



Onderzoek naar de effectiviteit van KlankKr8

Eindrapportage

Suzanne Gerritsen, Marijn Vleeskens, Wendy de Geus (Oberon)
Hans Luyten, Lyset Rekers-Mombarg (Universiteit Twente)
Elise de Bree (Universiteit van Amsterdam)

Inhoudsopgave

1	Inleiding en onderzoeksvragen	7
2	Beschrijving van de interventie KlankKr8 op school en onderliggende theorie/rationale	9
	Inzet van KlankKr8 in de klas	9
	Achtergrond van letterkennisbevordering en interventielogica van KlankKr8 in de klas	9
3	Methode van het effectonderzoek	11
3.1	Onderzoeksdesign	11
3.2	Werving, randomisatie en tussentijdse uitval	12
3.3	Deelnemers	14
3.4	Instrumenten	16
3.5	Analysestrategie	21
3.6	Verloop van het onderzoek: communicatie, respons, aanpassingen	22
4	Resultaten	25
4.1	Informatie over interventiebeleving en interventiegebruik	25
4.2	In hoeverre bevordert KlankKr8 de letterkennis in groep 2?	27
4.3	Welke factoren zijn van belang in de leeropbrengsten van KlankKr8?	31
4.4	In hoeverre heeft KlankKr8 een effect op technisch lezen, spelling en rekenen in groep 3? ...	38
5	Samenvatting, discussie en conclusie	43
6	Literatuurverwijzingen.....	47
Bijlage 1	Items van de achtergrondkenmerkenvragenlijst over school, klas en leerlingen (voormeting)	49
Bijlage 2	Items van de implementatievragenlijst voor interventiegroep (nameting 1)	53
Bijlage 3	Items van de zelfgemaakte letterkennistoets (nameting 1).....	55
Bijlage 4	Missing values.....	57
Bijlage 5	Pilot eyetracking	59
Bijlage 6	Aanvullende analyses met pre-test letterkennis in groep 2 als controlevariabele	61

Publieksvriendelijke samenvatting

Met deze studie werd de effectiviteit van het programma KlankKr8 op school bepaald, een letter-klank-interventie gericht op het aanleren en automatiseren van letter-klank-koppelingen als opmaat naar de technische leesvaardigheid. Aan de studie namen kleuters deel uit groep 2. Op de helft van de deelnemende scholen werd KlankKr8 aangeboden. Op de andere helft vond *business as usual* plaats. Scholen waren aan elkaar gematcht op leerlingaantal en schoolweging (d.w.z. sociaaleconomische achtergrond van de schoolpopulatie). De prestaties van de kleuters werden na afloop van de interventieperiode vergeleken op letterkennis en in groep 3 op technische leesvaardigheid en spelling-taalverzorging. Na de interventieperiode was de leerwinst op het gebied van letterkennis hoger bij de groep kleuters die met KlankKr8 had geoefend na controle voor enkele contextkenmerken. Binnen de context van onderwijs kan het effect worden aangemerkt als gemiddeld tot groot (Cohen's d ligt tussen .13 en .19). Bij de kinderen die met KlankKr8 hebben gewerkt is een verband gevonden tussen letterkennistoename en oefentijd: meer oefentijd leidt tot betere letterkennis. Het is opvallend dat het gemiddelde aantal sessies dat kinderen werkten met KlankKr8 beperkt was, en beduidend lager dan beoogd door de makers van KlankKr8, en dat een verschil tussen de controle en interventiegroep toch zichtbaar was. Voor een subset van de leerlingen waren Cito-LVS-gegevens beschikbaar op de nametingen in groep 3 op technische leesvaardigheid en spelling-taalverzorging. Daar waren geen significante verschillen te zien tussen de KlankKr8-groep en de controlegroep. Letterkennis was wel gecorreleerd aan deze uitkomsten. KlankKr8 kan worden ingezet om de vroege letterkennis te bevorderen, maar het bieden van voldoende oefentijd is een voorwaarde voor positieve opbrengsten.

KlankKr8 is onderzocht in het kader van een onderzoeksprogramma van het NRO: effectmeting kansrijke interventies. Het onderzoek is onafhankelijk uitgevoerd door onderzoekers van Oberon, Universiteit Twente en Universiteit van Amsterdam.

1 Inleiding en onderzoeksvragen

Leesvaardigheid is een essentiële vaardigheid in het dagelijks bestaan in onze geletterde maatschappij. De gevolgen van laaggeletterdheid en leesproblemen zijn groot. In Nederland gaat er in het schoolse curriculum daarom veel aandacht uit naar optimale instructie van geletterdheid. De jaarlijkse rapporten van de Onderwijsinspectie laten zien dat de opbrengst van het leesonderwijs een aandachtspunt is, wat ook naar voren komt in de recente nadruk op verwerving van basisvaardigheden (zie Masterplan Basisvaardigheden). De vraag is dus hoe bijgedragen kan worden aan het bevorderen van geletterdheid.

Werken aan een stevige basis is belangrijk, zeker ook wat betreft letterkennis. Letterkennis is een belangrijke voorspeller van latere leesvaardigheid (Van Viersen et al., 2018). Onvoldoende geautomatiseerde letterkennis vormt een belangrijke oorzaak van leesachterstanden, terwijl het versterken van de gevoeligheid voor letter-klankcombinaties juist de leesontwikkeling bevordert (Bonte & Brem, 2024). In het versterken van die letterkennis is het van belang om een hoger tempo aan te houden dan regelmatig wordt gehanteerd in de kleuterklassen (Sunde et al., 2019).

KlankKr8 is een mogelijk kansrijke aanpak gericht op het bevorderen van vroege geletterdheid. Het is een preventieve aanpak om letterkennis te trainen en te automatiseren om zo de leesontwikkeling van jonge kinderen te bevorderen. De interventie bestaat uit drie onderdelen: een klassikaal startprogramma, een serious game (app) waarmee leerlingen zelfstandig oefenen, en het LeerKr8-portaal met informatie voor leerkrachten. Eerder is de effectiviteit van het concept van KlankKr8 al op kleine schaal aangetoond (Aravéna et al., 2018; Verwimp et al., 2023).

Het doel van het huidige onderzoek was om de effectiviteit van KlankKr8 in groep 2 ten opzichte van het reguliere onderwijsprogramma te bepalen, door middel van een grootschalig effectonderzoek. Dit onderzoek was gericht op het onderzoeken van de effectiviteit in daadwerkelijk gebruik in de schoolcontext. In het onderscheid tussen efficacy en effectiveness (Bauer et al., 2015) was ons onderzoek dus met name gericht op effectiveness. We maakten gebruik van een *Randomized Control Trial* waarbij scholen ofwel het reguliere onderwijsprogramma (*business as usual*, BAU) ofwel KlankKr8 volgen. Om de effectiviteit van KlankKr8 in kaart te kunnen brengen onderzochten we of de opbrengsten op letterklankkennis (groep 2 en 3) en woordleesvaardigheid (groep 3) groter zijn voor leerlingen op scholen die in groep 2 met KlankKr8 werken dan voor leerlingen op scholen die het reguliere onderwijsprogramma volgen. We onderzochten daarnaast of die opbrengsten gelijk zijn voor leerlingen met een risico op (aanvankelijke) leesproblemen. Tot slot gingen we na of bepaalde leerling-, school- of implementatievariabelen invloed hebben op de leeropbrengsten van KlankKr8.

Samengevat zijn de onderzoeksvragen de volgende:

1. In hoeverre bevordert KlankKr8 de letterkennis in groep 2?

Op basis van eerder onderzoek (Aravéna et al., 2018; Verwimp et al., 2023) was de verwachting dat kleuters die oefenden met KlankKr8 meer vooruit zouden gaan op letterkennis dan kleuters die het reguliere programma volgden.

2. Welke factoren beïnvloeden de leeropbrengst van KlankKr8?

Binnen de interventiegroep was de verwachting dat met name de hoeveelheid oefen/speeltijd van invloed zou zijn op de uitkomsten. Kleuters die meer oefenden met KlankKr8 werden verwacht meer vooruitgang te laten zien op letterkennis dan kleuters die minder hadden geoefend. Gerelateerd aan deze vraag werd onderzocht of motivatie/aandacht tijdens het werken met KlankKr8 gemeten kon worden bij een subset van leerlingen middels facetracking. We inventariseerden ook of school-, klas- en leerlingkenmerken invloed hadden op de uitkomsten.

3. In hoeverre heeft KlankKr8 een effect op technisch lezen en spelling in groep 3?

Gegeven het belang van letterkennis voor geletterdheid (Bonte & Brem, 2024; Van Viersen et al., 2018) was de verwachting dat de lees- en spellinguitkomsten in groep 3 van kinderen die als kleuters hadden geoefend met KlankKr8 hoger zouden zijn dan kinderen die in de kleuterklas het reguliere programma hadden gevolgd.

2 Beschrijving van de interventie KlankKr8 op school en onderliggende theorie/rationale

Inzet van KlankKr8 in de klas

KlankKr8 bestaat uit drie onderdelen: een klassikaal startprogramma, een serious game (app), en het LeerKr8-portaal. Het **klassikaal startprogramma** is software voor het digibord om de hele klas kennis te laten maken met de letters en klanken uit de Nederlandse taal en bestaat uit 12 lessen van ongeveer 5 minuten. De leerkracht voert het startprogramma zelf uit. De aanbevolen uitvoering is om de lessen verspreid over 3 à 4 weken uit te voeren. Het klassikale startprogramma wordt gevolgd door het inzetten van de **serious game** 'KlankKr8 op school'. In deze game oefenen leerlingen individueel verder met de klanktekenkoppelingen. De aanbevolen hoeveelheid is om leerlingen vier keer per week 10 minuten te laten oefenen of drie keer per week 12 tot 14 minuten. Idealiter is dit verspreid over 16 weken. Leerlingen kunnen ook thuis oefenen met de game. De leerkracht kan in het **LeerKr8-portaal** alle benodigdheden vinden voor de interventie. Het portaal geeft ook per leerling een overzicht van de voortgang en de eventuele knelpunten.

Bij regulier gebruik van KlankKr8 kunnen leerlingen ook thuis oefenen met de game. Vanuit dit onderzoek werd thuisgebruik niet nadrukkelijk gestimuleerd, omdat we de effectiviteit van het schoolgebruik onderzochten. Er zijn gedurende de looptijd van de interventie geen wijzigingen doorgevoerd in de app. Hoewel de makers van KlankKr8 de app normaalgesproken continu doorontwikkelen, is afgesproken deze doorontwikkeling gedurende de looptijd van het onderzoek te pauseren. Dit bevordert de betrouwbaarheid van het onderzoek.

Achtergrond van letterkennisbevordering en interventielogica van KlankKr8 in de klas

KlankKr8 op school beoogt een stevige basis te bieden voor het leren lezen en spellen, specifiek door letter-klankkoppelingen te oefenen. De uitgangspunten van het programma zijn dat er op tijd wordt begonnen met het letteronderwijs, dat er meer aandacht uitgaat naar het aanleren en automatiseren van letterkennis en dat dat in een passend tempo gaat. Dit passende tempo verwijst zowel naar een adaptieve benadering voor individuele leerlingen als naar de noodzaak van een hoger tempo dan wat regelmatig wordt gehanteerd in de kleuterklassen en dat nodig is voor (zwakkere) leerlingen (Sunde et al., 2019). Daarnaast hopen de makers dat KlankKr8 op school bijdraagt aan een gestandaardiseerde aanpak voor letteronderwijs in Nederland. Die is er op dit moment nog niet en dat kan leiden tot onwenselijke verschillen tussen scholen en leerkrachten. KlankKr8 op school is ontwikkeld voor alle leerlingen in het basisonderwijs en kan zowel als preventie of als interventie worden ingezet. KlankKr8 richt zich primair op preventie, voor leerlingen in groep 2, omdat onderzoek toont dat kinderen beter leren lezen wanneer zij alle letter-klankcombinaties al beheersen voordat het formele leesonderwijs begint (Piasta et al., 2012). Door vroeg te starten met letteronderwijs krijgen leerlingen een stevigere basis voor hun leesontwikkeling.

Het verwerven van letterkennis voorafgaand aan het formele leesonderwijs is dus van groot belang. Eerder onderzoek heeft laten zien dat geautomatiseerde letterkennis (letterklankkoppelingen) een kernvoorwaarde is om vloeiend te leren lezen en dat bij kinderen met lees- en spellingproblemen onvoldoende geautomatiseerde letterkennis één van de belangrijkste onderliggende symptomen is (Bonte & Brem, 2024). In KlankKr8 wordt die letterkennis op een adaptieve manier aangeboden.

Het interactieve gamedesign van KlankKr8 biedt kleuters een speelse en veilige omgeving om fouten te maken en te leren. Binnen de serious game wordt de complexiteit stap voor stap opgebouwd. Die geleidelijke opbouw is nodig voor *mastery learning*, waarbij kleuters de kans krijgen om zich kennis volledig eigen te maken. Tegelijkertijd zijn er verschillende manieren om de cognitieve belasting van kleuters in balans te houden: de intrinsieke cognitive load wordt bewaakt door het adaptieve karakter en de visueel en auditief rustige game-omgeving, waarin spelelementen in dienst staan van de kerntaak van de interventie, beperkt de extrinsieke cognitieve load. Zo ontstaat er voldoende cognitieve ruimte om de opgedane kennis te verwerken, consolideren en integreren. Daarnaast past de game zich via het gebruik van een adaptief leersysteem aan op het niveau van de kleuter. Dit sluit aan bij Vygotsky's *zone van naaste ontwikkeling* (Vygotsky, 1978), waardoor leermotivatie en leerrendement versterkt worden. De directe feedback is tot slot bedoeld om gevoelens van frustratie en hulpeloosheid zoveel mogelijk te beperken of te voorkomen.

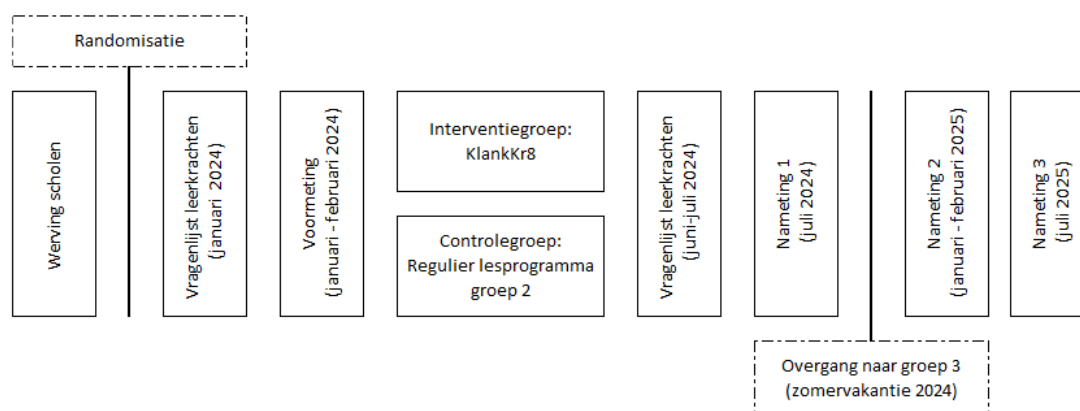
De toevoeging van een startprogramma met digibordlessen en een online leerkrachtportaal is gebaseerd op eerdere bevindingen dat digitale interventies in het onderwijs het meest effectief zijn wanneer zij worden geïntegreerd met klassikale leerkrachtinstructies (Salmerón et al., 2024). Hoewel de game individueel door kleuters kan worden gespeeld, wordt deze daarom gecombineerd aangeboden met klassikale (leerkracht)instructies. Via het leerkrachtportaal kunnen de vorderingen en uitdagingen van de leerlingen worden gemonitord. Op basis van de geautomatiseerde analyse van prestatiegegevens ontvangt de leerkracht bovendien leerlinggerichte adviezen. Dit maakt het mogelijk voor leerkrachten om instructie en feedback in de klas nauwkeuriger af te stemmen op de specifieke behoeften van individuele leerlingen.

KlankKr8 bevat zowel binnen de game als binnen de klassikale startlessen expliciete instructies, gecombineerd met impliciete instructie. Deze combinatie van impliciete en expliciete instructie is namelijk effectiever dan uitsluitend impliciete instructie, omdat expliciete instructies het leerproces doelgericht maken (Verwimp et al., 2023). Binnen zowel de game als in de klassikale lessen zijn de expliciete instructies speels vormgegeven zodat de gameplay en de lessen doelgericht blijven.

3 Methode van het effectonderzoek

3.1 Onderzoeksdesign

Dit effectiviteitsonderzoek betrof een gerandomiseerde gecontroleerde studie (RCT) van een grootschalige implementatie die werd uitgevoerd in groep 2 van het reguliere basisonderwijs. Het speciaal basisonderwijs en speciaal onderwijs waren geen onderdeel van dit onderzoek. Nadat scholen hadden toegezegd om deel te nemen, vond randomisatie plaats tussen deelnemende scholen. Scholen werden at random verdeeld tussen een controlegroep (BAU qua leesonderwijs) en een interventiegroep (KlankKr8). Een school had dus nooit een combinatie van interventie-en controlegroepklassen of leerlingen. Op deze manier werd beoogd *spill-over*-effecten tegen te gaan. Vooraf werd een poweranalyse¹ uitgevoerd, waaruit bleek dat door dit rigoureuze design 70 deelnemende scholen nodig waren en 1680 deelnemende kleuters. Scholen kregen een jaar lang gratis gebruik van KlankKr8 voor hun (nieuwe) groep 2 leerlingen. Voor de scholen in de interventiegroep gold deze periode tijdens de interventieperiode van het onderzoek en voor scholen in de controleconditie gold deze periode na afloop van de interventieperiode. Het tijdspad van de RCT staat hieronder visueel weergegeven. Ethische toetsing voor dit onderzoek heeft plaatsgevonden aan de Universiteit Twente (dossiernummer 230037).



Tijdens de RCT is het programma KlankKr8 van tevoren beschikbaar gesteld aan de interventiescholen, zodat leerkrachten het programma konden klaarzetten, de klassikale lessen konden voorbereiden en het programma en het leerkrachtportaal konden verkennen. De accounts van de leerlingen werden geopend in januari 2024 en afgesloten in de zomervakantie van 2024 (parallel aan de start van de zomervakantie van de school), zodat leerlingen niet thuis zouden doorspelen en er daardoor ruis in de resultaten zou verschijnen. Gerelateerd hieraan werd thuisgebruik van KlankKr8 tijdens de RCT niet gestimuleerd, omdat de metingen effectiviteit van schoolgebruik betreffen.

¹ Deze analyse is in de tussenrapportage in meer detail beschreven.

Het dient te worden opgemerkt dat we in het design bewust hebben gekozen om zoveel mogelijk de natuurlijke setting aan te houden om de daadwerkelijke ongestuurde implementatie in kaart te brengen (zie bijv. Bauer et al., 2015; Harn et al., 2013). We hebben dus geen instructie gegeven over het gebruik van KlankKr8; alleen de aanbieder heeft informatie en aanbevelingen aan school gegeven voor gebruik. Verder is alleen bij een paar scholen op de start/voortgang van de implementatie aangedrongen door de onderwijsadviseur van RID of de onderzoekers, als na enige tijd bleek dat ze nog niet begonnen waren met de digibordlessen of met het oefenen met de app. Verder heeft geen sturing plaatsgevonden, alleen ondersteuning door RID naar behoefte van de scholen, niet anders dan bij andere scholen.

Daarnaast hebben we het aantal metingen beperkt tot een minimum, door zoveel mogelijk gebruik te maken van gangbare toetsen. Tijdens nameting 2 en 3 hebben we bijvoorbeeld alleen leerlingvolgsysteemgegevens (LVS) gebruikt. Ook zijn de testafnames van alle meetmomenten uitgevoerd door de leerkrachten van de scholen zelf, om zoveel mogelijk aan te sluiten bij de natuurlijke setting. Voor de voormeting en voor de eerste nameting kregen de scholen een draaiboek toegestuurd met uitgebreide testinstructies. Er werden online vragensessies georganiseerd om scholen de mogelijkheid te bieden vragen te stellen over de instrumenten, afname en scoring en er konden per e-mail vragen worden gesteld. Voor nameting 2 en 3 (alleen leerlingvolgsysteemgegevens) werd ook een draaiboek met uitleg en instructie toegestuurd en konden scholen bellen of mailen als ze vragen hadden.

3.2 Werving, randomisatie en tussentijdse uitval

Werving

Het onderzoek startte in juli 2022 met een wervingsperiode, die liep tot en met oktober 2023. In 2022 is een oproep aan scholen geplaatst in de nieuwsbrief en op de website van het Nationaal Regieorgaan Onderwijsonderzoek (NRO) om deel te nemen aan het onderzoek. Vervolgens hebben de makers van KlankKr8 in 2023 scholen gericht benaderd per mail met een uitnodiging om deel te nemen. Dit gebeurde in drie tranches, waarbij een gelijke verdeling van scholen werd gehanteerd naar lage (onderste 45%), hoge (bovenste 10%) en overige (middelste 45%) achterstandsscores. Ook zijn in 2023 verschillende andere kanalen ingezet om scholen te werven, namelijk via LinkedIn van de makers van KlankKr8 en Oberon, nieuwsbrieven van belangenclubs (zoals LBBO of AVS) en twee online voorlichtingsbijeenkomsten over het onderzoek. Als een school interesse had in deelname, werd een informatiebrief per mail naar de contactpersoon van de school gestuurd waarin meer details (bijv. planning) over het onderzoek beschreven werden. Voor deelname aan het onderzoek was er voor iedere school een vergoeding beschikbaar mits de deelname volgens de afspraken werd uitgevoerd. De hoogte van de vergoeding was afhankelijk van het aantal deelnemende klassen per school: €1.500 voor scholen met 1 of 2 deelnemende klassen, €2.000 voor scholen met 3 of 4 deelnemende klassen en €2.500 voor scholen met 5 of 6 deelnemende klassen. Om commitment aan het onderzoek te stimuleren, werd gevraagd of scholen een samenwerkingsovereenkomst wilden ondertekenen waarin de verschillende

verantwoordelijkheden van de school, het onderzoeksteam en de makers van KlankKr8 vermeld stonden.

De wervingsperiode resulteerde in 81 aangemelde scholen in oktober 2023. In de fase voorafgaand aan het ondertekenen van de samenwerkingsovereenkomst zijn tien scholen afgefallen (begin november 2023). Redenen hiervoor waren: weinig draagvlak voor deelname onder het personeel, te veel personeelsuitval of andere verwachtingen van het onderzoek dan na de voorlichting. Half november 2023 gaven 71 scholen aan deel te nemen aan het onderzoek naar de effectiviteit van KlankKr8, waarmee de doelstelling van minstens 70 scholen werd bereikt. Omdat de aangeboden versie van KlankKr8 gericht is op het bevorderen van letterkennis voor alle leerlingen werden alle leerlingen van de deelnemende kleuterklassen uitgenodigd voor deelname aan het onderzoek. De uitnodiging, inclusief informatiebrief (in verschillende talen mogelijk) en toestemmingsformulier, werd verzonden naar ouders/verzorgers door de contactpersoon van de deelnemende scholen.

Randomisatie

Na het ondertekenen van de samenwerkingsovereenkomst is met de overgebleven 71 scholen de randomisatie uitgevoerd. Variabelen die werden betrokken bij de randomisatie waren schoolweging en leerlingaantallen. Schoolweging verwijst naar een weging van variabelen die kunnen bijdragen aan de kans op onderwijsachterstanden, waarbij een hogere uitkomst een indicatie is van grotere kans op onderwijsachterstanden (Centraal Bureau voor de Statistiek, 2023). De schoolweging loopt van 20 tot 40, waarbij landelijk de gemiddelde weging iets onder de 30 ligt (Ministerie van Onderwijs, Cultuur en Wetenschap, 2023). De gemiddelde schoolweging² van de deelnemende scholen was 30,89. Scholen zijn via loting aan de interventie- of controlegroep toegewezen. Door een procedure, vastgesteld in overleg met Prof. Dr. Chris van Klaveren³, is bewerkstelligd dat de groepen minimaal verschillen qua schoolweging en aantal leerlingen. Bij de randomisatie zijn scholen ingedeeld in koppels op basis van schoolgrootte en schoolweging. Uit elk koppel is door middel van loting één school ingedeeld in de controlegroep en één in de interventiegroep. De enige school die geen deel uitmaakte van een koppel is door middel van loting ingedeeld in de controlegroep. De controlegroep bestond uit 36 scholen met in totaal 1043 leerlingen. De interventiegroep bestond uit 35 scholen met in totaal 1004 leerlingen.⁴

Tussentijdse uitval

Na de randomisatie zijn in totaal nog zeven scholen uitgevallen. De redenen daarvoor zijn:

- Eén school (controlegroep) kreeg het organisatorisch niet voor elkaar om de tests af te nemen vanwege uitval van personeel en viel eind november 2023 af.
- Eén school (controlegroep) kreeg de toestemming van ouders voor deelname niet georganiseerd, waardoor zij zich moesten afmelden (december 2023).

² De schoolweging is gebaseerd op onder andere het opleidingsniveau van de ouders en het land van herkomst.

³ Amsterdam Center for Learning Analytics, Vrije Universiteit Amsterdam;
<https://sites.google.com/site/chrispbjvanklaveren/>

⁴ Ten tijde van de randomisatie was het leerlingaantal op één school in de interventiegroep onbekend. Het leerlingaantal was op dat moment 987 leerlingen in de interventiegroep.

- Drie scholen (één interventiegroep en twee controlegroep) meldden zich af na ontvangst van de testmaterialen (december 2023/januari 2024). Reden hiervoor was dat het testen van kleuters niet aansloot bij de onderwijsvisie van de school.
- Van twee scholen (controlegroep) is de reden voor uitval onbekend.

Deze zeven scholen (zes controle en een interventie) zijn alle uitgevallen voor de daadwerkelijke start van het onderzoek in januari 2024. Ten tijde van de start van het onderzoek namen 64 scholen deel aan het onderzoek, waarvan 30 in de controlegroep (662 leerlingen) en 34 scholen in de interventiegroep (842 leerlingen).

De zeven scholen die gekoppeld waren aan een school die na de randomisatie is afgehaakt, zijn uitgesloten in analyse van de vergelijking tussen de interventiescholen en de controlescholen. Dit is gedaan om eventuele vertekening door selectieve uitval te voorkomen. Dat betekent dat in de vergelijking naar de effectiviteit van KlankKr8 (onderzoeksvraag 1) 57 scholen van de oorspronkelijke 71 scholen die hadden toegezegd mee te doen aan het onderzoek meegenomen worden. Het gaat om 28 scholen in de interventiegroep en 29 controlescholen, en in totaal 120 groepen/leerkrachten (60 in de interventiegroep en 60 in de controlegroep).

In de analyse naar de relatie tussen de mate van KlankKr8-gebruik en leerwinst op letterkennis (onderzoeksvraag 2) namen we wel de gegevens mee van de zes interventiescholen die voor de vergelijking met de controlescholen zijn uitgesloten vanwege uitval van hun gekoppelde controleschool.

Voor beantwoording van de derde onderzoeksvraag (naar spelling en technische leesvaardigheid in groep 3) zijn de leerlingen uit de controlescholen en interventiescholen vergeleken van de 57 scholen. De steekproef van deze metingen (januari-februari en mei-juni groep 3) is beperkt t.o.v. de voor- en nameting in groep 2. Bij de overgang van groep 2 naar groep 3 bleef een deel van de leerlingen nog een jaar langer in groep 2. Daarbij gaat het om 69 leerlingen van wie we dit weten. Deze leerlingen konden niet worden meegenomen in de vergelijking. Dit heeft onvermijdelijk geleid tot selectieve uitval. Ook was er variatie in de gebruikte leerlingvolgsysteemtoetsen, naast incidenteel niet toetsen van de vorderingen van leerlingen via een LVS. Omdat de verschillende toetsen inhoudelijk niet gelijk zijn en omdat alleen de Cito-leerlingvolgsysteemtoetsen een doorlopende uitkomst hebben, zijn alleen de gegevens van leerlingen geïncludeerd van scholen waarop deze toetsing plaatsvond. Door deze uitval dienen de uitkomsten op onderzoeksvraag 3 hoogstens te worden opgevat als een indicatie.

3.3 Deelnemers

In totaal zijn 64 scholen, 145 klassen en 1504 leerlingen gestart met de uitvoering van het onderzoek. De vergelijking tussen de interventie- en de controlegroep heeft betrekking op 57 scholen. De aantallen scholen, klassen en leerlingen zijn daarbij nagenoeg gelijk verdeeld over beide groepen. De interventiegroep omvatte aan de start van het onderzoek 28 scholen, 60 klassen en 609 leerlingen. De aantallen voor de controlegroep zijn 29 scholen, 60 klassen en

607 leerlingen. Van alle deelnemende scholen is verder de denominatie bekend (Rooms-Katholiek, Protestants-Christelijk of anders) en de stedelijkheid van de locatie. Wat betreft stedelijkheid hanteert het CBS vijf categorieën die variëren van niet stedelijk tot zeer stedelijk. Deze indeling is gebaseerd op de dichtheid van het aantal adressen per km². Op slechts twee achtergrondkenmerken is er sprake van een significant verschil tussen de interventiescholen en de controlescholen (bij een tweezijdig significantieniveau van .05). Dit gaat om het tekort aan onderwijsassistenten in groep 2. Vijf respondenten van interventiescholen (met in totaal 95 leerlingen) geven aan dat dit op hun school van toepassing is. Dit geldt voor slechts één school uit de controlegroep (met slechts zes leerlingen). Verder geven de leerkrachten op de controlescholen aanmerkelijk vaker aan dat ze zich zorgen maken over bepaalde leerlingen in hun klas. Dit gaat om verschillende aspecten van gedrag en ontwikkeling (mondelinge taalverwerving, beginnende geletterdheid, cognitieve vaardigheden, gedrag in het algemeen, internaliserend probleemgedrag, externaliserend probleemgedrag). Tenslotte ligt het percentage meisjes in de interventiegroep lager dan in de controlegroep (46% vs. 51%). Dit verschil is niet significant ($p = .097$; tweezijdig). Nadere details worden gerapporteerd in Tabel 4.2 (paragraaf 4.3).

Tabel 3.1 Schema met aantal deelnemende scholen, klassen en leerlingen per fase en onderzoeksvraag

		Totaal aantal scholen	Aantal interventie scholen	Aantal controle- scholen
November 2023				
	toezegging deelname akkoord	71		
	randomisatie in groepen		35	36
	uitval na randomisatie		-1	-6
Januari 2024				
	start uitvoering onderzoek	64	34 (842 leerlingen)	30 (662 leerlingen)
	verwijderen uit analyse i.v.m. uitval koppelschool		-6	-1
Onderzoeksvraag 1 (effectiviteit Klankkr8 groep 2)	beschikbare scholen voor vergelijkende analyse	57	28 (60 klassen, 609 leerlingen)	29 (60 klassen, 607 leerlingen)
Onderzoeksvraag 2 (effect van speelduur)	alle beschikbare interventiescholen bij start onderzoek		34	
Januari 2025				
Onderzoeksvraag 3 (effect op spelling en technische leesvaardigheid groep 3)	beschikbare scholen voor vergelijkende analyse met Cito LVS	29	15	14

(Beginnend) leesonderwijs op de scholen

Op basis van vragenlijstgegevens van de voormeting (zie 3.4) is geïnventariseerd hoe er in de controleklassen werd gewerkt aan het beginnend leesonderwijs. Beginnende geletterdheid kan verwijzen naar woordenschat, grammatica, vertellen/luisteren, letterkennis, fonemisch bewustzijn. Die inventarisatie liet grote onderlinge variatie zien. Zo gaf een deel (ongeveer 1/4) van de respondenten aan dat er werd gewerkt aan beginnende geletterdheid via monitoring en doelen, maar zonder verdere toelichting *hoe* dat gebeurde. Als er werd verwezen naar methodes of bestaande materialen (ongeveer 2/3), dan ging dit bijvoorbeeld om Onderbouwd, Kleuterplein, Schatkist, werkbladen fonemisch bewustzijn CPB, Bouw! en Logo300. Bij een deel van de klassen (ongeveer 1/8) werd geen gebruik gemaakt van een methode en stond geen informatie over de activiteiten die plaatsvonden. Het is opvallend dat twee respondenten expliciet aangaven niet tevreden te zijn met de aanpak en materialen (ontbreken van een vaste leidraad). Daarnaast gaf een respondent aan dat er per twee weken een nieuwe letter werd aangeboden. Dit tempo is lager dan wordt aangeraden voor het aanleren van letters (Sunde et al., 2019). Er is dus veel variatie in de controlegroep in het vormgeven van het beginnend leesonderwijs. Deze bevindingen verhouden zich direct tot de doelen die de makers van KlankKr8 hebben, omdat ze willen bijdragen aan een uniformer/gestandaardiseerd aanbod voor de basis van het leesonderwijs waarin een hoger tempo van letterkennisoefening plaatsvindt dan nu gemiddeld genomen het geval is.

De uitvraag over het lees-en spellingonderwijs liet zien dat de scholen in zowel de controlegroep als de interventiegroep vanaf groep 3 allemaal gebruikmaakten van een bestaande leesmethode (merendeel van de scholen) of een methode die was gebaseerd op een gefundeerde bestaande methode (ZLKS) (één school in de controlegroep). Deze bevindingen waren hetzelfde voor de spellingmethode. De controle en interventiegroepen maakten dus alle gebruik van een gedegen lees- en spellingmethode.

3.4 Instrumenten

Achtergrondkenmerken: Vragenlijst voormeting

De voormetingvragenlijst ging over school-, klas- en leerlingkenmerken en werd ingevuld door leerkrachten en intern begeleiders (zie Bijlage 1). Het responspercentage op deze vragenlijst was 100%. Uitgevraagde *schoolkenmerken* betroffen het aantal leerlingen, personeelstekorten (leerkrachten, onderwijsassistenten, intern begeleiders), aanwezigheid van een leescoördinator, eigen beoordeling over het kunnen leveren van lees- en spellingscurriculum (op een Likertschaal van 1-5) en een inventarisatie van de gebruikte lees- en spellingmethoden.

Uitgevraagde *leerkrachtkenmerken* betroffen aantal jaar werkervaring in het primair onderwijs, in een kleuterklas en de tevredenheid met het curriculum beginnende geletterdheid (op een Likertschaal van 1-5). Aan de interventiescholen werd gevraagd wat hun inschatting was van de effectiviteit van KlankKr8 bovenop het reguliere leescurriculum, in een cijfer van 1 tot 5.

Uitgevraagde *leerlingen* betroffen leeftijd, beheersing van de Nederlandse taal van ouders/verzorgers en dyslexie bij biologische ouders, broers of zussen (allemaal niet verschillend tussen de controle- en interventiegroepen), geslacht, geschiedenis van VVE, thuistaal, ondersteuning voor beginnende geletterdheid en vragen over zorgen om de ontwikkeling (op een Likertschaal van 1-5). De vragen betroffen de mate waarin er zorgen waren over de mondelinge taalverwerving, de beginnende geletterdheid, de cognitieve leervaardigheid, gedrag in het algemeen, internaliserende problematiek en externaliserende problematiek. Deze vragen werden gecombineerd tot een schaal 'zorgen over leerling'. De Cronbach's alfa van de schaal was .86, wat duidt op een goede betrouwbaarheid.

Achtergrondkenmerken: Leerlingvaardigheden voormeting

Wat betreft leerlingvaardigheden zijn vaardigheden gemeten die invloed kunnen hebben op het kunnen verwerken van letter-klankkennis en leren lezen. De toetsen van Cito Kleuter in Beeld en de foneemdiscriminatietaak werden klassikaal afgenomen, in de volgorde letterkennis, woordenschat, fonologisch bewustzijn. De foneemdiscriminatie-test en snelbenoemtaak werden individueel afgenomen. Alle taken waren pen-en-papier-taken.

Foneemdiscriminatie: Om te bepalen of kinderen de spraakklanken van het Nederlands konden onderscheiden werd een aangepaste versie van de ADIT-A (Auditieve discriminatietest versie A, Crul & Peters 1976) gebruikt. In de huidige aanpassing is een selectie gemaakt van 15 items (i.p.v. de oorspronkelijke 30) en zijn de woorden ingesproken door een van de onderzoekers om de afnameomstandigheden zoveel mogelijk gelijk te houden. De geselecteerde items bevatten klinkerverschillen (bijv. mes-mus), medeklinkerverschillen in de onset (bijv. pan-kan) en medeklinkerverschillen in de coda (bijv. pot-pop). Een voorbeeld is: 'Mes. Bij welk plaatje hoort dat woord?' met keuze uit mes en mus. Per item werd ingevoerd of het antwoord goed/fout was. Op basis daarvan werd het aantal correcte antwoorden bepaald. Afnametijd was ongeveer 5 minuten. Betrouwbaarheid van de taak was onvoldoende (Cronbach's alfa = .443).

Woordenschat: De woordenschattaak van de Cito Kleuter in Beeld werd gebruikt, versie eind groep 2. Het kind hoort een woord in context en ziet drie verschillende plaatjes. Het kind moet het plaatje omcirkelen dat het woord weergeeft. Een voorbeeld is: 'Op welk plaatje heeft de auto een deuk?' met een keuze uit een kapot raam, een deuk en krassen. De taak bestaat uit 16 items. Per item werd ingevoerd of het antwoord goed of fout was. Op basis daarvan werd het aantal correcte antwoorden bepaald (minimum 0-maximum 16). Afnametijd was ongeveer 20 minuten. Betrouwbaarheid was voldoende (Cronbach's alfa = .681).

Fonologisch bewustzijn: De fonologisch bewustzijntaak van de Cito Kleuter in Beeld versie eind groep 2 werd gebruikt. Deze taak bestaat uit drie onderdelen, namelijk rijm (items 1-3), beginfoneem (items 4-9) en synthese (items 10-15). Voor elk item krijgt het kind gesproken informatie en moet het op basis van die informatie uit drie verschillende plaatjes het juiste plaatje selecteren. De antwoordopties bestaan uit het correcte antwoord en afleiders. De taak bestaat uit 15 items. Een voorbeelditem is: 'Wat rijmt op lip? wip, vis, big', waarbij het kind het antwoord moest omcirkelen. Per item werd ingevoerd of het antwoord goed of fout was. Op

basis daarvan werd het aantal correcte antwoorden bepaald (minimum 0 - maximum 15). Afnametijd was ongeveer 20 minuten. Cronbach's alfa van deze taak was goed (alfa = .757).

Snelbenoemen: Snelbenoemen werd gemeten door middel van twee taken van de *Continu Benoemen & Woorden lezen* (van den Bos & Lutje Spelberg, 2007), namelijk plaatjes en kleuren benoemen. Er wordt een kaart met 5 kolommen met respectievelijke bekende plaatjes (boom, eend, stoel, schaar, fiets) en kleuren (zwart, geel, rood, groen, blauw) getoond. Het kind moet zo snel mogelijk de visuele symbolen correct benoemen. De taak is genormeerd vanaf 6 jaar, wat dus ouder is dan de leeftijd van de meeste leerlingen in onze steekproef. We kunnen daarom alleen het aantal goed benoemde afbeeldingen per seconde berekenen en rapporteren. Afnametijd was ongeveer 3 minuten in totaal. De correlatie (r) tussen de twee snelbenoemtaken was .709, $p < .001$, wat een indicatie is van voldoende betrouwbaarheid.

Primaire uitkomstmaat: letterkennis (voormeting en nameting 1)

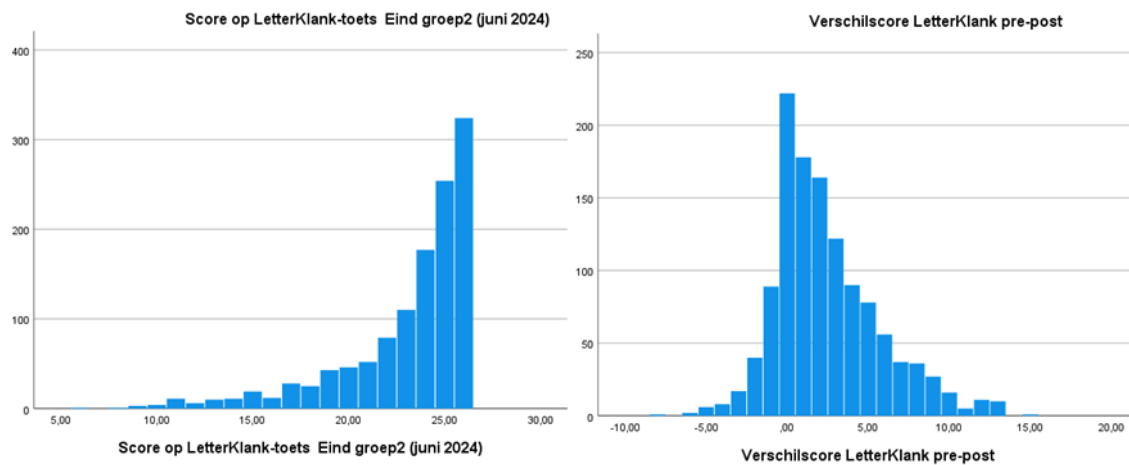
KlankKr8 is gericht op het aanleren van letterkennis. De primaire uitkomstmaat was daarom letterkennis, gemeten op de voormeting en nameting 1. De gehanteerde leerlingvolgssysteemtoets (Cito Letterkennis uit Kleuter in Beeld) bleek richting een plafondscoor te gaan (zie figuur 3.1). Daarom is bij nameting 1 een zelfgemaakte letterkennis toegevoegd. Deze extra letterkennistoets is in het begin van het volgende schooljaar (augustus/september 2024) afgenomen. Hieronder staat toelichting over de twee toetsen.

De *letterkennistaak* van de Cito Kleuter in Beeld versie eind groep 2 werd gebruikt voor de voormeting en nameting 1. De taak bestaat uit 26 items. Het kind werd gevraagd de gesproken klank te koppelen aan de letter, waarbij er per klank drie letters werden aangeboden. Een voorbeelditem is: 'waar zie je de k van koek?' met een keuze uit letters 'k' 't' 'f'. De taak werd klassikaal afgenomen en duurde ongeveer 20 minuten. De taak bestaat uit 26 items, waarvan 22 ééntekenklanken (bijvoorbeeld 'k', 's', 'j') en 4 tweetekenklanken, alle horend bij lange klinkers ('aa', 'uu', 'oo', en 'ee'). Per item werd ingevoerd of het antwoord goed of fout was. Op basis daarvan werd het aantal correcte antwoorden bepaald (minimum 0 - maximum 26).

De verschilscore tussen de voormeting en nameting 1 letterkennis (januari vs. juni 2024) is gebruikt als afhankelijke variabele in de analyse om het effect van KlankKr8 te bepalen. In principe komt ook de score op de post-test hiervoor in aanmerking. Gezien de gerandomiseerde toewijzing mag immers verwacht worden dat er geen sprake van een significant verschil tussen interventiegroep en de controlegroep. De scores op nameting 1 letterkennis laten echter een sterk plafondeffect zien: 26.6% van de leerlingen haalt de maximumscore. Deze tekortkoming is niet van toepassing op de frequentieverdeling van de verschilcores (zie Figuur 3.1).⁵

⁵ Bij zeer sterke plafondeffecten is de kans groot dat lineaire modellen vertekende of misleidende correcties produceren, omdat de regressiehellings de werkelijke relatie niet correct weergeeft. Door te werken met verschilcores is ervoor gezorgd dat we te maken met een uitkomstvariabele die bij benadering wel standaard normaal verdeeld is.

Figuur 3.1 Frequentieverdelingen nameting 1 letterkennis (links) en verschilscore voormeting-nameting 1 (rechts)



De zelfgemaakte letterkennistaak die werd opgesteld voor nameting 1 (zie 3.5) volgde de opzet van de letterkennistaak van Kleuter in Beeld: de gesproken klank moest worden gekoppeld aan een van drie tekens. Waar de Kleuter in Beeld letterkennistaak voornamelijk ééntekenklanken uitvroeg en alleen lange klinkers, werd in deze zelfgemaakte taak nadrukkelijk gebruikgemaakt van complexere twee- en drietekenklanken. Inclusie van deze items werd nodig geacht vanwege de plafondscores op de Kleuter in Beeld-letterkennistoets, waar minder verschillende letters in werden uitgevraagd dan aangeboden in KlankKr8. In de zelfgemaakte letterkennistaak werden 20 items getoetst, waarvan 4 eentekenklanken ('i', 'f', 'g', 'd') en een tweetekenklank voor lange klinker ('oo'), naast 15 complexere koppelingen ('oe', 'ou', 'ei', 'ij', 'ie', 'ui', 'uw', 'eu', 'eeuw', 'auw', 'aai', 'ooi', 'ng', 'nk', 'ch'). Alle items zijn te vinden in de Bijlage 3. Instructie was gelijk aan de Kleuter in Beeld letterkennistoets. Een voorbeelditem is: 'Waar zie je de 'oe' van koek? met keuzes 'oe', 'eo', en 'ou'. De makkelijkere items (items 1, 5, 10, 14, 18) werden afgewisseld met de complexere koppelingen. De positie van het correcte antwoord werd willekeurig bepaald, met 7x positie 1, 6x positie 2 en 7x positie 3. Per item werd ingevoerd of het antwoord goed of fout was. Afnametijd was ongeveer 20 minuten. Cronbach's alfa op alle items was onvoldoende (.519), wat aansloot bij een gemiddelde lage score (9.42 correct). De taak bleek voor veel leerlingen (veel) te moeilijk en (mogelijk daardoor) nauwelijks betrouwbaar. Alleen door negen van de oorspronkelijk twintig items te verwijderen kon een enigszins betrouwbare score worden verkregen. Daarom werd besloten om een score te berekenen op basis van 11 items uit de oorspronkelijke set van 20 (zie ook de Bijlage 3). Dit levert een Cronbach's alfa van .574. Een alternatieve maatstaf voor de betrouwbaarheid van een test is GLB (Greatest Lower Bound to Reliability). Deze biedt een ondergrens voor de werkelijke betrouwbaarheid van een schaal en geldt als nauwkeuriger dan Cronbach's alfa. Voor de toets met de 11 gehandhaafde items komt de GLB uit op .647. Dit betekent dat ongeveer 64.7% van de variantie in de testscore is toe te schrijven aan letterkennis. De betrouwbaarheid is daarmee matig.

Secundaire uitkomstmaten: implementatie (nameting 1)

Dosage. Omdat bekend is dat dosage invloed heeft op interventie-uitkomsten, ook van vroege geletterheidsinterventies (bijv. van der Weijden, 2024) is speelfrequentie opgevraagd bij GameTailors. De verwachting was dat meer spelen leidt tot sterkere toename in letterkennis. De gebruikte variabele is aantal sessies op school vanaf de start van de interventie (eind januari 2024) tot de zomervakantie 2024 (wisselend per schoolregio). De verschillende mogelijke uitkomstvariabelen, naast aantal sessies ook de exacte gespeelde tijd en het aantal bereikte levels, lieten een zeer hoge onderlinge correlatie zien ($>.80$). Daarom is gekozen voor de meest eenvoudige optie in de analyse en het hanteren van aantal sessies.

Ervaringen met KlankKr8. Om zicht te krijgen op de implementatie werd een aantal vragen gesteld aan de groep 2 leerkrachten van de interventiescholen (Bijlage 2). De vragen gingen over hun persoonlijke ervaringen met KlankKr8. Het betrof de inschatting van de effectiviteit van KlankKr8, de tevredenheid over de digibordlessen, de serious game-app waarmee de leerlingen zelfstandig oefenen op een device en het werken met de leerkrachtportaal LeerKr8. Ook is gevraagd of ze KlankKr8 zouden aanbevelen aan collega's, KlankKr8 een verbetering en/of verrijking vinden ten opzichte van het reguliere beginnend leesonderwijsprogramma, en of KlankKr8 heeft geleid tot een toename van de letterkennis van de leerlingen. Met behulp van de 7 items uit deze lijst is een schaal geconstrueerd die aangeeft hoe tevreden een leerkracht is over KlankKr8. Deze schaal heeft een hoge betrouwbaarheid (Cronbach alfa = .92).

Ontwikkelingen. Tenslotte zijn de leerkrachten gevraagd naar andere recente ontwikkelingen op hun school (bijvoorbeeld: Zijn er nog andere recente ontwikkelingen op school wat betreft bemensing in jouw kleuterklas? Zijn er nog andere recente ontwikkelingen op school wat betreft leerlingpopulatie in de kleuterklas?). Het totaal aantal genoemde ontwikkelingen had een maximum van tien.

Secundaire uitkomstmaten: Leerlingvolgsysteemgegevens (nameting 2 en 3)

De verwachting was dat het bevorderen van letterkennis een positieve invloed zou hebben op het technisch lezen en spellen. Daarnaast werd geïnventariseerd of de rekenuitkomsten van de controle- en interventiescholen gelijk waren om de uitkomsten van lezen en spellen te kunnen relateren aan onderwijsuitkomsten in het algemeen. Daarom zijn de gegevens van scholen uit het leerlingvolgsysteem opgevraagd in groep 3 (midden en eind), wat gelijkstaat aan nameting 2 en 3. Gegevens uit het Cito leerlingvolgsysteem waren het meest toegankelijk om te gebruiken, omdat die vaardigheidsscores hebben en genormeerd zijn en omdat daar de voldoende tot goede betrouwbaarheid van bekend is (COTAN, 2010, 2015, zie <https://www.cotandocumentatie.nl>). De uitkomsten van de leerlingvolgsysteemoetsen bestaan uit ruwe scores en vaardigheidsscores.

Woordlezen. De Drieminutentoets (DMT, Krom et al., 2010) werd gebruikt om woordleesvlotheid te meten. Kinderen krijgen achtereenvolgend 3 kaarten te zien met elk 150 woorden erop. De eerste kaart bestaat uit monosyllabische woorden zonder consonantclusters

(bijv. *oom*), de tweede monosyllabische woorden met consonantclusters (bijv. *bloem*) en de derde multisyllabische woorden (bijv. *moeilijk*). Voor elke kaart kreeg het kind de instructie om de woorden zo snel en goed mogelijk op te lezen. Na één minuut werd het lezen per kaart beëindigd. De totale uitleg en afnametijd is 5 minuten. De taak wordt individueel afgenomen. Voor de meeste scholen en kinderen gold dat midden groep 3 kaart 1 en 2 werden afgenomen en eind groep 3 kaart 1, 2 en 3. De ruwe scores (correct aantal gelezen woorden binnen de tijd) zijn gebruikt voor de analyses.

Spellen. De Cito-spellingtaak Leerling in Beeld versie groep 3 is een woorddictée (de Wijs et al., 2010). De taak wordt klassikaal afgenomen. Op basis van de correctscore werd een vaardigheidsscore berekend.

Rekenen. De Cito-rekentaak Leerling in Beeld versie groep 3 meet verschillende onderdelen van rekenvaardigheid, zoals getallen, verbanden, meten en meetkunde en verhouding. Op basis van de correctscore werd een vaardigheidsscore berekend.

3.5 Analysestrategie

Het effect van KlankKr8 werd in kaart gebracht door de leerwinst bij leerlingen op de interventiescholen op het gebied van letterkennis in groep 2 te vergelijken met die van de leerlingen op de controlescholen. Daarbij is gecontroleerd voor eventuele verschillen tussen beide groepen, als daarvan sprake bleek ondanks de gerandomiseerde toewijzing. Deze analyses zijn herhaald voor de extra letterkennistaak.

Daarnaast werd bij de leerlingen op de scholen die met KlankKr8 hebben gewerkt nagegaan in hoeverre de hoeveelheid leerwinst samenhang met het aantal gespeelde KlankKr8-sessies. Hierbij werd gecontroleerd voor zoveel mogelijk leerling-, klas-, en schoolkenmerken.

Tenslotte werd bepaald of in groep 3 significante verschillen konden worden waargenomen tussen beide groepen op diverse leeruitkomsten. Dit betreft toetsen uit het Cito-leerlingvolgsysteem (LVS). Daarbij ging het om technisch lezen, spelling-taalverzorging en rekenen-wiskunde. De LVS-toetsen zijn afgenomen in de periodes januari-februari (M-toetsen) en mei-juni (E-toetsen).

Op basis van standard errors die de data-analyses opleveren kan een minimal detectable effect size worden bepaald. Anders gezegd: hoe groot had het verschil in de data minimaal moeten zijn om als statistisch significant te kunnen worden aangemerkt?

Om rekening te houden met het gegeven dat de leerlingen zijn gegroepeerd binnen scholen is steeds gebruik gemaakt van multilevel modelling. De resultaten van de multilevel-analyses die in dit rapport gepresenteerd worden, hebben allemaal betrekking op random-intercept modellen waarin naast het leerlingniveau alleen het schoolniveau wordt onderscheiden. Resultaten van analyses waarin eveneens het klasniveau is onderscheiden worden niet gerapporteerd. In enkele gevallen leidde het fitten van modellen met zowel school- als klasniveau tot convergentieproblemen. In de overige gevallen bleek de variantie op klasniveau

altijd bescheiden en statistisch niet-significant. Indien convergentie-problemen geen rol speelden, leverden modellen met klasniveau dezelfde regressiecoëfficiënten op als modellen waarin dit niveau ontbreekt.

3.6 Verloop van het onderzoek: communicatie, respons, aanpassingen

Communicatie en respons

In het uitvoeren van het onderzoek is actief ingezet op communicatie met scholen. Om scholen betrokken te houden bij het onderzoek, werd maandelijks een nieuwsbrief verzonden met inhoudelijke informatie, zoals informatie over het belang van het onderzoek of van leesvaardigheid en kennis over adaptieve online middelen voor leren. Er stond ook informatie in over de stand van zaken rondom het onderzoek en waardering voor de deelname van de scholen. Tot slot stond er praktische informatie in over wat nodig was of wat binnenkort nodig zou zijn als vervolgstap voor het onderzoek. Daarnaast was contact met onderzoekers van Oberon laagdrempelig, waardoor betrokkenen altijd contact konden opnemen met het onderzoeksteam. Onze inschatting is dat deze actieve communicatiestrategie en de beperkte onderzoeks-/testlast bij hebben gedragen aan de hoge respons en lage uitval gedurende het onderzoek.

De aantallen scholen verschillen per analyse. Wanneer een vergelijking werd gedaan tussen de interventiegroep en de controlegroep gaat het om 57 scholen (respectievelijk 28 en 29 scholen). Daarnaast zijn er zes scholen die wel met KlankKr8 hebben gewerkt, maar die gekoppeld waren aan uitgevallen controlescholen. In de analyses naar het *gebruik* van KlankKr8 worden deze scholen wel meegenomen (totaal 34 scholen).

In de 57 scholen zaten in schooljaar 2024-25 in totaal 1270 leerlingen in groep 2. In de controlegroep gaat het om 633 leerlingen in de interventiegroep om 637 leerlingen. Hiervan konden 49 leerlingen niet worden meegenomen in de vergelijkende analyse omdat ze de toets letterkennis aan het einde van het schooljaar niet hebben gemaakt. Het gaat om 24 leerlingen uit de controlegroep en 25 uit de interventiegroep. Tenslotte zijn nog vijf leerlingen buiten de vergelijking gelaten omdat zij de pre-test in januari niet hebben gemaakt. Dit gaat om drie leerlingen uit de controlegroep en twee leerlingen uit de interventiegroep. In Tabel 3.2 staat een overzicht van deelnemende scholen en uitval per groep.

Tabel 3.2 Omvang interventiegroep en controlegroep voor en na uitval

	Interventiegroep	Controlegroep	Totaal
Aantal scholen	28	29	57
Aantal leerlingen in groep 2	637	633	1270
Uitval leerlingen			
Geen post-test letterkennis	25	24	49
Geen pre-test letterkennis	3	2	5
Meegenomen in de analyses	609	607	1216

Aanpassingen in onderzoek t.o.v. de aanvraag

Er was een aantal aanpassingen in de uitvoering van het effectonderzoek ten opzichte van de oorspronkelijke aanvraag. Die aanpassingen hebben met name te maken met de keuze om toetsing aan te laten sluiten bij de reguliere toetsen uit het leerlingvolgsysteem, om zo aan te sluiten bij schoolpraktijken (bijv. Bauer et al., 2015). Dit betekent dat we letterkennis niet hebben gemeten door middel van zowel een passieve als een actieve letterkennistaak, maar alleen met de passieve letterkennistaak van Cito (voormeting en nameting 1, zie 3.4 en tussenrapportage). Omdat in eerder onderzoek naar de effectiviteit van KlankKr8 bij kleuters ook was gekozen voor een passieve letterkennistaak (Verwimp et al., 2023) leek dit geen probleem. Ook hebben we woordlezen gemeten door middel van de Drie-Minuten-Test die gangbaar is in het LVS. We hebben daarom geen aparte (ongetimede) leesaccuratesseset afgenomen. Uit eerder onderzoek is bekend dat letterkennis (latente variabele op basis van letterkennisproductie- en herkenning) sterker is gerelateerd aan woordleesvlotheid (fluency, van Viersen et al., 2018) dan aan accuratesse op een ongetimede taak. Daarnaast werden de leestaken midden en eind groep 3 afgenomen, dus met al een flinke hoeveelheid leesinstructie. We denken dus dat de meting alsnog inhoudelijk passend is, naast dat het aansluit bij de praktijk van school.

De keuze voor aansluiting bij de toetspraktijk en het LVS van school had ook onvoorziene gevolgen voor nameting 2 en 3. We beoogden de LVS-gegevens van technisch lezen en spellen op te vragen. Door de veranderingen in het LVS-landschap is er sprake van meerdere LVS-systemen. Niet alle uitkomsten kunnen even makkelijk worden geïnterpreteerd en worden vergeleken met elkaar, wat ook op de NRO-EKI-bijeenkomsten duidelijk werd voor alle lopende projecten. Omdat de DMT-uitkomsten en de Cito LVS het meest frequent worden gebruikt, hebben we op de follow-up meting alleen datapunten met deze uitkomstmaten geïnccludeerd.

Een andere aanpassing ten opzichte van het geplande design was dat we een extra passieve letterkennistaak hebben toegevoegd. Ten tijde van de nameting waren de gegevens van de voormeting net beschikbaar. Daaruit bleek dat de uitkomsten van de passieve letterkennistaak scheef naar rechts verdeeld waren en dat het raadzaam was een extra letterkennistaak toe te voegen met meer letters (zie 3.4 voor informatie). Daarom heeft na de zomervakantie van 2024, toen de deelnemende leerlingen net in groep 3 waren begonnen, een extra meting plaatsgevonden.

Daarnaast hebben we van de leerlingen geen informatie over sociaal-economische status (SES) kunnen uitvragen. De informatie over de leerlingen hebben we verkregen via de leerkrachten en zorgcoördinatoren. Zij hebben geen gegevens over SES. We hebben daarom alleen informatie over de schoolweging, welke is betrokken in de randomisatie. Wel hebben we via de leerkrachten achtergrondgegevens opgevraagd zoals thuistaal en VVE-indicatie. De VVE-indicatie zien we als een proxy voor SES.

Een laatste aanpassing die we hebben gedaan, is dat we na een uitgebreide pilot besloten hebben dat eyetracking niet geschikt is om de flow bij kinderen in groep 2 te meten. De reden

daarvan is dat de concentratie van een kind tijdens het spelen met KlankKr8 met toegepaste trackingsmethode(n) – als proxy voor de flow – niet valide gemeten kan worden. De persoonlijk observatie van de aanwezige onderzoeker van het speelgedrag van het kind bleek daarvoor meer geschikt. Jonge kinderen zijn tijdens het spelen op de tablet in een rustige ruimte op school heel erg beweeglijk en makkelijk afleidbaar. Ze kijken zeker niet de gehele tijd naar het scherm, maar zijn tegelijkertijd wel geconcentreerd: ze weten wat ze (moeten) doen of waar ze moeten klikken voor het juiste antwoord. Om redenen van privacy hebben we geen toestemming gekregen om videoregistraties van de leerlingen te maken en te analyseren. De beschikbare backenddata van KlankKr8 levert betrouwbaardere informatie op over het klikgedrag van een leerling dan gehanteerde trackingsmethode (face-tracking). In Bijlage 5 staat meer informatie over de stappen van de pilot.

4 Resultaten

We beschrijven hier eerst het verloop van de implementatie (4.1). In de daaropvolgende paragrafen (4.2, 4.3 en 4.4) beantwoorden we de onderzoeksvragen:

1. In hoeverre bevordert KlankKr8 de letterkennis in groep 2?
2. Welke factoren beïnvloeden de leeropbrengst van KlankKr8?
3. In hoeverre heeft KlankKr8 een effect op technisch lezen en spelling in groep 3?

4.1 Informatie over interventiebeleving en interventiegebruik

Evaluatie van het leerproces

We hebben geprobeerd scholen het programma KlankKr8 te laten volgen zoals voorgeschreven door de makers van KlankKr8 (zie Hoofdstuk 2), behalve dat we kinderen niet stimuleerden om thuis te oefenen. We zagen echter dat een aantal scholen was gestart met het individueel oefenen voordat de klassikale sessies hadden plaatsgevonden. Een deel van de scholen is later dan gepland gestart met individueel oefenen, waardoor de leerlingen op deze scholen minder dan 16 weken individueel met de serious game hebben gewerkt. De gemiddelde speeltijd per leerling per week op school (berekend over 16 weken) wordt gerapporteerd in Hoofdstuk 4 en oefentijd/speeltijd wordt geïncorporeerd in de analyses. Wat betreft thuisgebruik bleek dat op 24 scholen leerlingen thuis hebben geoefend. We controleren voor thuisspeeltijd in de analyses.

De groep 2 leerkrachten op de interventie-scholen ($n=82$; 60 klassen in de interventiegroep die betrokken zijn in de vergelijking met de controlegroep, aangevuld met 12 klassen uit 6 interventiescholen die niet gekoppeld waren) hebben bij de voormeting vrij positieve verwachtingen ten aanzien van de effectiviteit van KlankKr8 bovenop het reguliere leescurriculum (gem: 3,3 op een vijfpuntsschaal). Bij de nameting zijn ze in vergelijkbare mate positief: ze vinden het een verrijking van het regulier beginnend leesonderwijsprogramma (gem: 3,4 op een vijfpuntsschaal) en geven aan dat iets meer dan de helft van de leerlingen vooruit gaat in hun letterkennis als ze KlankKr8 gebruiken (gem: 3,2 op een vijfpuntsschaal). Verder zegt 40% van de leerkrachten dat ze KlankKr8 waarschijnlijk tot zeker wél zouden aanraden aan collega-leerkrachten in kleutergroepen. Diverse leerkrachten zien de toegevoegde waarde van KlankKr8 vooral voor de zwakkere leerlingen met extra ondersteuningsbehoefte bij het voorbereidend lezen.

KlankKr8 bestaat uit een aantal digibordlessen die worden gevolgd door individuele oefensessies op de devices (tablet, iPad of computer). De leerkrachten beschouwen de digibordlessen als nuttig (gem: 4,0 op een vijfpuntsschaal). Ze geven desgevraagd een aantal verbeterpunten voor de digibordlessen aan. Er worden veel letters in korte tijd aangeboden; vooral voor jongere leerlingen in groep 2 en NT2 leerlingen gaat dit te snel. Verder komen er moeilijk te onderscheiden dubbele klanken aan bod die normaliter pas in groep 3 aan de orde

zijn. De leerkrachten gaven daarnaast aan dat er weinig inoefenopdrachten waren, hetgeen de actieve deelname van alle kinderen bemoeilijkte. Bovendien misten ze de gebaren voor klanken, iets wat zij uit andere methodieken kennen als hulpmiddel bij het onthouden van klanken.

Volgens de leerkrachten vinden de leerlingen in groep 2 de digibordlessen heel erg leuk (gem: 4,8 op een vijfpuntsschaal). Het zelfstandig werken met KlankKr8 op de devices vinden ze over het algemeen iets minder leuk (gem: 4,1 op een vijfpuntsschaal). De leerkrachten lichten dit uitvoerig toe. Het thema van KlankKr8 (ruimte, raketten en de Loenie-figuurtjes) sluit goed aan bij de beleavingswereld van de leerlingen. De visuele elementen, zoals de grappige figuurtjes, geluiden en korte filmpjes, spraken de kinderen erg aan. Kinderen keken ernaar als een YouTube-filmpje. Verder bevordert de speelse interactieve opzet van de digibordlessen met klankposters, inkleur- en opzoekopdrachten de betrokkenheid en enthousiasme van de leerlingen volgens leerkrachten. Ook gaven zij aan dat het toewerken naar een concreet einddoel, zoals het lanceren van een raket, goed werkte om deze betrokkenheid bij de digibordlessen vast te houden. Na enkele weken merkten leerkrachten echter dat het vaste patroon van de lessen en het gebrek aan variatie leidden tot verminderde uitdaging en interesse bij de leerlingen: ze wisten al wat er ging gebeuren. Ook vonden de leerkrachten dat het lesaanbod weinig ruimte voor differentiatie bood, waardoor niet alle leerlingen voldoende werden uitgedaagd of betrokken waren.

Praktische uitvoerbaarheid van KlankKr8

De leerkrachten in groep 2 geven aan dat ze in totaal gemiddeld een klein uur besteden (gem: 52 min) aan het voorbereiden van de digibordlessen. Ze vinden de digibordlessen heel gebruikersvriendelijk (gem: 4,6 op een vijfpuntsschaal). Het installeren van KlankKr8 op de devices (gem: 3,5 op een vijfpuntsschaal) en het gebruik van de leerlingaccounts (gem: 3,6 op een vijfpuntsschaal) is volgens leerkrachten in het algemeen goed te doen. Er zijn geen problemen met de internetverbinding en het aantal werkende devices op school is meestal goed op orde. Vanuit KlankKr8 wordt aanbevolen dat de groep 2 leerlingen minstens 3 keer per week gedurende minimaal 10 minuten werken met KlankKr8 op de devices. De leerkrachten geven aan dat het meestal wel lukt om de leerlingen vaak genoeg te laten werken met KlankKr8 (gem 1,8 op een driepuntsschaal; met 1=nee, 2=meestal, 3=ja). De leerlingen spelen met KlankKr8 voornamelijk in de klas (55%). In de overige gevallen gebeurt dat op de gang (13%), zowel in de klas en de gang (18%) of in een aparte ruimte in combinatie met klas en gang (16%).

Er kan ook thuis gespeeld worden met KlankKr8: 65% van de leerkrachten heeft informatie hierover gegeven aan de ouders. Van de leerkrachten die deze informatie gedeeld hebben, heeft 68% geen zicht op of ouders enthousiast zijn over KlankKr8 of niet, 19% geeft aan dat ouders dat zeker wel zijn. Uit de backenddata van KlankKr8 is op te maken dat 15% van sessies thuis is gespeeld.

Voor het monitoren van KlankKr8 is een leerkrachtportaal beschikbaar, genaamd LeerKr8. Bijna alle leerkrachten geven aan hier weleens gebruik van te maken (soms: 71%; vaak: 22%). Zij gebruiken het portaal met name om te checken of een leerling voldoende heeft gespeeld en/of hoe de leervorderingen zijn (86%). Sommige leerkrachten zetten een kanttekening bij de betrouwbaarheid van het monitoren. Ze geven aan dat de motivatie van een aanzienlijke groep leerlingen om op devices te werken na enkele weken afneemt. Bij verminderde motivatie gaan leerlingen vaak gokken of zomaar klikken. Zo ontstaat niet altijd een goed beeld van de vordering van hun letterkennis.

De leerkrachten geven knelpunten bij de praktische uitvoerbaarheid van KlankKr8 aan. Ze ervaren dat het programma in totaal veel tijd vraagt—zowel in de voorbereiding van digibordlessen als bij de begeleiding van leerlingen op de devices. Ze noemen specifiek dat het steeds opnieuw moeten inloggen met QR-codes tijdrovend is. Verder vinden leerkrachten met de combinatiegroep 1/2 het lastig om KlankKr8 goed te organiseren voor de leerlingen in groep 2 met de leerlingen van groep 1 erbij. Ze hebben zorgen over het hergebruik van de lessen in het volgende schooljaar: de leerlingen uit groep 1 zullen deze dan voor de tweede keer doorlopen, wat vragen bij hen oproept over de effectiviteit en afwisseling op langere termijn.

4.2 In hoeverre bevordert KlankKr8 de letterkennis in groep 2?

Om de vraag te beantwoorden in hoeverre KlankKr8 de letterkennis in groep 2 bevordert is de leerwinst die de leerlingen op de interventiescholen in groep 2 hebben geboekt vergeleken met de leerwinst op de controlescholen. De verschilscore tussen voormeting en nameting 1 letterkennis (januari vs. juni) is gebruikt als afhankelijke variabele in de analyse om het effect van KlankKr8 te bepalen. De vergelijking is gebaseerd op 1216 leerlingen op 57 scholen. Het aantal leerlingen en scholen in beide groepen is nagenoeg gelijk. Zie Tabel 4.1 voor descriptieve statistieken m.b.t. de verschilcores.

Tabel 4.1 Descriptieve statistieken verschilcores letterkennis

	Totaal	Interventiescholen	Controlescholen
Aantal leerlingen	1216	609	607
Aantal scholen	57	28	29
Gemiddelde leerwinst letterkennis	2.42	2.63	2.20
Standaarddeviatie	3.28	3.21	3.34
Variantie op schoolniveau	.89	.79	.97
Variantie op leerlingniveau	9.84	9.49	10.19

Gemiddeld boeken de leerlingen 2.42 punt vooruitgang op de toets letterkennis. Vergeleken met de controlescholen is de gemiddelde verschilscore in de interventiescholen 0.43 punt hoger. Uit de standaarddeviaties valt af te lezen dat de spreiding iets hoger is in de controlescholen.

Verschillen tussen de interventie- en controlescholen

Ondanks de gerandomiseerde toewijzing blijft de kans aanwezig dat de groepen significant van elkaar verschillen op bepaalde kenmerken. Indien dat het geval is, is het nodig om in de analyses hiervoor te controleren. Wat betreft leerlingkenmerken is door middel van een aantal (logistische) multilevel-analyses bepaald of er sprake is van significante verschillen. Voor de leerkracht- en schoolkenmerken is dit gebeurd aan de hand van t-toetsen en χ^2 -toetsen.

De uitkomsten worden gerapporteerd in Tabel 4.2. Voor verreweg de meeste variabelen is er geen sprake van een significant verschil tussen de interventiegroep en de controlegroep. De meest in het oog springende uitzondering betreft de zorgen van de leerkracht over het gedrag en de ontwikkeling van hun leerlingen ($p = .005$). In de controlegroep tonen de leerkrachten zich bezorgder over het gedrag en de ontwikkeling van hun leerlingen. De score op deze schaal loopt van 1 tot 5. In de controlegroep ligt het gemiddelde op 2.16 tegen 1.67 in de interventiegroep (bij een standaarddeviatie van .97).

Daarnaast is er sprake van een significant verschil m.b.t. een tekort aan onderwijsassistenten in groep 1-2 ($p = .038$): In de interventiegroep geven vijf respondenten aan dat ze te kampen hebben met een tekort aan onderwijsassistenten in de groep 1 en 2, terwijl slechts één respondent in de controlegroep hiervan melding maakt. Tot slot is het percentage meisjes in de controlegroep (51.4%) hoger dan in de interventiegroep (46.1%) ($p = .098$). In de analyse naar het verschil tussen de interventiescholen en de controlescholen in de verschillen op letterkennis worden deze drie variabelen meegenomen als controlevariabelen. Voor alle andere achtergrondkenmerken ligt de p -waarde (ver) boven .10.

Tabel 4.2 Verschillen tussen interventie- en controlescholen (p -waardes) en descriptieve statistieken

Kwantitatieve leerlingkenmerken; p-waarden verschil bepaald d.m.v. multilevel regressieanalyse			
Variabele	p-waarde (2-zijdig)	Gemiddelde	SD
Voormeting/pre-test			
Pre-test letterkennis	.739	20.68	4.62
Foneemdiscriminatie	.273	13.64	1.29
Receptieve woordenschat	.868	11.50	2.88
Fonologisch bewustzijn	.824	11.33	3.16
Snelbenoemen kleuren	.238	.67	.18
Snelbenoemen afbeeldingen	.850	.69	.16
Zorgen over gedrag en ontwikkeling leerling bij leerkracht	.005	1.91	.97
Binaire leerlingkenmerken; p-waarden verschil bepaald d.m.v. logistische multilevel regressieanalyse			
Variabele	p-waarde (2-zijdig)	Proportie	
Ontbrekende score op een of meer pretesttoetsen	.852	.03	
Ontbrekende score vragenlijst leerkracht	.727	.01	

Versnelde schoolloopbaan (gezien leeftijd)	.410	.01	
Vertraagde schoolloopbaan (gezien leeftijd)	.898	.15	
Geslacht leerling (proportie meisjes)	.097	.49	
VVE-indicatie	.728	.14	
Thuis taal niet Nederlands	.700	.19	
Dyslexie in de familie	.659	.09	
Kwantitatieve leerkrachtkenmerken; p-waarden verschil bepaald d.m.v. t-toets			
Variabele	p-waarde (2-zijdig)	Gemiddelde	SD
Jaren werkervaring in primair onderwijs	.533	18.25	12.3
Jaren werkervaring met kleuters	.480	12.98	10.78
Tevredenheid met schoolcurriculum beginnende geletterdheid	.475	3.93	.73
Kwantitatieve schoolkenmerken; p-waarden verschil bepaald d.m.v. t-toets			
Variabele	p-waarde (2-zijdig)	Gemiddelde	SD
Vertrouwen schoolleider in leesonderwijs op school (max 5)	.150	4.17	.71
Aantal methoden technisch lezen op school	.635	2.17	.81
Aantal methoden spelling op school	.355	1.06	.32
Aantal leerlingen op school	.258	293.05	130.34
Aantal leerlingen in groep 2	.912	32.83	14.08
Schoolweging	.235	30.47	4.40
Stedelijkheid schoollocatie	.586	3.25	1.17
Binaire schoolkenmerken; p-waarden verschil bepaald d.m.v. χ^2-toets			
Variabele	p-waarde (2-zijdig)	Proportie	
Leerkrachtentekort groep 1-2	.574	.05	
Onderwijsassistententekort groep 1-2	.038	.08	
IB-tekort groep 1-2	.313	.02	
Leescoördinator op school	.723	.88	
Denominatie school: protestants-christelijk	.557	.26	
Denominatie school: rooms-katholiek	.631	.31	

Voor het bepalen van de effectiviteit van KlankKr8 is niet alleen een check op een succesvolle randomisatie bij start van de interventie van belang, ook gedurende de interventieperiode kan verdunning van het effect ontstaan als docenten op controlescholen, getriggerd door hun deelname aan het KlankKr8-onderzoek, toch in hun lessen extra aandacht besteden aan beginnend geletterdheid gedurende de interventieperiode. We hebben hiervoor geen aanwijzingen gevonden voor tijdens onze contactmomenten met de controlescholen, ook niet toen deze scholen na een jaar alsnog met KlankKr8 mochten starten (*delayed treatment design*) en daarover een implementatievragenlijst hebben ingevuld.

Verskil tussen interventiescholen vs. controlescholen op leerwinst letterkennis

Met een drietal multilevel-analyses is het verschil tussen de interventiescholen en de controlescholen bepaald. Er zijn 3 modellen gefit en getoetst:

Model 1 met alleen een binaire variabele die aangeeft of een school wel/niet tot de interventiegroep behoort als verklarende variabele

Model 2 met drie controlevariabelen. Dit zijn de variabelen waarop de interventiescholen en de controlescholen significant van elkaar verschillen (zorgen om leerling, onderwijsassistententekort, proportie meisjes)

Model 3 met alleen de controlevariabelen die een significant verband laten zien.

Missing values zijn geen groot probleem in deze analyses. Voor ruim twee derde van de variabelen in Tabel 4.3 zijn alle waarden bekend. Voor geen enkele variabele ligt het percentage ontbrekende waarden boven 1.4% (17 leerlingen). Zie Bijlage 4 voor nadere details. De resultaten van de multilevel-analyses worden gerapporteerd in Tabel 4.3.

Tabel 4.3 Verschil leerwinst letterkennis tussen interventie- en controlescholen

	MODEL 1		MODEL 2		MODEL 3	
	Coeff.	Sign.	Coeff.	Sign.	Coeff.	Sign.
Fixed Effects						
Intercept	2.267	<.001	1.102	.002	1.051	.001
Experimenteel – Controle*	.373	.123	.672	.027	.624	.030
Vrouwelijke leerling	---	---	-.067	.716	---	---
Zorgen over de leerling	---	---	.572	<.001	.576	<.001
Tekort onderwijsassistenten	---	---	-.363	.532	---	---
Random Effects						
Variantie op leerlingniveau	9.842	<.001	9.603	<.001	9.599	<.001
Variantie op schoolniveau	.878	.001	.954	<.001	.930	<.001

Afhankelijke variabele: verschilscore letterkennis (januari – juni); Tweezijdige significantieniveaus

* Eenzijdig significantieniveau voor verschil interventiescholen vs. Controlescholen

Wanneer geen rekening wordt gehouden met de variabelen waarop de interventiescholen en controlescholen significant verschillen, is het verschil op letterkennis tussen beide groepen scholen niet significant. Dit is echter wel het geval als in de analyse wordt gecontroleerd voor deze variabelen. Van de drie variabelen in kwestie (geslacht van de leerling, zorgen over de leerling en het tekort aan onderwijsassistenten op school) blijkt echter alleen de mate waarin een leerkracht zich zorgen maakt over het gedrag en de ontwikkeling van de leerlingen significant samen te hangen met de verschilcores op letterkennis. In Tabel 4.3 worden tweezijdige significantieniveaus gerapporteerd, m.u.v. het verschil tussen interventiescholen en controlescholen. Omdat vooraf werd verwacht dat de interventiescholen betere resultaten zouden boeken op letterkennis, is voor het effect van de experimentele conditie een eenzijdige toets gerechtvaardigd. De uitkomsten geven aan dat wanneer gecontroleerd wordt voor de zorgen die leerkrachten over hun leerling hebben de vooruitgang in letterkennis .624 hoger ligt in de interventiescholen. Gezien de standaarddeviatie in leerwinst op de controlescholen (zie Tabel 4.1), komt dit overeen met een effect size (Cohen's d) van .19. Om de robuustheid te testen van de analyses met verschilcores als uitkomstvariabele zijn deze analyses

eveneens uitgevoerd met score op de post-test als afhankelijke variabele en de pre-test als controle (zie bijlage 6). Het effect van de interventie is dan eveneens significant, maar wel kleiner dan het effect dat gerapporteerd is in tabel 4.3. Het effect van de interventie wordt dan geschat op .450. Dit komt overeen met een effect size van .13.

4.3 Welke factoren zijn van belang in de leeropbrengsten van KlankKr8?

Binnen de interventiegroep was de verwachting dat de hoeveelheid oefen/speeltijd van invloed zou zijn op de uitkomsten. Kleuters die meer oefenden met KlankKr8 werden verwacht meer vooruitgang te laten zien op letterkennis dan kleuters die minder hadden geoefend. In Tabel 4.4 valt af te lezen dat er sprake is van een aanzienlijke spreiding tussen de leerlingen in het aantal gespeelde sessies. Het gemiddeld aantal sessies is slechts 16, terwijl op basis van het voorschrift van de makers van KlankKr8 in deze periode tenminste 48 sessies plaats hadden moeten vinden. Een sessie duurde gemiddeld 13 minuten en 35 seconden. Slechts 16 leerlingen (2%) hebben meer dan 48 sessies gedaan en drie leerlingen meer dan 64 (0.4%). Het is in dat licht opvallend dat uit de vragenlijstgegevens bleek dat het leerkrachten meestal lukte om leerlingen volgens plan te laten werken aan KlankKr8 (zie paragraaf 4.1).

De samenhang tussen het aantal gespeelde KlankKr8-sessies en de geboekte leerwinst is nader onderzocht. De data-analyse betreft uitsluitend leerlingen op scholen die gewerkt hebben met KlankKr8. Naast de 28 interventiescholen die betrokken zijn in de vergelijking tussen de interventiegroep en de controlegroep heeft de data-analyse betrekking op zes interventiescholen die niet opgenomen konden worden in de eerdere vergelijking. Dit zijn de scholen die bij de randomisatie gekoppeld zijn aan controlescholen die zich na de randomisatie alsnog uit het onderzoek terugtrokken. Zodoende heeft de analyse betrekking op 803 leerlingen uit 34 scholen. In deze 34 scholen zaten 835 leerlingen in groep 2 in leerjaar 2024-25. Hiervan hadden 803 zowel een score op pre-test als op de posttest letterkennis. Voor 27 leerlingen is geen score op de post-test beschikbaar. Daarnaast ontbreekt bij vijf leerlingen met een geldige posttestscore een score op de pre-test.

Bij de leerlingen op de interventiescholen is via backenddata van de KlankKr8-app geregistreerd hoeveel sessies ze hebben gespeeld in de periode van 30 januari tot 19 juli, zowel op school als thuis. In de data-analyse wordt gecontroleerd voor een groot aantal leerling-, leerkracht- en schoolkenmerken. Dit betreft alle variabelen die zijn terug te vinden in Tabel 4.2 met een aantal aanvullende gegevens die alleen zijn verzameld op de scholen die met KlankKr8 hebben gewerkt. Deze variabelen zijn in Tabel 4 gearceerd.

Het gaat in de eerste plaats om drie variabelen die betrekking hebben op het aantal gespeelde KlankKr8-sessies. Voor 50 leerlingen zijn geen backenddata beschikbaar. Aangenomen is dat deze leerlingen zowel op school als thuis geen enkele sessie hebben gespeeld. In de data-analyse wordt een dummy-variabele meegenomen die het ontbreken van backenddata aangeeft. Hiermee wordt gecontroleerd of deze leerlingen wat betreft behaalde leerwinst afwijken van de overige leerlingen.

Een ruime meerderheid van de leerlingen (93.8%) heeft minstens één KlankKr8-sessie gespeeld op school. Slechts een kleine minderheid (32.3%) heeft ook thuis minstens één sessie gespeeld. Het gemiddelde aantal op school gespeelde sessies is ook veel hoger dan het aantal sessies dat leerlingen thuis hebben gespeeld. De leerlingen van wie backenddata beschikbaar zijn, hebben gemiddeld ruim 17 sessies op school gespeeld (range 0-87). Deze leerlingen hebben thuis gemiddeld minder dan drie sessies gespeeld.

Tabel 4.4 Beschrijvende statistieken over de interventiescholen

	Gemiddelde of proportie	Std. Dev.	N
Leerlingenkenmerken			
Leerwinst letterkennis	2.42	3.11	803
Aantal sessies KlankKr8 op school*	17.10	13.27	753
Aantal sessies KlankKr8 thuis	2.94	7.59	753
Proportie geen backenddata KlankKr8 beschikbaar	.06	.24	803
Foneemdiscriminatie (pretest)	13.60	1.30	800
Receptieve woordenschat (pretest)	11.59	.275	802
Fonologisch bewustzijn (pretest)	11.49	3.10	797
Snelbenoemen kleuren (pretest)	.68	.18	796
Snelbenoemen afbeeldingen (pretest)	.68	.16	793
Leerlingen met ontbrekende score pretesttoetsen	.02	.15	803
Zorgen leerkracht over gedrag en ontwikkeling leerling (pretest)	1.63	.33	790
Proportie versnelde schoolloopbaan (gezien leeftijd)	.01	.12	803
Proportie vertraagde schoolloopbaan (gezien leeftijd)	.14	.35	803
Geslacht leerling (proportie meisjes)	.48	.50	803
Proportie VVE-indicatie	.11	.31	803
Proportie thuistaal niet Nederlands	.14	.34	803
Proportie dyslexie in de familie	.11	.31	803
Leerkrachtkenmerken			
Jaren werkervaring in primair onderwijs	17.56	12,21	82
Jaren werkervaring met kleuters	12.04	10.74	82
Tevreden met schoolcurriculum beginnende geletterdheid (max 5)	4.04	.66	82
Inschatting effectiviteit KlankKr8 boven reguliere curriculum (max 5)	3.42	.81	82
Tevredenheid met KlankKr8 (alfa=.92)	3.38	.90	82
Aantal recente ontwikkelingen op school volgens leerkracht (max 5)	2.49	1.41	82
Proportie ontbrekende leerkrachtgegevens	.02	.13	82

Schoolkenmerken			
Vertrouwen schoolleider in leesonderwijs op school (max 5)	4.32	.66	57
57Aantal methoden technisch lezen op school	2.20	.84	57
Aantal methoden spelling op school	1.08	.27	57
Aantal leerlingen op school	397.26	289.43	57
Aantal leerlingen in groep 2	36.69	16.29	57
Schoolweging	29.48	4.04	57
Stedelijkheid schoollocatie	3.32	1.14	57
Proportie leerkrachtentekort groep 1-2	.04	.21	57
Proportie onderwijsassistententekort groep 1-2	.16	.36	57
Proportie IB-tekort groep 1-2	.03	.18	57
Proportie leescoördinator op school	.90	.30	57
Proportie Protestants-Christelijke school	.23	.42	57
Proportie Rooms-Katholieke school	.45	.49	57

Gearceerd: Kenmerken waarover uitsluitend informatie is verzameld op interventiescholen

* Hier wordt de gemiddeld aantal levels gerapporteerd voor de leerlingen van wie backend data beschikbaar zijn. Onder de aanname dat de leerlingen zonder backend data geen enkele sessie met de KlankKr8-app hebben gespeeld (aantal sessies is 0), bedraagt het gemiddelde voor *alle* leerlingen 16.03 sessies.

De overige aanvullende variabelen in Tabel 4.4 hebben betrekking op informatie die afkomstig is van de leerkrachten. Dit betreft een inschatting van de leerkracht over de effectiviteit van KlankKr8 en andere recente ontwikkelingen op hun school

Bij de analyses zijn ontbrekende waarden van leerlingkenmerken vervangen door de mediane waarde in hun groep. Alleen voor het aantal gespeelde sessies op school of thuis is aangenomen dat een ontbrekende waarde betekent dat de leerling geen enkele sessie heeft gespeeld. In het geval van leerkrachtkenmerken zijn de ontbrekende waarden vervangen door de mediaan in de totale dataset.

De samenhang tussen het aantal gespeelde sessies en de geboekte leerwinst is met de volgende aanpak in kaart gebracht. Ten eerste is een multilevel model gefit met daarin alle 36 verklarende variabelen uit Tabel 4.4. Gezien het hoge aantal verklarende variabelen ligt het risico van overfitting op de loer. Er is een reële kans dat toevallige afwijkingen gemodelleerd worden als een effect van de verklarende variabelen. De volgende stappen zijn doorlopen om uit te komen bij een model met minder verklarende variabelen: In de eerste stap zijn alle verklarende variabelen die een regressiecoëfficiënt opleveren met een p-waarde groter dan .50 uit het multilevel model verwijderd. Daarbij is een uitzondering gemaakt voor de drie variabelen die betrekking hebben op het gebruik van KlankKr8 (aantal sessies op school, aantal sessie thuis, geen backenddata beschikbaar). In de volgende stap zijn de variabelen verwijderd die een regressie- coëfficiënt opleveren met een p-waarde grote dan .20 (wederom m.u.v. van de variabelen over het KlankKr8-gebruik).

Hierna zijn de resterende verklarende variabelen één voor één verwijderd. Daarbij is begonnen met de variabele die de coëfficiënt met de hoogste p-waarde opleverde. Dit proces is herhaald totdat er alleen coëfficiënten resteren met een p-waarde kleiner dan .05 (tweezijdig). De drie variabelen die betrekking hebben op het KlankKr8-gebruik zijn bij deze stappen steeds buiten schot gebleven. Pas in de allerlaatste stap zijn ook deze variabelen verwijderd indien ze een coëfficiënt opleverden met p-waarde groter dan .05. Voor de variabelen over het aantal sessies op school en aantal sessies thuis is overigens een eenzijdige toetsing toegepast. Het onderzoek gaat immers uit van aanname dat KlankKr8 de leerwinst bevordert. De resultaten worden gerapporteerd in Tabel 4.5 en 4.6.

Tabel 4.5 laat weinig verschil zien tussen de diverse multilevel-modellen. De coëfficiënt van het aantal sessies gespeeld op school is in alle gevallen nagenoeg gelijk. Deze impliceert dat bij één sessie meer gespeeld op school, de leerwinst met ongeveer .02 punt hoger komt te liggen. Bij tien extra sessies ligt de leerwinst .20 punt hoger, bij twintig .40 enzovoorts.

Het aantal sessies dat leerlingen thuis spelen lijkt weinig toe te voegen. In dit verband liggen de coëfficiënten zeer dicht bij nul. Bovendien zijn ze tegen de verwachting in allemaal negatief. Omdat een eenzijdige toetsing is aangehouden, resulteert dit in significantieniveaus hoger dan .50. We kunnen dus met meer dan 50% zekerheid concluderen dat het aantal thuis gespeelde sessies geen positieve samenhang vertoont met de geboekte leerwinst.

Tabel 4.5 Samenhang KlankKr8-gebruik met leerwinst Letterkennis per multilevelmodel

	Aantal sessies op school		Aantal sessies thuis		Ontbrekende backenddata	
	Coeff.	Sign.	Coeff.	Sign.	Coeff.	Sign.*
Model met alle controlevariabelen	.020	.049	-.000	.504	.282	.587
Alleen controlevariabelen met $p < .50$	.020	.030	-.001	.518	.197	.692
Alleen controlevariabelen met $p < .20$	.021	.019	-.002	.564	.099	.839
Alleen controlevariabelen met $p < .05$	.021	.019	-.003	.583	-.195	.679
Alleen controlevariabelen met $p < .05$ en KlankKr8-variabelen met $p < .05$	.021	.013	---	---	---	---

Niet-gestandaardiseerde multilevel regressie-coëfficiënten; eenzijdige significantieniveaus

*Tweezijdige significantieniveaus voor ontbrekende backenddata

De samenhang tussen ontbrekende backenddata en leerwinst is eveneens in geen van de modellen significant. In dit geval zijn de resultaten van een tweezijdige significantietoets gerapporteerd. Deze dummy-variabele is opgenomen in de data-analyse om rekening te houden met de mogelijkheid dat leerlingen met ontbrekende backenddata afwijken van de overige leerlingen (positief of negatief).

Tabel 4.6 Multilevel analyse leerwinst Letterkennis

	Coëfficiënt	Significantie
Fixed Effects		
Intercept	3.508	<.001
Aantal sessies op school*	.021	.013
Toets fonologisch bewustzijn	-.198	<.001
Zorgen leerkracht over gedrag en ontwikkeling leerling	.462	.005
VVE-indicatie	1.026	.006
Random Effects		
Variantie op leerlingniveau	8.222	<.001
Variantie op schoolniveau	.312	.082

Niet-gestandaardiseerde multilevel regressie-coëfficiënten; tweezijdige significantieniveaus*Eenzijdig significantieniveau voor aantal sessies op school

In Tabel 4.6 staan de resultaten voor het uiteindelijk multilevel-model. Drie controlevariabelen leveren een coëfficiënt op die op een significante samenhang met leerwinst duiden. Het gaat om een negatieve coëfficiënt van de score op de toets fonologisch bewustzijn. Hoe hoger de score op deze toets, des te geringer de leerwinst op letterkennis. De twee andere coëfficiënten zijn positief. Het blijkt dat naarmate een leerkracht zich meer zorgen maakt over de cognitieve ontwikkeling en het gedrag van een leerling, deze meer leerwinst hebben geboekt op het gebied van letterkennis. Ook blijken leerlingen met een VVE-indicatie meer leerwinst te hebben geboekt dan de overige leerlingen. De multilevel-analyse laat tenslotte ook zien dat het merendeel van de onverklaarde variantie leerlingsspecifiek is. De hoeveelheid schoolvariantie is na controle voor de vier verklarende variabelen niet significant voor $p < .05$. Naast de analyses met verschilcores als uitkomstvariabele zijn deze analyses eveneens uitgevoerd met score op de post-test als afhankelijke variabele en de pre-test als controle (zie bijlage 6). In beide gevallen is het geschatte effect van het aantal op school gespeelde sessies nagenoeg gelijk (.021 vs. .022).

Samenhang implementatie met aantal gespeelde sessies op school

Uit bovenstaande multilevel-analyse volgt dat het aantal gespeelde sessies op school een belangrijke voorspeller is voor letterkennis leerwinst van leerlingen. Vanuit KlankKr8 is het advies gegeven om de leerlingen minstens 3 keer per week gedurende minimaal 10 minuten te laten werken met KlankKr8 op de devices. We hebben nader gekeken naar hoe KlankKr8 op de interventiescholen werd geïmplementeerd. Daaruit bleek dat er veel variatie is tussen leerlingen in het totaal aantal gespeelde sessies op school (zie Tabel 4.4). Dat geldt ook voor het klasgemiddelde van het aantal gespeelde sessies op school (gem = 15,3, $SD = 11,4$, min = 0, max = 46).

Ter verdieping zijn we nagegaan in hoeverre dit klasgemiddelde geassocieerd is met verschillen tussen leerkrachtpercepties over KlankKr8. Informatie hierover is afkomstig uit de leerkrachten vragenlijst van nameting 1. Zo bleek dat het klasgemiddelde van het aantal gespeelde sessies op school significant (positief) gecorreleerd is met de volgende kenmerken:

- De inschatting van de effectiviteit van KlankKr8 boven op het reguliere leescurriculum op de school (Pearson's $r=0.28$)
- De tevredenheid met het programma KlankKr8 (Pearson's $r=0.25$)
- De totale voorbereidingstijd voor de digibordlessen van KlankKr8 (Pearson's $r=0.29$)
- Het gemak waarmee KlankKr8 op de devices te installeren is (Pearson's $r=0.31$).
- Het gemak waarmee leerlingaccounts van KlankKr8 te gebruiken zijn (Pearson's $r=0.36$).
- De mate waarin het lukt om leerlingen vaak genoeg te laten werken met KlankKr8 (Pearson's $r=0.39$).
- De plek op school waar leerlingen werken met KlankKr8 op de devices. In klassen waar leerlingen met de device spelen in de gang of klas in combinatie met een aparte ruimte worden gemiddeld genomen 12,6 meer sessies gespeeld dan in klassen waar de leerlingen alleen in de klas op de devices spelen (Met Bonferonni correctie voor multiple comparison: $p=0.005$)
- De mate waarin LeerKr8 is gebruikt (Pearson's $r=0.39$).
- De mate waarin letterkennis volgens de leerkracht verbetert als leerlingen KlankKr8 gebruiken (Pearson's $r=0.34$).

Met een multiple lineaire regressieanalyse zijn we vervolgens nagegaan welke implementatiekenmerken de sterkste samenhang hebben met het klasgemiddelde van het aantal gespeelde sessies op school. De toegepaste selectieprocedure (stepwise, backward, forward) was hierbij niet van belang: het resulteerde steeds in hetzelfde spaarzame model. Tabel 4.7 toont de resultaten van het spaarzame model (met alleen significante predictoren)

Tabel 4.7 Lineaire regressieanalyse van implementatiekenmerken op het klasgemiddelde aantal gespeelde sessies op school

	Ongestandaardiseerde coëfficiënt (B)	Significantie
Intercept	-14,630	0,003
De totale voorbereidingstijd voor de digibordlessen van KlankKr8 (in minuten)	0,039	0,017
Het gemak waarmee KlankKr8 op de devices te installeren is.	2,366	0,005
De mate waarin het lukt om leerlingen vaak genoeg te laten werken met KlankKr8	5,324	0,002
De mate waarin letterkennis verbetert als leerlingen KlankKr8 gebruiken	2,429	0,025
Werkplek gang	5,017	0,105
Werkplek gang en klas	2,827	0,323
Werkplek gang/klas in combinatie met apart ruimte	9,564	0,005
Verklaarde variantie (adjusted R²)	0,398	

Noot: (alleen) de klas als werkplek is referentie-categorie.

Met de gestandaardiseerde regressiecoëfficiënten (β) is de sterkte van de samenhang van de implementatiekenmerken met het klasgemiddelde van het aantal gespeelde sessies op school direct onderling te vergelijken. Het aantal gespeelde sessies hangt samen met hoe goed een leerkracht KlankKr8 heeft weten in te passen in het onderwijsprogramma. Tijd, speellocatie en digitale vaardigheid van de leerkracht spelen hierbij een belangrijke rol. Naarmate het vaker lukt om leerlingen genoeg te laten werken met KlankKr8 op de devices ($\beta = 0,30$), dit werken gebeurt in een apart ruimte in combinatie met de gang/klas ten opzichte van alleen in de klas ($\beta = 0,29$), en het installeren van KlankKr8 op de devices gemakkelijker verliep ($\beta = 0,27$), worden er gemiddeld genomen meer sessies door leerlingen op school gespeeld. Iets minder relevant maar wel significant zijn voorbereidingstijd voor de digibordlessen ($\beta = 0,23$) en perceptie van vooruitgang in letterkennis door het spelen met KlankKr8 ($\beta = 0,22$).

Extra letterkennistaak

De tweede letterkennis test die direct na de zomervakantie van 2024 is afgenomen is niet door alle kinderen gemaakt: van 126 leerlingen ontbreken de gegevens. Dit komt doordat twee interventiescholen de toets niet hebben afgenomen en doordat een aantal leerlingen in groep 2 zijn gebleven. Dit betreft 58 leerlingen (30 uit de experimentele groep en 28 uit de controlegroep). In totaal ontbreken de gegevens van 10.4% van de groep waarop de oorspronkelijke vergelijking tussen de controlegroep en interventiegroep is gebaseerd. De uitval is in de interventiegroep ruim tweemaal zo groot als in de controlegroep (14.0% vs. 6.8%). Uit een logistische multilevel-analyse blijkt dat het verschil in uitval statistisch gezien niet significant is ($p = .681$).

De uitkomsten laten zien dat deze letterkennistoets aanmerkelijk moeilijker was voor de leerlingen dan de Cito Kleuter in Beeld toets die was afgenomen voor de zomervakantie van 2024 in groep 2. De gemiddelde score bedraagt slechts 5.98 op een schaal van 0 tot 11. Tabel 4.8 geeft een overzicht van de aantallen leerlingen op de interventiescholen en de controlescholen en nadere beschrijvende statistieken m.b.t. de scores op de test letterkennis in groep 3. De leerlingen in de interventiescholen scoren ook in groep 3 iets beter op letterkennis.

Wat betreft achtergrondkenmerken is er nauwelijks sprake van significante verschillen tussen de interventiegroep en de controlegroep. In dit opzicht is het beeld vrijwel gelijk aan de verschillen die in Tabel 4.7 zijn gerapporteerd. Het meest opvallende verschil is nog steeds dat op de controlescholen de leerkrachten (uit groep 2) meer zorgen hadden over hun leerlingen ($p = .003$).

Tabel 4.8 Beschrijvende statistieken letterkennis groep 3

	Totaal	Interventiescholen	Controlescholen
Aantal leerlingen	1090	524	566
Aantal scholen	55	26	29
Gemiddelde (max 11)	5.98	6.12	5.86
Standaarddeviatie	2.03	1.99	2.05
Variantie op schoolniveau	.56	.26	.76
Variantie op leerlingniveau	3.51	3.73	3.32

Met een drietal multilevel-analyses is het verschil tussen de interventiescholen en de controlescholen bepaald. Hierbij is dezelfde procedure gevolgd als bij de leerwinst letterkennis in groep 2. Ondanks dat de leerlingen op de interventiescholen ook in groep 3 iets betere resultaten laten zien op letterkennis, is het verschil met de controlescholen niet significant (zie Tabel 4.9). Gezien de standaarddeviatie in de controlescholen (zie Tabel 4.9) komt dit verschil overeen met een effect size van .07. Als het effect significant was geweest, zou het nog steeds een effect van bescheiden omvang zijn.

Bij de analyses m.b.t. leerwinst in groep 2 was er geen sprake van een significant verschil tussen jongens en meisjes. Op de letterkennisuitkomsten na de zomer scoren meisjes significant slechter. De coëfficiënt voor zorgen over een leerling is significant voor zowel leerwinst in groep 2 en letterkennis in groep 3. Voor leerwinst letterkennis is deze positief, maar negatief voor het niveau in groep 3.

Tabel 4.9 Verschil letterkennis in groep 3 tussen interventiescholen en controlescholen

	MODEL 1		MODEL 2		MODEL 3	
	Coeff.	Sign.	Coeff.	Sign.	Coeff.	Sign.
Fixed Effects						
Intercept	5.786	<.001	6.692	<.001	6.700	<.001
Experimenteel – Controle*	.333	.083	.104	.339	.143	.280
Vrouwelijke leerling	---	---	-.241	.039	-.240	.040
Zorgen over de leerling	---	---	-.376	<.001	-.377	<.001
Tekort onderwijsassistenten	---	---	.327	.475	---	---
Random Effects						
Variatie op leerlingniveau	3.510	<.001	3.411	<.001	3.411	<.001
Variatie op schoolniveau	.546	<.001	.588	<.001	.583	<.001

Afhankelijke variabele: toets letterkennis september 2024

Tweezijdige significantieniveaus

* Eenzijdig significantieniveau voor verschil interventiescholen vs. controlescholen

4.4 In hoeverre heeft KlankKr8 een effect op technisch lezen, spelling en rekenen in groep 3?

De lees- en spellinguitkomsten van de leerlingen zijn tweemaal uitgevraagd in groep 3 (schooljaar 2024-2025), in januari/februari en in mei/juni. Als controlemaat zijn ook de gegevens van rekenen-wiskunde uitgevraagd op deze momenten. Niet alle scholen hebben op beide momenten alle drie de toetsen (DMT, spelling en rekenen) afgenomen: slechts voor 15 interventiescholen en 14 controlescholen is deze data volledig beschikbaar. Deze variatie heeft te maken met de gehanteerde leerlingvolgsystemen op school (Cito of anders) en vooral met de keuzes van aantal momenten van toetsing op school. Op slechts drie scholen is geen enkele toets technisch lezen of spelling-taalverzorging afgenomen. Dit betreft twee scholen uit de interventiegroep en één uit de controlegroep. Voor 214 (17.6%) van alle 1216 leerlingen die in

groep 2 zijn gevolgd is geen enkele LVS-score in groep 3 bekend. Voor slechts 482 leerlingen (35.2%) zijn LVS-scores beschikbaar voor alle drie de vakken in beide perioden. Naast de uitval door het toetscurriculum op de scholen is er leerlinguitval: niet alle leerlingen op een school hebben de toetsen gemaakt, bijvoorbeeld doordat ze nog een jaar in groep 2 bleven of omdat ze zijn verhuisd.

Al met al is er sprake van aanzienlijke en waarschijnlijk selectieve uitval. De resultaten van een vergelijking tussen de interventiegroep en de controlegroep moeten daarom uiterst omzichtig geïnterpreteerd worden. Tabel 4.10 laat de beschrijvende statistieken zien voor de LVS-toetsen.

Tabel 4.10 Descriptieve statistieken LVS-toetsen in groep 3

	Januari-Februari			Mei-Juni		
	Overall	Exp.	Contr.	Overall	Exp.	Contr.
Technisch lezen (Drie-Minuten-Toets)						
Gemiddelde	15.56	15.07	16.03	29.79	29.33	30.14
Standaarddeviatie	12.04	11.28	12.72	17.36	17.22	17.47
Aantal leerlingen	792	386	406	726	318	408
Aantal scholen	48	23	25	49	21	28
Spelling-Taalverzorging						
Gemiddelde	163.44	163.46	163.43	220.69	218.56	222.90
Standaarddeviatie	58.89	59.75	57.97	48.05	48.32	47.74
Aantal leerlingen	840	449	391	847	431	416
Aantal scholen	40	22	18	42	22	20
Rekenen-Wiskunde						
Gemiddelde	121.41	123.32	119.25	146.49	148.43	144.29
Standaarddeviatie	30.34	29.87	30.77	30.86	32.61	28.65
Aantal leerlingen	816	433	383	781	414	367
Aantal scholen	40	22	18	39	21	18

Met een aantal multilevel-analyses is nagegaan of er sprake is van statistisch significante verschillen tussen de interventiescholen en de controlescholen. De resultaten zijn terug te vinden in Tabel 4.11 en Tabel 4.12. Tabel 4.11 laat de resultaten zien van analyses waarin alleen het verschil tussen de interventiegroep en de controlegroep wordt geschat. De cijfers in Tabel 4.12 hebben betrekking op analyses waarin wordt gecontroleerd voor de drie variabelen waarvan eerder is vastgesteld dat de beide groepen hierop van elkaar verschillen. De belangrijkste conclusie is dat in groep 3 geen significante verschillen zijn te vinden tussen de interventiescholen en de controlescholen. Dit geldt zowel voor de analyses met en zonder controlevariabelen. De analyses laten verder zien dat meisjes in groep 3 significant slechter scoren dan jongens op technisch lezen en rekenen-wiskunde, maar niet op spelling-taalverzorging. Leerlingen over wie de leerkrachten zich in groep zorgen maakten (over hun ontwikkeling en/of gedrag) doen het in groep 3 significant slechter op vijf van de zes LVS-toetsen. Alleen voor technisch lezen in mei-juni is het verschil niet significant.

Tabel 4.11 Multilevel analyses LVS-toetsen in groep 3

		Januari-Februari		Mei-Juni	
		Technisch lezen (Drie-Minuten-Toets)			
Fixed Effects					
Intercept		15.329	<.001	28.216	<.001
Experimenteel – Controle		.012	.993	1.526	.580
Random Effects					
Variantie op leerlingniveau		129.348	<.001	248.770	<.001
Variantie op schoolniveau		16.745	.002	63.067	<.002
		Spelling-Taalverzorging			
Fixed Effects					
Intercept		159.092	<.001	220.139	<.001
Experimenteel – Controle		5.510	.605	-3.019	.618
Random Effects					
Variantie op leerlingniveau		2858.105	<.001	2053.269	<.001
Variantie op schoolniveau		651.763	<.001	253.576	.002
		Rekenen-Wiskunde			
Fixed Effects					
Intercept		116.886	<.001	143.433	<.001
Experimenteel – Controle		5.461	.273	4.052	.269
Random Effects					
Variantie op leerlingniveau		751.888	<.001	871.563	<.001
Variantie op schoolniveau		191.740	<.001	74.290	.006

Tweezijdige significantieniveaus

Tabel 4.12 Multilevel analyses LVS-toetsen in groep 3 met controlevariabelen (Model 2)

	Januari-Februari		Mei-Juni	
	Technisch lezen (Drie-Minuten-Toets)			
Fixed Effects				
Intercept	19.867	<.001	34.088	<.001
Experimenteel – Controle	-1.351	.377	.964	.724
Vrouwelijke leerling	-2.971	<.001	-5.832	<.001
Zorgen over de leerling	-1.530	.002	-1.337	.059
Tekort onderwijsassistenten	3.147	.184	-2.272	.666
Random Effects				
Variantie op leerlingniveau	127.068	<.001	242.200	<.001
Variantie op schoolniveau	14.665	.004	58.027	.002
	Spelling-Taalverzorging			
Fixed Effects				
Intercept	189.664	<.001	241.187	<.001
Experimenteel – Controle	-3.947	.710	-9.215	.197
Vrouwelijke leerling	1.866	.620	1.182	.283

Zorgen over de leerling	-14.747	<.001	-10.473	<.001
Tekort onderwijsassistenten	9.030	.579	2.757	.803
Random Effects				
Variantie op leerlingniveau	2706.452	<.001	1961.491	<.001
Variantie op schoolniveau	801.409	<.001	325.338	<.001
Rekenen-Wiskunde				
Fixed Effects				
Intercept	138.654	<.001	164.926	<.001
Experimenteel – Controle	.061	.990	-.135	.971
Vrouwelijke leerling	-9.595	<.001	-14.997	<.001
Zorgen over de leerling	-8.454	<.001	-6.757	<.001
Tekort onderwijsassistenten	12.077	.166	1.689	.789
Random Effects				
Variantie op leerlingniveau	699.450	<.001	806.512	<.001
Variantie op schoolniveau	182.454	<.001	67.331	.010

Tweezijdige significantieniveaus

Hoewel dit onderzoek geen bewijs levert voor een effect van deelname aan de Klankkr8-interventie in groep 2 op taal- en rekenvaardigheid in groep 3, is er wel sprake van een onmiskenbare samenhang tussen letterkennis in groep 2/begin groep 3 en de diverse LVS-scores in de tweede helft van groep 3. Tabel 4.13 rapporteert de correlaties tussen letterkennis op diverse meetmomenten en de LVS-scores op Technisch lezen, Spelling-Taalverzorging en Rekenen-wiskunde in januari/februari en mei/juni. Alle correlaties zijn zowel statistisch significant ($p < .01$; tweezijdig) als inhoudelijk betekenisvol (variërend van .312 tot .488).

Tabel 4.13 Correlaties (Pearson r) tussen letterkennis en LVS-scores in groep 3 (interventiescholen en controle)

	Letterkennis in groep 2 (jan.)		Letterkennis in groep 2 (juni)		Letterkennis in groep 3 (sept.)	
	r	n	r	n	r	n
Technisch lezen (jan.-feb. - M3)	.370	792	.334	792	.475	744
Spelling-Taalverzorging (jan.-feb. M3)	.358	840	.353	840	.402	774
Rekenen-Wiskunde (jan.-feb. M3)	.409	816	.389	816	.382	770
Technisch lezen (mei-juni E3)	.354	726	.348	726	.488	696
Spelling-Taalverzorging (mei-juni E3)	.370	847	.336	847	.380	781
Rekenen-Wiskunde (mei-juni E3)	.333	781	.312	781	.368	733

Alle correlaties zijn significant voor $p < .01$ (tweezijdig)

5 Samenvatting, discussie en conclusie

Gegeven dat onze samenleving in sterke mate geletterdheid vereist, is het van belang om leerlingen een optimale start te bieden in het leesonderwijs. Het bevorderen van een stevige basis voor lezen en spellen in de kleutergroepen verdient daarom aandacht. Letterkennis vormt een belangrijk onderdeel van die basis. KlankKr8 op school is een adaptief programma gericht op het aanleren van letter-klankkoppelingen. De effectiviteit van dit programma is eerder in kleinschaligere onderzoeken vastgesteld. Het doel van het huidige onderzoek was om in kaart te brengen of die effectiviteit ook kon worden aangetoond in een grootschalig onderzoek met groep 2 leerlingen. De reguliere schoolcontext werd zoveel mogelijk gevolgd en er werd niet gestuurd op gebruik van het programma. Deelnemende scholen werden aan elkaar gekoppeld op basis van achtergrondkenmerken en vervolgens werd één school van het koppel toegekend aan de interventiegroep en één school aan de controlegroep. De letterkennisopbrengsten werden gemeten eind groep 2 (na een half jaar KlankKr8 gebruikt te hebben) en geletterdheid werd gemeten in midden en eind groep 3.

Effect van KlankKr8 gemeten door groepsvergelijkingen op letterkennis

Onderzoeksvraag 1 betrof in hoeverre KlankKr8 de letterkennis bevordert in groep 2. Deze vraag is beantwoord door de letterkennisdata van de Cito LVS en door een extra letterkennistaak.

In zowel de controlegroep als de interventiegroep was er bij de leerlingen sprake van een toename in de passieve letterkennisuitkomst gemeten door de Cito kleuter in beeld letterkennistaak na afloop van de interventieperiode. Er was een toename van ruim 2 letters correct. De toename van de letterkennis van de leerlingen in de interventiegroep was significant hoger dan die van de leerlingen in de controlegroep. Het verschil bedroeg echter gemiddeld minder dan 1 letter in analyses waarbij gecontroleerd werd voor de variabelen waarop de controle- en interventiescholen van elkaar verschilden. Die variabelen betroffen de proportie meisjes, de mate van zorgen van de leerkracht over de ontwikkeling van de leerling en het tekort aan onderwijsassistenten op school. In de controlegroep lag het percentage meisjes in groep 2 hoger dan in de interventiegroep. De leerkrachten in de controlegroep maakten zich ook meer zorgen over de ontwikkeling van hun leerlingen. In de interventiegroep was het tekort aan onderwijsassistenten groter. De effect size (Cohen's d : .13 - .19) kan in de context van het onderwijs worden aangemerkt als een gemiddeld tot groot effect (Kraft, 2020).

Vanwege een plafondeffect op de voormeting, is er bij een extra nameting een extra passieve letterkennistaak afgenomen, waarin meer letters werden bevraagd dan tijdens de eerste nameting. Het verschil in scores van de leerlingen in de interventie- en de controlegroep op deze extra nameting was statistisch niet significant.

Daarom dient het te worden opgemerkt dat het significante effect van KlankKr8 is gevonden op een letterkennistoets die niet ideaal bleek: de uitkomsten op de letterkennistoets waren gemiddeld relatief hoog op de voormeting (wat we hadden vastgesteld nadat de nameting al was uitgezet) en bevatten niet alle mogelijke letters en lettercombinaties. In eerder

kleinschaliger en kortdurender interventie-onderzoek naar KlankKr8 bij kleuters (Verwimp et al., 2023) werd letterkennis gemeten aan de hand van 44 letters/lettercombinaties, terwijl onze letterkennistoets opgaven met 26 verschillende letters bevatte. Tegelijkertijd bleek de extra letterkennistoets van 20 items, inclusief moeilijke combinaties (zoals eu en ng) juist weer te moeilijk te zijn voor de leerlingen in de eerste weken van groep 3. Het meten van letterkennis vraagt dus verdere aandacht: in de toetsing is een balans in complexiteit nodig, om bodemscores en bijbehorende frustratie van kleuters te vermijden, maar ook om plafondscores te vermijden.

Effect van KlankKr8 gemeten door groepsvergelijkingen op lezen en spellen in groep 3

Onderzoeksvraag 3 betrof in hoeverre KlankKr8 een effect heeft op technisch lezen en spelling in groep 3. Een vergelijking op de lees- en spellinguitkomsten van de leerlingen in groep 3 kon alleen plaatsvinden voor kinderen die naar groep 3 doorstroonden (dus niet voor kleuters die groep 2 herhaalden), en voor leerlingen bij wie de Cito-leerlingvolgssysteemtoetsen werden afgenomen, namelijk de Drie-Minuten-Toets en de Cito spelling van Leerling in Beeld. Dit betekent dat er sprake is van een veel kleinere dataset gebaseerd op selectieve uitval, en dat de bevindingen met voorzichtigheid moeten worden geïnterpreteerd. Er waren geen verschillen tussen deze subset van de interventiegroep en de controlegroep op deze geletterdheidstoetsen. Letterkennis (voormeting, nameting, extra meting) was wel positief zwak-matig significant gecorreleerd aan lees- en spellinguitkomsten in groep 3. Een positieve correlatie werd verwacht op basis van conceptuele literatuur en empirische bevindingen (bijv Bonte & Brem, 2024; Van Viersen et al., 2018). Op de controlemaat van rekenen-wiskunde verschilden de twee groepen ook niet van elkaar.

Een mogelijke verklaring voor de bevindingen op lezen en spellen in groep 3 is dat de (significante) leeropbrengst van de interventiegroep t.o.v. de controlegroep beperkt was (namelijk gemiddeld minder dan 1 letter). Dit had te maken met de speeltijd (zie volgende paragraaf). Het relatief kleine verschil tussen de groepen betekent mogelijk dat dit te klein was om invloed te hebben nadat het geletterdheidsonderwijs start. Omdat de lees- en spellingsuitkomsten zijn afgenomen op M3, dus een half jaar na de letterkennistaak, is het mogelijk dat onderwijs in groep 3 bepalender is geworden voor de lees-spellinguitkomsten. Tegelijkertijd heeft eerder kleinschaliger onderzoek laten zien dat er ook geen verschil was op leesaccuratesse-uitkomsten van kleuters direct na een korte interventieperiode van 3 weken met KlankKr8 (Verwimp et al., 2023). Dit sluit aan bij onderzoek naar een andere letterkennisinterventie, GraphoGame: een review naar de effectiviteit van dat programma liet zien dat de interventie effectief was voor leesvoorwaarden, maar niet op de leesvaardigheid (McTigue et al., 2020). Nader onderzoek is dus nodig om de KlankKr8opbrengst in relatie tot het geletterdheidscurriculum te duiden.

KlankKr8-effecten in relatie tot contextkenmerken en aantal sessies

Onderzoeksvraag 2 betrof welke factoren de leeropbrengst van KlankKr8 beïnvloedden. Om de uitkomsten van de groepsvergelijkingen te duiden werd onderzocht of contextkenmerken een rol speelden in de KlankKr8-opbrengst. Er was een aantal contextkenmerken dat bijdroeg aan de behaalde letterkenniswinst in de interventiegroep. Zo had een hoger fonologisch bewustzijn een negatief effect op de letterkenniswinst. Daarentegen hadden de mate van

leerkrachtzorgen over de ontwikkeling en het gedrag van de leerling, en als leerling VVE heeft gevolgd een positief effect. Na controle voor deze contextkenmerken vonden we dat leerlingen die meer hadden geoefend ook meer vooruit gingen op letterkennis dan leerlingen die minder oefenden. Er was aanzienlijke variatie in het aantal oefensessies. Dat had te maken met het moment waarop scholen startten met KlankKr8 en met het aantal momenten dat kleuters konden werken met het programma. Het gemiddelde aantal sessies was met 16 beduidend lager dan het minimaal beoogde doel van 48 sessies dat werd aanbevolen door de makers van KlankKr8. In dat licht is het opvallend dat zelfs met deze gemiddeld beperkte hoeveelheid sessies er al een groepseffect werd gevonden in de groepsvergelijking. De bevinding dat het aantal oefensessies beperkter was dan aanbevolen sluit niet direct aan bij de rapportage van scholen dat het gemiddeld meestal lukte om de aanbevolen speeltijd te halen.

Aandacht voor implementatie van KlankKr8

Een verklaring voor de relatief beperkte effecten op letterkennis kan zijn dat de scholen in de interventiegroep al voldoende goede aandacht besteden aan letterkennis, waardoor de ruimte voor een effect van KlankKr8 beperkt is.

Een in ons ogen waarschijnlijker interpretatie is dat implementatie van KlankKr8 bevorderd kan worden. Eerder onderzoek naar KlankKr8 (Aravena et al., 2018; Verwimp et al., 2023) vond plaats in een context met meer sturing van onderzoekers en met regelmatigere en meer oefensessies. Het is dus waarschijnlijk dat de KlankKr8-opbrengsten in onze steekproef groter zouden zijn geweest als de kleuters meer sessies hadden gewerkt met KlankKr8. Onze bevindingen lieten immers zien dat het aantal aanbevolen oefensessies niet werd behaald en dat het aantal oefensessies invloed had op de letterkennistoename. De implementatie van KlankKr8 kan dus worden geoptimaliseerd.

In de uitvraag aan leerkrachten bleek dat zij gemiddeld genomen positief waren over het programma, zowel de klassikale startlessen als de individuele sessies, en dat ze positief waren over de ingeschatte effecten van het programma. Ze gaven echter ook praktische knelpunten rond het werken met de devices aan en inhoudelijke suggesties voor verbetering. Deze uitvraag biedt mogelijkheden voor KlankKr8 om het programma nog beter aan te laten sluiten bij de onderwijscontext, om zo de effectiviteit te kunnen optimaliseren (zie ook Harn et al., 2013). Een andere mogelijkheid is dat de effectiviteit van de interventie toeneemt als scholen meer ervaring hebben met de implementatie en inbedding van KlankKr8 op hun school. In dit eerste jaar moesten leerkrachten zich het hele programma eigen maken, inclusief de software, de klassikale lessen, het leerlingprogramma en het Leerkr8-portal. Meer ervaring en tijd dragen mogelijk bij aan een betere implementatie en effectiviteit. De interventie Bouw! voor leerlingen met (een risico op) vroege leesproblemen bleek bijvoorbeeld effectiever te worden naarmate het langer op school werd gebruikt (Van der Weijden et al., 2024).

Een optie die we hier niet hebben onderzocht is of de aangeboden inhoud invloed heeft op de effectiviteit. In het proefschrift van Van der Weijden (2024) naar de effectiviteit van Bouw! werd bijvoorbeeld gevonden dat het aantal en soort lessen dat een kind uitvoerde in Bouw! invloed had op de leesuitkomsten. Voor KlankKr8 zijn er voorzichtige aanwijzingen dat inhoud ook invloed kan hebben: Verwimp et al. (2023) vonden dat er aanzienlijke variatie was in de levels die kinderen selecteerden tijdens het oefenen (naar een moeilijker level of het huidige

level herhalen). De bevindingen sluiten aan bij de literatuur die aangeeft dat het opschalen van interventies en aandacht voor implementatie in interventies en interventie-onderzoek van groot belang zijn (bijv Harn et al., 2013; Nilsen & Bernhardsson, 2019; Ryan et al., 2025). Samengenomen geven de resultaten aan dat KlankKr8 een kansrijke interventie is voor het bevorderen van letterkennis, maar dat aandacht voor de implementatie van wezenlijk belang is.

6 Literatuurverwijzingen

Aravena, S., Tijms, J., Snellings, P., & van der Molen, M. W. (2017). Predicting Individual Differences in Reading and Spelling Skill With Artificial Script–Based Letter–Speech Sound Training. *Journal of Learning Disabilities*, 51(6), 552–564.

<https://doi.org/10.1177/0022219417715407>

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology*, 3(1).

<https://doi.org/10.1186/s40359-015-0089-9>

Bonte, M., & Brem, S. (2024). Unraveling individual differences in learning potential: A dynamic framework for the case of reading development. *Developmental Cognitive Neuroscience*, 66, 101362. <https://doi.org/10.1016/j.dcn.2024.101362>

van den Bos, K. P., & Lutje Spelberg, H. C. (2007). *Continu Benoemen en Woorden Lezen* [Continuous Naming and Word Reading]. Boom Test Uitgevers.

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, 79(3), 181–193. <https://doi.org/10.1177/001440291307900204>

Kraft, M.A. (2020). Interpreting Effect Sizes of Education Investigations. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *DMT en AVI. Groep 3 tot en met 8*. [DMT and AVI. Grade 1 through 6]. Cito.

Lane, K. L., & Menzies, H. M. (2005). Teacher-Identified Students with and without Academic and Behavioral Concerns: Characteristics and Responsiveness. *Behavioral Disorders*, 31(1), 65–83. <https://doi.org/10.1177/019874290503100103>

Piasta, S. B., Petscher, Y., & Justice, L. M. (2012). How many letters should preschoolers in public programs know? The diagnostic efficiency of various preschool letter-naming benchmarks for predicting first-grade literacy achievement. *Journal of Educational Psychology*, 104(4), 945–958. <https://doi.org/10.1037/a0027757>

Salmerón, L., Altamura, L., Delgado, P., Karagiorgi, A., & Vargas, C. (2024). Reading comprehension on handheld devices versus on paper: A narrative review and meta-analysis of the medium effect and its moderators. *Journal of Educational Psychology*, 116(2), 153–172. <https://doi.org/10.1037/edu0000830>

Sunde, K., Furnes, B., & Lundetræ, K. (2019). Does Introducing the Letters Faster Boost the Development of Children’s Letter Knowledge, Word Reading and Spelling in the First Year of School? *Scientific Studies of Reading*, 24(2), 1–18.

<https://doi.org/10.1080/10888438.2019.1615491>

van Viersen, S., de Bree, E. H., Zee, M., Maassen, B., van der Leij, A., & de Jong, P. F. (2018). Pathways into literacy: The role of early oral language abilities and family risk for dyslexia. *Psychological Science*, 29(3), 418-428. <https://doi.org/10.1177/0956797617736886>

Verwimp, C., Snellings, P., Wiers R.W., & Tijms J. (2023a). A randomised proof-of-concept trial on the effectiveness of a game-based training of phoneme-grapheme correspondences in pre-readers. *Journal of Computer Assisted Learning*, 39(5), 1607-1619. <https://doi.org/10.1111/jcal.12821>

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. p. 86 Harvard University Press. (vertaling: Verhofstadt-Denève L., Vyt, A., Van Geert, P. (1991). *Handboek ontwikkelingspsychologie. Grondslagen en theorieën*. p. 168. Bohn Stafleu Van Loghum).

Van der Weijden, F. A., Van den Boer, M., Zijlstra, B. J. H., & De Jong, P. F. (2024). Implementation Takes Time: Reduction of Literacy Problems in Schools Implementing an Early-Literacy Intervention. *Journal of Research on Educational Effectiveness*, 1–33. <https://doi.org/10.1080/19345747.2024.2384365>

de Wijs, A., Kamphuis, F., Kleintjes, F., & Tomesen, M. (2010). *Leerling- en onderwijsvolg-systeem. Spelling Groep 3 t/m 6..* Cito.

Bijlage 1 Items van de achtergrondkenmerkenvragenlijst over school, klas en leerlingen (voormeting)

Klasniveau: Digitale vragenlijst over kleuterklas en leerlingkenmerken [voor leerkracht]

Achtergrondinformatie

1. Hoeveel jaar werkervaring heb je als bevoegde leerkracht in het primair onderwijs
2. Hoeveel jaar werkervaring heb je in groep 1 en/of groep 2 in totaal?
3. Over welke (aan dit onderzoek deelnemende) kleuterklas vul je deze vragenlijst in?

Beginnende geletterdheid

Toelichting: Deze volgende vragen gaan over beginnende geletterdheid in uw kleuterklas.

4. Hoe tevreden ben je met het curriculum op jouw school voor beginnende geletterdheid/voorbereidend lezen in de groepen 1 en 2? (Slidervraag van heel ontevreden (links) naar zeer tevreden (rechts)
 - a. Kun je je antwoord toelichten?
5. [Voor interventiegroep] Wat is jouw inschatting van de effectiviteit van KlankKr8 boven op het reguliere leescurriculum op de school? (Slidervraag van heel weinig effect (links) naar heel groot effect (rechts) (1-5))
 - a. Kun je je antwoord toelichten?

Algemene achtergrondkenmerken

Toelichting: Dit blok van vragen gaat over de leerlingen uit jouw klas die bij het KlankKr8-onderzoek betrokken zijn.

6. Dit gaat om onderstaande leerlingen. De namen van de leerlingen zijn door de onderzoekers overgenomen van de toestemmingsformulieren van ouders. Mogelijk staan hier spelfouten in. Mochten er wijzigingen nodig zijn, kun je dat hieronder aangeven.
[alle namen van deelnemende kinderen]

Invullen naam kind	→	Kind 1	Kind 2	Kind 3	Kind 4 (etc)
Algemene achtergrond leerling					
7	Geslacht en dyslexie thuis				
7.1	Geslacht van de leerling	Meisje/jongen/anders			
7.2	Geboortemaand en -jaar leerling	Maand & jaar			
7.3	Heeft de leerling een VVE-indicatie gehad?	• Ja, • Nee • Onbekend			
7.4	Heeft de leerling voorheen de VVE bezocht?	• Ja • Nee • Onbekend			

8 Taal en geletterdheid thuis					
8.1	Welke taal spreekt het kind thuis het meest?	- Nederlands (of Fries of een Nederlands dialect) - Een andere, Europese, taal (bijvoorbeeld: Duits, Engels, Pools, Zweeds) - Een andere, niet-Europese, taal (bijvoorbeeld: Arabisch, Berber, Chinees, Papiaments, Turks) - Onbekend			
8.2	Komt dyslexie voor bij biologische ouders en/of brussen van het kind?	- Ja, - Nee - Onbekend			
9 Ouders/verzorgers					
9.1	Er volgt een aantal vragen over de ouders/verzorgers (vanaf nu: <i>ouders</i> genoemd). Wie neemt u hiervoor in gedachten?	Ouder/verzorger 1 - Biologische moeder - Biologische vader - Niet-biologische moeder - Niet-biologische vader - Anders, namelijk.... Ouder/verzorger 2 - Biologische moeder - Biologische vader - Niet-biologische moeder - Niet-biologische vader - Er is geen 2^e ouder/verzorger in beeld - Anders, namelijk....			
9.2	Hoe goed beheerst ouder1 het Nederlands?	5-puntsschaal van: de ouder spreekt nauwelijks Nederlands (1) naar de ouder spreekt vloeiend Nederlands (5)			
9.3	Hoe goed beheerst ouder2 het Nederlands? (indien 2 ^e ouder/verzorger in beeld)	5-puntsschaal van: de ouder spreekt nauwelijks Nederlands (1) naar de ouder spreekt vloeiend Nederlands (5)			

10	Kind in de klas				
10.1	In hoeverre heeft u bij deze leerling zorgen over... 5-puntsschaal van: helemaal niet (1) in grote mate (5)	• De mondelinge taalverwerving			
		• De beginnende geletterdheid			
		• De cognitieve leervaardigheid			
		• Het gedrag in het algemeen			
		• Internaliserend probleemgedrag (bijv. teruggetrokken, weinig zelfvertrouwen)			
		• Externaliserend probleemgedrag (bijv. ruzie maken, woede-aanvallen, niet luisteren)			
10.2	Krijgt de leerling extra ondersteuning voor beginnende geletterdheid?	• Ja, namelijk ... • Nee			

Schoolniveau: Digitale vragenlijst over schoolkenmerken [voor IB/taalcoördinator of directeur]

1. Is er in de groep 1 en 2 van uw school sprake van een...
 - a. Leerkrachttekort (ja/nee)
 - b. Onderwijsassistenttekort (ja/nee)
 - c. IB-tekort (ja/nee)
2. Is er een leescoördinator op school?
 - a. Ja
 - b. Nee
 - c. Weet ik niet
3. In hoeverre is jouw school in staat om het lees- en spellingonderwijs vorm te geven met ondersteuningsniveaus 1, 2 en 3 in de groepen 3 t/m 8? (Schaal van helemaal niet (links) naar in grote mate/helemaal wel (rechts))
4. Welke methode(n) wordt gebruikt voor technisch lezen? (Meerkeuze)
 - a. Atlantis
 - b. Blink
 - c. Estafette
 - d. Karakter
 - e. Leeslink
 - f. LetterKlankStad
 - g. Lijn3
 - h. Nieuw Nederlands Junior
 - i. Pit
 - j. Staal
 - k. Taal Actief

- l. Taal Oceaan
 - m. Veilig Leren lezen
 - n. Zo leer je kinderen lezen en spellen
 - o. Anders, namelijk...
 - p. Geen methode
5. Welke methode(n) wordt gebruikt voor spelling?
- a. Blink
 - b. Estafette
 - c. Pit
 - d. Staal
 - e. Taal Actief
 - f. Taaljacht
 - g. Taal Oceaan
 - h. Zo leer je kinderen lezen en spellen
 - i. Anders, namelijk...
 - j. Geen methode
6. Welke methode(n) of programma('s) wordt gebruikt voor beginnende geletterdheid/ voorbereidend lezen? (bijvoorbeeld de Lettertrein, Schatkist, Bouw!)

Bijlage 2 Items van de implementatievragenlijst voor interventiegroep (nameting 1)

Klasniveau: vragen over de implementatie van KlankKr8 in de klas

Algemeen over KlankKr8

1. Hoe tevreden ben je met het programma KlankKr8? (Heel ontevreden (1) – heel tevreden (5))
2. Wat is jouw inschatting van de effectiviteit van KlankKr8 boven op het reguliere leescurriculum op de school? (Heel weinig effect (1) – heel veel effect (5))

Gebruik

Digibordlessen

1. Hoeveel tijd (in minuten) heb je in totaal aan het voorbereiden van de digibordlessen van KlankKr8 besteed? (invullen in minuten)
2. Hoe *gebruiksvriendelijk* waren de digibordlessen van KlankKr8 volgens jou? (Helemaal niet, niet zo, neutraal, een beetje, heel)
3. Hoe *nuttig* vond je de digibordlessen van KlankKr8? (Helemaal niet, niet zo, neutraal, een beetje, heel)
4. Hoe *leuk* vonden de leerlingen de digibordlessen van KlankKr8? (Helemaal niet, niet zo, neutraal, een beetje, heel)
 - a. Kun je kort toelichten waarom de leerlingen de digibordlessen van KlankKr8 (helemaal niet, niet zo, neutraal, een beetje, heel leuk) vonden?

KlankKr8 en devices

5. Hoe makkelijk was het om KlankKr8 op de 'devices' (tablets, laptops, computers) te *installeren*? (Helemaal niet, niet zo, neutraal, een beetje, heel)
6. Hoe makkelijk was het om de leerlingaccounts van KlankKr8 te gebruiken? (Helemaal niet, niet zo, neutraal, een beetje, heel)
7. Zijn er voldoende werkende 'devices' in de klas om met KlankKr8 te werken? (ja / meestal / nee)
8. Is de internetverbinding op school stabiel genoeg om in de klas met KlankKr8 te werken? (ja / meestal / nee)

Zelfstandig gebruik door leerlingen

9. Lukt het om de leerlingen vaak genoeg te laten werken met KlankKr8? (ja / meestal / nee)
10. Waar werken de leerlingen op school met KlankKr8 als ze er zelfstandig mee aan de slag zijn? (meerdere opties mogelijk: in de klas, op de gang, in een aparte ruimte, anders, namelijk...)
11. Hoe leuk vinden leerlingen het om zelfstandig te werken met KlankKr8? (Helemaal niet, niet zo, neutraal, een beetje, heel)
12. Hebben ouders vanuit school gehoord dat kinderen ook thuis met KlankKr8 kunnen oefenen? (ja, nee, onbekend)
13. Zijn ouders enthousiast over het thuisgebruik van KlankKr8? (ja/nee/onbekend)

Gebruik van LeerKr8 (het leerkrachtportaal van KlankKr8)

14. Hoe vaak gebruik je Leerkr8? (vaak, soms, nooit)
15. Waarvoor gebruikt u Leerkr8? (Meerdere antwoorden mogelijk: Bekijken of leerlingen voldoende spelen, Vorderingen van leerlingen monitoren, Anders, namelijk...)

Evaluatie van KlankKr8

16. Zou je KlankKr8 aanraden aan collega-leerkrachten in de kleuterklas en groep 3? (zeker niet, waarschijnlijk niet, neutraal, waarschijnlijk wel, zeker wel)
17. Vind je KlankKr8 een verbetering van het reguliere beginnend leesonderwijsprogramma? (nee, niet zo, neutraal, enigszins, ja)
18. Vind je KlankKr8 een verrijking van het reguliere beginnend leesonderwijsprogramma? (nee, niet zo, neutraal, enigszins, ja)
19. Gaan leerlingen die KlankKr8 gebruiken naar uw idee vooruit op hun letterkennis? (ja, allemaal / het merendeel van de leerlingen / ongeveer de helft van de leerlingen / een beperkt deel van de leerlingen / geen van de leerlingen)

Schoolniveau: vragen over beginnend leesonderwijs op school

Zijn er, naast het gebruik van KlankKr8, nog andere recente ontwikkelingen op school rondom het gebied van beginnend leesonderwijs? Wat betreft...

- ...bemensing in jouw kleuterklas
- ...bemensing van interne begeleiding op school
- ...bemensing van leesspecialist op school
- ...basisonderwijsprogramma van beginnend leesonderwijs (dus naast KlankKr8)
- ...extra onderwijsprogramma van beginnend leesonderwijs voor leerlingen met een risico op leesproblemen (bv. Andere methode/materialen, indeling of werkwijze)
- ...toetsing van beginnend leesonderwijs
- ...IT-specialist op school
- ...hardware voor leerlinggebruik (tablets, laptops, computers)
- ...softwareprogramma's voor onderwijs (nieuwe apps/games voor lezen, rekenen, spellen)
- ...leerlingpopulatie in de kleuterklas (aantallen leerlingen, NT-leerlingen, leerlingen met speciale onderwijsbehoeften)

Bijlage 3 Items van de zelfgemaakte letterkennistoets (nameting 1)

Correcte antwoorden zijn gemarkeerd in oranje. Makkelijkere koppelingen zijn gemarkeerd in blauw. Moeilijke afleiders zijn gemarkeerd in paars (alleen 'au', 'ouw', 'oei' en 'ieuw' zijn niet uitgevraagd als doelvorm). De items die uiteindelijk zijn geëxcludeerd voor de analyse zijn in grijze letters weergegeven.

	Kind ziet en moet omcirkelen			Wat moet leerkracht zeggen:
1	l	i	ie	Waar zie je de i van vis?
2	oe	eo	ou	Waar zie je de oe van koek?
3	nk	nn	ng	Waar zie je de ng van lang?
4	gh	ch	cc	Waar zie je de ch van lach?
5	s	f	v	Waar zie je de f van fiets?
6	uu	iu	ui	Waar zie je de ui van huis?
7	ou	ouw	au	Waar zie je de ou van koud?
8	aaj	ei	aai	Waar zie je de aai van haai?
9	ie	i	ei	Waar zie je de ie van mier?
10	o	oo	ou	Waar zie je de oo van boot?
11	auw	uaw	ou	Waar zie je de auw van blauw?
12	uuw	uv	uw	Waar zie je de uw van duw?
13	nk	ngk	ng	Waar zie je de nk van bank?
14	g	j	gh	Waar zie je de g van gum?
15	euw	ieuw	eeuw	Waar zie je de eeuw van leeuw?
16	ooj	oi	ooi	Waar zie je de ooi van mooi?
17	ei	ie	ij	Waar zie je de ei van trein?

18	b	d	p	Waar zie je de d van doos?
19	eeu	eu	ue	Waar zie je de eu van leuk?
20	ei	eij	ij	Waar zie je de ij van kijk?

Bijlage 4 Missing values

Missing values zijn geen groot probleem in dit onderzoek. Voor slechts 9 van de 30 variabelen in onderstaande tabel ontbreken de scores bij een beperkt aantal leerlingen. Indien het leerlingkenmerken betreft, zijn de ontbrekende waarden vervangen door de mediane waarde in hun groep. In het geval van leerkrachtkenmerken zijn de ontbrekende waarden vervangen door de mediaan in de totale dataset. Het maximumaantal missing values bedraagt 17. Dit komt neer op 1.4% van alle leerlingen. Zie onderstaande tabel voor een overzicht.

Ontbrekende waarden op leerlingniveau

Leerlingkenmerken	Totaal	Contr.	Exp.
Toets auditieve discriminatie (ADIT)	8	5	3
Toets fonologisch bewustzijn	5	2	3
Toets receptieve woordenschat	6	1	5
Toets correct benoemen kleuren	17	10	7
Toets correct benoemen afbeeldingen	17	7	10
Zorgen over gedrag en ontwikkeling leerlingen door leerkracht	17	4	13
Ontbrekende score op een of meer pretests	---	---	---
Ontbrekende score vragenlijst leerkracht	---	---	---
Versnelde schoolloopbaan (gezien leeftijd)	---	---	---
Vertraagde schoolloopbaan (gezien leeftijd)	---	---	---
Geslacht leerling	---	---	---
VVE-indicatie	---	---	---
Thuis taal niet Nederlands	---	---	---
Dyslexie in de familie	---	---	---
Leerkrachtkenmerken			
Jaren werkervaring in primair onderwijs	---	---	---
Jaren werkervaring met kleuters	---	---	---
Tevredenheid met schoolcurriculum beginnende geletterdheid	---	---	---
Schoolkenmerken			
Vertrouwen schoolleider in leesonderwijs op school	---	---	---
Aantal methoden technisch lezen op school	---	---	---
Aantal methoden spelling op school	---	---	---
Aantal leerlingen op school	---	---	---
Aantal leerlingen in groep 2	---	---	---
Schoolweging	---	---	---
Stedelijkheid schoollocatie	---	---	---
Leerkrachtentekort groep 1-2	---	---	---
Onderwijsassistententekort groep 1-2	---	---	---
IB-tekort groep 1-2	---	---	---

Leescoördinator op school	---	---	---
Denominatie school: protestants-christelijk	---	---	---
Denominatie school: rooms-katholiek	---	---	---

Bijlage 5 Pilot eyetracking

Als de kleuters met KlankKr8 werken op de devices op school of thuis, dan geeft de backenddata van KlankKr8 informatie over de duur van speelsessies, de frequentie van het spelen en reactiesnelheid van ieder kind. Het doel van de pilot eyetracking was om een stap verder te gaan door informatie ook de flow (concentratie, motivatie) tijdens het spelen te meten. Het idee achter een serious game als KlankKr8 is dat het kind dusdanig in het spel opgaat dat hij/zij het wil blijven spelen en zo onbewust letterkennis opdoet. Een face-tracking meting is te beschouwen als een proxy van de flow.

In eerste instantie hebben we voor het meten van de flow de toepassingsmogelijkheden van *Tobii eye tracker glasses* onderzocht. De enige beschikbare bril van Tobii (Zie figuur B1 links) is ontwikkeld voor volwassenen en bleek te groot voor kinderen van ongeveer 5 jaar oud: hij zakte af en kon de pupilbewegingen niet adequaat volgen.

Wearable eye trackers (glasses)

Glasses 3



Figuur B1: Bril voor eyetracking en camera op voorhoofd voor face-tracking.

Vervolgens hebben we zelf een (simpeler) alternatief met bijbehorende software ontwikkeld: een *webcam face-tracking applicatie* (ziein de figuur rechts). Deze applicatie gebruikt een techniek waarbij een kleine camera op het voorhoofd van het kind registreert waar het kind naar kijkt als het speelt op een tablet met KlankKr8. Zolang tenminste één gemarkeerde hoek van de tablet in beeld is, indiceert dit dat een kind geconcentreerd aan het spelen is. Als dit niet het geval is en dit duurt geruime tijd, dan is het kind afgeleid en de flow (concentratie) tijdelijke weg. Met de bijbehorende software kan worden vastgesteld voor welke deel van de totale speeltijd dit het geval is. Een face-tracking meting is in principe te beschouwen als een goede indicator van de flow.

We hebben de *webcam face-tracking applicatie* toegepast bij 15 groep 2 leerlingen op reguliere basisscholen. De kinderen werden uit de groep gehaald en speelden in een rustige ruimte of op de gang met KlankKr8 op een tablet. De aanwezige onderzoeker plaatste de hoofdband met camera op het hoofd van het kind en zette de programmatuur aan: de camera en de laptop voor

de videoregistratie van de opname. Het kind ging verder spelen daar waar het de vorige keer gestopt was en speelde 15-20 minuten, hetgeen overeenkomt met ongeveer drie sessies. Om privacy redenen is er geen videoregistratie gemaakt van het kind en/of omgeving, alleen maar van waar het kind naar kijkt.

Uit de resultaten hebben we moeten concluderen dat het face-tracking bij jonge kinderen problematisch is. Uit de persoonlijke observaties van de onderzoeker bleek dat alle kinderen tijdens het spelen op de tablet heel erg beweeglijk en makkelijk afleidbaar zijn. Ze kijken zeker niet de gehele tijd naar het scherm, maar zij spelen tegelijkertijd wel geconcentreerd: ze weten wat ze (moeten) doen of waar ze moeten klikken voor het juiste antwoord als ze weer wél kijken naar het scherm. De *webcam face-tracking flow* meting van hoeveel tijd en hoe vaak tenminste één hoek van het scherm in beeld is, bleek geen valide meting van de flow. De persoonlijke observaties van de aanwezige onderzoeker van het speelgedrag van het kind bleek daarvoor geschikter. Dertien van de vijftien kinderen waren ondanks het regelmatig wegstijgen van het scherm, wel geconcentreerd en gemotiveerd met KlankKr8 aan het spelen. De twee overige kinderen waren dat niet: ze vroegen na korte tijd hoe lang ze nog moesten spelen, klikten zomaar ergens op om het muziekje te horen, of verzonnen een smoesje om weer terug te kunnen gaan naar hun groep. De beschikbare backenddata van KlankKr8 levert betrouwbaardere informatie op over het klikgedrag van een leerling dan de face-tracking.

Bijlage 6 Aanvullende analyses met pre-test letterkennis in groep 2 als controlevariabele

Om de robuustheid te testen van de analyses met verschilscores als uitkomstvariabele zijn deze analyses eveneens uitgevoerd met score op de post-test als afhankelijke variabele en de pre-test als controle. Dit is gedaan voor de uitkomsten naar het verschil in leerwinst letterkennis tussen de interventiescholen en de controlescholen in groep 2 (zie tabel. 4.3) en de samenhang tussen Klankkr8-gebruik met leerwinst letterkennis in groep 2 (zie tabel 4.6).

Wat betreft het verschil tussen beide groepen blijkt de pre-test score (letterkennis in januari) naast de interventie de enige variabele die een significante samenhang laat zien met de afhankelijke variabele (letterkennis in juni). Het effect van de interventie is in deze analyse eveneens significant, maar wel kleiner dan het effect dat gerapporteerd is in tabel 4.3.

Verskil leerwinst letterkennis tussen interventiescholen en controlescholen met post-test letterkennis als uitkomstvariabele

	Coëfficiënt	Significantie
Fixed Effects		
Intercept	12.208	.001
Experimenteel-Controle*	.450	.047
Pre-test letterkennis	.516	.001
Random Effects		
Variantie op leerlingniveau	5.481	.001
Variantie op schoolniveau	.676	.001

Tweezijdige significantieniveaus

*Eenzijdig significantieniveau voor verschil Interventie vs. controlescholen

Wat betreft de samenhang tussen Klankkr8-gebruik (aantal gespeelde sessies op school) en letterkennis is het resultaat vrijwel gelijk aan de analyses met leerwinst als uitkomstvariabele (.021 vs. .022).

Samenhang Klankkr8-gebruik met post-test letterkennis als uitkomstvariabele

	Coëfficiënt	Significantie
Fixed Effects		
Intercept	11.581	.001
Aantal sessies op school*	.022	.001
Pre-test letterkennis	.464	.001
Toets fonologisch bewustzijn	.148	.001
VVE-indicatie	.592	.030
Random Effects		
Variantie op leerlingniveau	4.833	.001
Variantie op schoolniveau	.098	.259

Tweezijdige significantieniveaus

*Eenzijdig significantieniveau voor aantal sessies op school

Oberon

Postbus 1423, 3500 BK Utrecht

t 030 230 60 90

info@oberon.eu | www.oberon.eu

Utrecht, 14 november 2025

Dit project is (mede) gefinancierd door het NRO. Projectidentificatie KKOSOTUD

