL&S-thuis* ($X^2(1) = 4.62$, $p = .032$). Leerlingen die laag scoorden op de beginmeting van spellingvaardigheid, profiteerden meer per gemaakte oefening in de thuissoftware ($B = -0.69$, $SE = 0.32$, $p = .031$). Hoewel we gemiddeld genomen een negatief effect vonden van het aantal gemaakte oefeningen in de thuissoftware, was dit effect positief voor leerlingen met een lage spellingvaardigheid.

Tabel 8. *Resultaten multilevel groeimodel VLL Letterkennistoetsen kern Start en kern 3.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + hoofdeffect conditie		*Analysemodel 3:* + conditie*tijd		*Analysemodel 4:* + aantal oefeningen school-software		*Analysemodel 5:* + aantal oefeningen thuissoftware	
Fixed part	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
intercept	18.06***	0.63	19.31***	0.87	21.61***	1.39	21.61***	1.39	21.62***	1.39
tijd (weken)	0.91***	0.04	0.91***	0.04	0.70***	0.11	0.70***	0.11	0.70***	0.11
geslacht:meisje	1.12**	0.43	1.11**	0.43	1.12**	0.43	1.12**	0.43	1.12**	0.43
geslacht:meisje*tijd	-0.11***	0.03	-0.11***	0.03	-0.11***	0.03	-0.11***	0.03	-0.11***	0.03
leerlinggewicht:ja	-1.28**	0.42	-1.31**	0.42	-1.31**	0.42	-1.31**	0.42	-1.33**	0.42
leeftijd (jaren)[a]	1.83**	0.59	1.81**	0.59	1.79**	0.59	1.79**	0.59	1.79**	0.59
leeftijd[a]*tijd	-0.15***	0.04	-0.15***	0.04	-0.14***	0.04	-0.14***	0.04	-0.15***	0.04
zittenblijver	11.94***	1.28	11.99***	1.28	12.05***	1.28	12.05***	1.28	12.05***	1.28
zittenblijver*tijd	-0.59***	0.10	-0.59***	0.10	-0.60***	0.10	-0.60***	0.10	-0.60***	0.10
tijdsbesteding taalonderwijs (uren/dag)[a]	-5.45**	1.81	-5.29**	1.74	-4.83**	1.71	-4.83**	1.71	-4.83**	1.71
tijdsbesteding taalonderwijs[a]*tijd	0.47***	0.14	0.47***	0.14	0.43**	0.14	0.43**	0.14	0.43**	0.14
schoolpercentage leerlinggewicht	-0.06**	0.02	-0.06**	0.02	-0.07**	0.02	-0.07**	0.02	-0.07**	0.02
schoolregio:midden (ref = noord)	0.94*	0.45	0.66	0.49	0.69	0.48	0.69	0.48	0.69	0.48
schoolregio:zuid (ref = noord)	-0.11	0.53	0.00	0.54	0.03	0.53	0.03	0.53	0.03	0.54
conditie:SoftwareSchool			-1.32[†]	0.76	-3.73*	1.64	-3.73*	1.64	-3.74*	1.64
conditie:SoftwareSchoolThuis			-1.37*	0.63	-4.12**	1.44	-4.12**	1.44	-4.12**	1.44
conditie:SoftwareSchool*tijd					*0.22**	*0.13*	*0.22**	*0.13*	*0.22**	*0.13*
conditie:SoftwareSchoolThuis*tijd					*0.25**	*0.12*	*0.25**	*0.12*	*0.26**	*0.12*
aantaloefeningenL&S-school (x10)							*0.00*	*0.03*	*0.00*	*0.03*
aantaloefeningenL&S-thuis (x10)									*-0.02*	*0.04*

Random part[b]	*Var (SE)*	*Var (SE)*	*R²*	*R*	*Var (SE)*	*R²*	*R*	*Var (SE)*	*R²*	*R*	*Var (SE)*	*R²*	*R*
Meetpuntniveau: var intercept	9.71 (0.47)	9.67 (0.46)	.004	.065	9.70 (0.47)	- [c]		9.70 (0.47)	- [c]		9.69 (0.46)	.001	.029
Leerlingniveau: var intercept	37.32 (2.12)	37.37 (2.11)	- [c]		37.31 (2.11)	.002	.041	37.31 (2.11)	- [c]		37.31 (2.11)	- [c]	
Leerlingniveau: var slope	0.05 (0.01)	0.06 (0.01)	- [c]		0.06 (0.01)	.009	.095	0.06 (0.01)	- [c]		0.06 (0.01)	- [c]	
Leerlingniveau: covar slope/intercept	-1.99 (0.14)	-2.00 (0.14)			-1.99 (0.14)			-1.99 (0.14)			-1.99 (0.14)		
Klasniveau: var intercept	11.51 (2.68)	10.45 (2.49)	.093	.305	9.66 (2.34)	.075	.273	9.66 (2.34)	.000	.009	9.68 (2.34)	- [c]	
Klasniveau: var slope	0.07 (0.02)	0.07 (0.02)	- [c]		0.06 (0.02)	.111	.333	0.06 (0.02)	.002	.040	0.06 (0.02)	- [c]	
Klasniveau: covar slope/intercept	-0.85 (0.20)	-0.81 (0.20)			-0.73 (0.18)			-0.73 (0.18)			-0.73 (0.18)		
Modelfit													
-2loglikelihood	11706.23	11702.25			11697.77			11697.76			11697.50		
X^2		$X^2(2) = 3.98$			$X^2(2) = 4.49$			$X^2(1) = 0.00$			$X^2(1) = 0.26$		

Noot. De analysemodellen bevatten drie niveaus: meetpuntniveau (*n* = 2007 datapunten), leerlingniveau (*n* = 1066 leerlingen) en klasniveau (*n* = 59 klassen). De modellen bevatten random slopes op leerling- en klasniveau.
[a]Gecentreerd rond het gemiddelde. [b]Var = variantie, cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).
[†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. (*schuingedrukte* waarden werden eenzijdig getoetst)

Tabel 9. *Resultaten multilevel groeimodel VLL Woordleestoetsen kern 1 t/m kern 6.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + hoofdeffect conditie		*Analysemodel 3:* + conditie*tijd		*Analysemodel 4:* + Aantal oefeningen school-software		*Analysemodel 5:* + Aantal oefeningen thuissoftware	
Fixed part	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Intercept	18.79***	0.51	20.98***	1.46	21.02***	1.47	21.05***	1.46	21.04***	1.46
tijd (weken)	0.81***	0.04	0.81***	0.04	0.86***	0.11	0.86***	0.12	0.86***	0.11
leerlinggewicht:ja	-4.80***	1.20	-4.87***	1.20	-4.87***	1.20	-5.04***	1.20	-5.00***	1.20
leeftijd (jaren)[a]	1.49	0.92	1.49	0.92	1.49	0.91	1.52†	0.92	1.54†	0.92
leeftijd[a]*tijd	-0.14*	0.06	-0.14*	0.06	-0.14*	0.06	-0.14*	0.06	-0.14*	0.06
zittenblijver	17.56***	1.90	17.51***	1.90	17.52***	1.90	17.36***	1.90	17.39***	1.90
zittenblijver*tijd	-0.36**	0.12	-0.36**	0.12	-0.36**	0.12	-0.34**	0.12	-0.33**	0.12
conditie:SoftwareSchool			-3.11†	1.74	-3.06†	1.74	-3.55*	1.74	-3.52*	1.74
conditie:SoftwareSchoolThuis			-2.20	1.57	-2.28	1.58	-2.85†	1.58	-2.90†	1.58
conditie:SoftwareSchool*tijd					0.07	0.13	-0.09	0.14	-0.08	0.14
conditie:SoftwareSchoolThuis*tijd					-0.10	0.12	-0.27*	0.13	-0.28*	0.13
aantaloefeningenL&S-school (x10)							*0.42****	*0.07*	*0.40****	*0.07*
aantaloefeningenL&S-thuis (x10)									*0.14**	*0.08*

Random part[b]	*Var (SE)*	*Var (SE)*	R^2	*R*	*Var (SE)*	R^2	*R*	*Var (SE)*	R^2	*R*	*Var (SE)*	R^2	*R*
Meetpuntniveau: var intercept	21.22 (0.43)	21.22 (0.43)	- [c]		21.22 (0.43)	.000	.004	21.04 (0.43)	.008	.092	21.02 (0.43)	.001	.028
Leerlingniveau: var intercept	130.51 (5.64)	130.48 (5.64)	.000	.017	130.44 (5.64)	.000	.016	130.58 (5.64)	- [c]		130.58 (5.64)	.000	.004
Leerlingniveau: var slope	0.42 (0.02)	0.43 (0.02)	- [c]		0.43 (0.02)	- [c]		0.42 (0.02)	.010	.098	0.42 (0.02)	- [c]	
Leerlingniveau: cov slope/intercept	2.33 (0.26)	2.33 (0.26)			2.33 (0.26)			2.28 (0.25)			2.29 (0.25)		
Klasniveau: var intercept	8.98 (2.84)	8.29 (2.72)	.076	.277	8.31 (2.72)	- [c]		8.24 (2.71)	.008	.090	8.24 (2.71)	.001	.026
Klasniveau: var slope	0.06 (0.02)	0.06 (0.02)	.003	.058	0.05 (0.01)	.087	.295	0.07 (0.02)	- [c]		0.07 (0.02)	.048	.219
Modelfit													
-2loglikelihood	50209.31	50206.18			50201.77			50162.94			50159.47		
X^2		$X^2(2) = 3.13$			$X^2(2) = 4.40$			$X^2(1) = 38.84$***			$X^2(1) = 3.46$†		

Noot. De analysemodellen bevatten drie niveaus: meetpuntniveau (*n* = 7376 datapunten), leerlingniveau (*n* = 1321 leerlingen) en klasniveau (*n* = 71 klassen). De modellen bevatten random slopes op leerling- en klasniveau (op klasniveau werd de covariantie tussen slope en intercept niet meegenomen, omdat deze niet significant bijdroeg aan de modelfit).

[a]Gecentreerd rond het gemiddelde. [b]Var = variantie, cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. (*schuingedrukte* waarden werden eenzijdig getoetst)

Tabel 10. *Resultaten multilevel ANCOVA regressieanalyse met DMT M3 als eindmeting (afhankelijke variabele) en VLL Letterkennistoets kern Start als beginmeting.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + conditie				*Analysemodel 3:* + Aantal oefeningen schoolsoftware				*Analysemodel 4:* + Aantal oefeningen thuissoftware			
Fixed part	*B*	*SE*	*B*	*SE*			*B*	*SE*			*B*	*SE*		
intercept	26.54***	0.92	24.02***	1.54			24.16***	1.58			24.17***	1.58		
score Letterkennistoets kern Start[a]	0.78***	0.04	0.79***	0.04			0.79***	0.04			0.79***	0.04		
tijd tussen begin- en eindmeting (dagen)[a]	0.02	0.05	0.01	0.04			0.00	0.05			0.00	0.05		
geslacht:meisje	-1.14*	0.58	-1.16*	0.58			-1.09†	0.58			-1.09†	0.58		
leerlinggewicht:0.3	-2.18†	1.20	-2.01†	1.21			-2.04†	1.20			-2.05†	1.20		
aantal jaar in Nederland[a]	-1.13**	0.41	-1.14**	0.41			-1.16**	0.41			-1.16**	0.41		
zittenblijver	-6.80*	2.74	-6.91*	2.74			-6.63*	2.73			-6.65*	2.74		
Friestalig	4.55†	2.37	4.81*	2.36			4.43†	2.37			4.44†	2.38		
tijdsbesteding taalonderwijs (uren/dag)[a]	6.45***	1.83	6.43***	1.76			6.32***	1.80			6.32***	1.80		
gebruik andere software (uren/week)	14.99**	5.05	14.47**	4.95			14.30**	4.99			14.31**	4.99		
schoolpercentage leerlinggewicht 0.3	-0.44***	0.10	-0.45***	0.10			-0.48***	0.10			-0.48***	0.10		
conditie:SoftwareSchool			*3.50**	*1.72*			*2.60†*	*1.79*			*2.59†*	*1.80*		
conditie:SoftwareSchoolThuis			*2.83**	*1.58*			*1.60*	*1.70*			*1.62*	*1.71*		
aantaloefeningenL&S-school (x10)							*0.16***	*0.07*			*0.16***	*0.07*		
aantaloefeningenL&S-thuis (x10)											-0.01	0.07		
Random part[b]	*Var (SE)*		*Var (SE)*		*R²*	*R*	*Var (SE)*		*R²*	*R*	*Var (SE)*		*R²*	*R*
Leerlingniveau: var intercept	99.75 (5.47)		99.58 (5.47)		.002	.041	98.76 (5.43)		.008	.091	98.73 (5.43)		.000	.018
Leerlingniveau: var slope	0.21 (0.06)		0.21 (0.06)		- [c]		0.21 (0.06)		.006	.078	0.21 (0.06)		- [c]	
Leerlingniveau: cov slope/intercept	3.69 (0.39)		3.67 (0.38)				3.64 (0.38)				3.64 (0.38)			
Schoolniveau: var intercept	8.15 (2.66)		6.87 (2.37)		.157	.397	7.46 (2.49)		- [c]		7.52 (2.51)		- [c]	
Modelfit														
-2loglikelihood	7694.56		7690.49				7684.79				7684.75			
X^2			$X^2(2) = 4.06$				$X^2(1) = 5.71$*				$X^2(1) = 0.03$			

Noot. De analysemodellen bevatten twee niveaus: leerlingniveau (*n* = 1030 leerlingen) en schoolniveau (*n* = 45 scholen). De modellen bevatten een random slope op leerlingniveau op de beginmeting.
[a]Gecentreerd rond het gemiddelde. [b]Var = variantie, Cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. (*schuingedrukte* waarden werden eenzijdig getoetst)

Tabel 11. *Resultaten multilevel ANCOVA regressieanalyse met DMT M3 als eindmeting (afhankelijke variabele) en VLL Woordleestoets kern 1 als beginmeting.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + conditie				*Analysemodel 3:* + Aantal oefeningen schoolsoftware				*Analysemodel 4:* + Aantal oefeningen thuissoftware			
Fixed part	*B*	*SE*	*B*	*SE*			*B*	*SE*			*B*	*SE*		
Intercept	24.27***	0.47	24.48***	1.23			24.47***	1.27			24.46***	1.26		
score Woordleestoets kern 1[a]	0.87***	0.02	0.87***	0.02			0.87***	0.02			0.87***	0.02		
tijd tussen begin- en eindmeting (dagen)[a]	0.10**	0.03	0.10**	0.03			0.10**	0.03			0.10**	0.03		
leeftijd (jaren)[a]	-1.98***	0.59	-1.98***	0.59			-2.01***	0.59			-1.99***	0.59		
zittenblijver	-7.48***	1.53	-7.48***	1.53			-7.31***	1.53			-7.29***	1.53		
leerlinggewicht:1.2	1.08	1.28	1.10	1.28			1.10	1.28			1.14	1.28		
tijdsbesteding leesonderwijs (uren/dag)[a]	4.81**	1.75	4.95**	1.72			4.80**	1.76			4.79**	1.75		
klaspercentage leerlingewicht 1.2	-26.43***	7.65	-24.73**	7.65			-25.31**	7.82			-25.30**	7.78		
gebruik andere software (uren/week)	5.78*	2.70	5.22[†]	2.72			5.22[†]	2.77			5.21[†]	2.76		
conditie:SoftwareSchool			*0.67*	*1.44*			*0.09*	*1.52*			*0.12*	*1.51*		
conditie:SoftwareSchoolThuis			-0.58	1.33			-1.23	1.41			-1.28	1.41		
aantaloefeningenL&S-school (x10)							*0.10**	*0.06*			*0.10**	*0.06*		
aantaloefeningenL&S-thuis (x10)											*0.04*	*0.06*		
Random part[b]	*Var (SE)*		*Var (SE)*		R^2	*R*	*Var (SE)*		R^2	*R*	*Var (SE)*		R^2	*R*
Leerlingniveau: var intercept	63.99	(3.16)	64.02	(3.16)	- [c]		63.8	(3.15)	.003	.059	63.82	(3.15)	- [c]	
Leerlingniveau: var slope	0.02	(0.01)	0.02	(0.01)	- [c]		0.02	(0.01)	.050	.224	0.02	(0.01)	- [c]	
Leerlingniveau: cov slope/intercept	1.09	(0.14)	1.09	(0.14)			1.09	(0.14)			1.09	(0.14)		
Schoolniveau: var intercept	6.69	(1.85)	6.28	(1.78)	.061	.246	6.86	(1.89)	- [c]		6.74	(1.87)	.017	.132
Modelfit														
-2loglikelihood	9119.75		9118.18				9115.15				9114.84			
X^2			$X^2(2) = 1.57$				$X^2(1) = 3.04$[†]				$X^2(1) = 0.30$			

Noot. De analysemodellen bevatten twee niveaus: leerlingniveau (*n* = 1302 leerlingen) en schoolniveau (*n* = 55 scholen). De modellen bevatten een random slope op leerlingniveau op de beginmeting. [a]Gecentreerd rond het gemiddelde. [b]Var = variantie, cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).

[†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. (*schuingedrukte* waarden werden eenzijdig getoetst)

Tabel 12. *Resultaten multilevel ANCOVA regressieanalyse met Cito Spelling 3.0 M3 als eindmeting (afhankelijke variabele) en VLL Letterkennistoets kern Start als beginmeting.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + conditie			*Analysemodel 3:* + Aantal oefeningen schoolsoftware			*Analysemodel 4:* + Aantal oefeningen thuissoftware		
Fixed part	*B*	*SE*	*B*	*SE*		*B*	*SE*		*B*	*SE*	
intercept	171.78***	2.74	159.41***	6.80		159.43***	6.80		159.72***	6.86	
score Letterkennistoets kern start[a]	2.79***	0.21	2.83***	0.21		2.83***	0.21		2.80***	0.21	
tijd tussen begin- en eindmeting (dagen)[a]	-0.91*	0.37	-0.82*	0.36		-0.83*	0.37		-0.82*	0.37	
leerlinggewicht:ja	-24.66***	5.68	-23.89***	5.67		-23.88***	5.67		-24.79***	5.66	
leeftijd (jaren)[a]	-16.75***	3.85	-16.84***	3.85		-16.84***	3.85		-17.75***	3.85	
tijdsbesteding taalonderwijs (uren/dag)[a]	23.54*	9.60	20.80*	9.41		20.85*	9.42		21.23*	9.51	
conditie:SoftwareSchool			*12.28*[†]	*8.77*		*12.14*[†]	*8.92*		*11.58*[†]	*9.00*	
conditie:SoftwareSchoolThuis			*15.13**	*7.54*		*14.79**	*8.47*		*16.44**	*8.56*	
aantaloefeningenL&S-school (x10)						*0.04*	*0.45*		*0.12*	*0.45*	
aantaloefeningenL&S-thuis (x10)									-1.08*	0.43	
Random part[b]	*Var (SE)*		*Var (SE)*	*R²*	*R*	*Var (SE)*	*R²*	*R*	*Var (SE)*	*R²*	*R*
Leerlingniveau: var intercept	1783.6 (126.8)		1785.8 (126.8)	- [c]		1785.9 (126.8)	- [c]		1773.8 (125.8)	.007	.082
Leerlingniveau: var slope	3.7 (1.6)		3.7 (1.6)	.004	.060	3.7 (1.6)	.000	.009	3.6 (1.5)	.017	.129
Leerlingniveau: cov slope/intercept	-14.4 (6.6)		-14.7 (6.6)			-14.7 (6.6)			-15.1 (6.5)		
Klasniveau: var intercept	249.3 (72.8)		220.1 (66.9)	.117	.342	219.8 (66.8))	.001	.039	226.8 (68.1)	- [c]	
Modelfit											
-2loglikelihood	9761.18		9757.32			9757.31			9750.88		
X^2			$X^2(2) = 3.86$			$X^2(1) = 0.01$			$X^2(1) = 6.43$*		

Noot. De analysemodel bevatten twee niveaus: leerlingniveau (*n* = 929 leerlingen) en klasniveau (*n* = 53 klassen). De modellen bevatten een random slope op leerlingniveau op de beginmeting.
[a]Gecentreerd rond het gemiddelde. [b]Var = variantie, cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).
[†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$. (*schuingedrukte* waarden werden eenzijdig getoetst)

Tabel 13. *Resultaten multilevel ANCOVA regressieanalyse met Cito Spelling 3.0 M3 als eindmeting (afhankelijke variabele) en VLL Spellingtoets kern 3 als beginmeting.*

	Analysemodel 1: Nulmodel + significante covariaten		*Analysemodel 2:* + conditie				*Analysemodel 3:* + Aantal oefeningen schoolsoftware				*Analysemodel 4:* + Aantal oefeningen thuissoftware			
Fixed part	*B*	*SE*	*B*	*SE*			*B*	*SE*			*B*	*SE*		
intercept	170.02***	3.54	170.12***	6.60			170.46***	6.57			170.42***	6.57		
score Spellingtoets kern 3[a]	13.88***	0.85	13.94***	0.85			13.95***	0.85			13.84***	0.85		
tijd tussen begin- en eindmeting (dagen)[a]	0.62*	0.24	0.60*	0.24			0.60*	0.24			0.60*	0.24		
leerlinggewicht:ja	-18.00***	5.38	-17.77***	5.37			-17.74**	5.37			-18.48***	5.36		
anderstalig (behalve Fries)	-12.89*	5.70	-13.47*	5.72			-13.64*	5.72			-13.76*	5.71		
schoolregio:midden (ref = noord)	-2.96	4.63	-3.59	4.64			-3.98	4.64			-3.81	4.63		
schoolregio:zuid (ref = noord)	19.66**	6.19	19.53**	6.02			18.67**	6.07			18.51**	6.07		
stedelijkheid (schaal van 1 tot 5)[a]	4.55**	1.52	4.26**	1.50			4.37**	1.50			4.32**	1.50		
conditie:SoftwareSchool			-5.31	6.84			-3.94	6.98			-4.18	6.98		
conditie:SoftwareSchoolThuis			*2.33*	*5.93*			*4.21*	*6.29*			*5.42*	*6.31*		
aantaloefeningenL&S-school (x10)							-0.53	0.62			-0.40	0.62		
aantaloefeningenL&S-thuis (x10)											-1.56*	0.67		
Random part[b]	*Var (SE)*		*Var (SE)*		*R²*	*R*	*Var (SE)*		*R²*	*R*	*Var (SE)*		*R²*	*R*
Leerlingniveau: var intercept	2043.8	(86.1)	2043.6	(86.1)	.000	.009	2043.6	(86.1)	.000	.002	2033.5	(85.7)	.005	.070
Klasniveau: var intercept	76.4	(58.9)	75.8	(58.0)	.007	.086	73.0	(57.1)	.037	.192	77.0	(58.0)	- [c]	
Schoolniveau: var intercept	25.7	(52.4)	16.8	(49.9)	.344	.587	17.2	(49.3)	- [c]		14.4	(49.4)	.167	.409
Modelfit														
-2loglikelihood	12435.86		12433.14				12432.41				12427.02			
X^2			$X^2(2) = 2.72$				$X^2(1) = 0.73$				$X^2(1) = 6.39$*			

Noot. De analysemodellen bevatten drie niveaus: leerlingniveau (*n* = 1185 leerlingen), klasniveau (*n* = 62 klassen) en schoolniveau (*n* = 48 scholen). De modellen bevatten geen random slopes.
[a]Gecentreerd rond het gemiddelde. [b]Var = variantie, cov = covariantie, R^2 = proportie verklaarde variantie ten opzichte van het voorgaande model. X^2 toetst de verbetering van de modelfit ten opzichte van het voorgaande model. [c]Geen variantie verklaard ($R^2 \leq 0$).
† *p* < .10. * *p* < .05. ** *p* < .01. *** *p* < .001. (*schuingedrukte* waarden werden eenzijdig getoetst.