

**Leesgerelateerde kenmerken van Leerlingen in het Speciaal Basisonderwijs:
Een Steekproefanalyse ter Voorbereiding op Effectonderzoek naar GraphoGame**

Student: Meike Wekamp (S4455789)

Masterthesis Orthopedagogiek

Faculteit der Gedrags- en Maatschappijwetenschappen

Rijksuniversiteit Groningen

Eerste begeleider: dr. B.J.A. de Groot

Tweede beoordelaar: dr. M.C.S. Paap

Datum: 9 juni 2025

Totaal aantal woorden: 8065

Abstract

In the Netherlands, reading skills are declining. Therefore, it is very important to conduct research into effective ways to stimulate reading development. Several interventions have already been developed and researched. However, there is a lack of research on special primary education. GraphoGame is a promising intervention concerning reading development. This digital learning environment aims to automate letter knowledge, naming speed and phonemic analysis through repeated practice. The effectiveness of this program was previously investigated in pupils in regular education. Given the characteristics of GraphoGame that seem to fit in extremely well with special primary education, an effect study will also be conducted in this population. This research serves as an exploration of the sample in preparation for this future effect study. The One-Minute Test, the Klepel-R, the FAT-R and the Continuous Naming & Reading Words were administered to 51 pupils from year 3, 4 and 5. The data from an earlier baseline measurement was also used. The results of this study indicate that the pupils within the sample score low on average on the administered tests compared to the norm group in regular education. The results also show that the experimental and control groups within this study are comparable on all variables. Finally, a correlation was found between phonemic awareness and continuous naming speed with (pseudo)word reading.

Samenvatting

De leesvaardigheid in Nederland daalt. Het is dan ook van groot belang om onderzoek te doen naar effectieve manieren om de leesontwikkeling te stimuleren. Meerdere interventies zijn reeds ontwikkeld en onderzocht. Echter ontbreekt het hierbij aan onderzoek binnen het speciaal basisonderwijs. GraphoGame is een veelbelovende interventie op het gebied van de leesontwikkeling. Deze digitale leeromgeving heeft als doel het automatiseren van letterkennis, benoemsnelheid en fonemische analyse door middel van herhaalde oefening. De effectiviteit van dit programma werd eerder al onderzocht bij leerlingen op het regulier onderwijs. Gezien de kenmerken van GraphoGame die uiterst goed lijken aan te sluiten op het speciaal basisonderwijs, zal er ook een effectonderzoek worden uitgevoerd bij deze populatie. Dit onderzoek dient als verkenning van de steekproef ter voorbereiding op dit toekomstige effectonderzoek. Bij 51 leerlingen uit de groepen 3, 4 en 5 werden de Eén-Minuuut-Test, de Klepel-R, de FAT-R en de Continu Benoemen & Woorden Lezen afgenomen. Ook werd er gebruik gemaakt van de data van een eerdere nulmeting. De resultaten van dit onderzoek wijzen uit dat de steekproef gemiddeld laag scoort op de afgenomen tests ten opzichte van de normgroep uit het regulier onderwijs. Tevens blijkt uit de resultaten dat de experimentele en controlegroepen binnen dit onderzoek vergelijkbaar zijn op alle variabelen. Tenslotte werd er samenhang gevonden tussen het fonemisch bewustzijn en de continue benoemsnelheid met het (pseudo-)woordlezen.

Inhoudsopgave

Abstract	2
Samenvatting.....	3
Inleiding.....	5
Methode	12
Resultaten.....	16
Discussie	24
Literatuurlijst.....	29
Bijlage	35

Inleiding

Het is een wereldwijd probleem: de daling van het niveau van leesvaardigheid. De uitslagen van de PISA-onderzoeken (Meelissen et al., 2023), die de vaardigheden van 15-jarigen meten op internationaal niveau, lieten de afgelopen jaren een afname van het gemiddelde leesniveau zien. Hoewel deze negatieve trend zichtbaar is in vele landen in Europa, valt Nederland op. Het leesniveau in Nederland daalde ten opzichte van de meting in 2018 zelfs zo sterk, dat Nederland recentelijk bijna alle andere landen van de Europese Unie voor moest laten gaan. Binnen het Nederlandse basisonderwijs zijn daarmee flinke zorgen over de leesvaardigheid. Ook uit het landelijke peilingsonderzoek van de Inspectie van het Onderwijs (2022) werd duidelijk dat het technisch leesniveau in het primair onderwijs in Nederland daalt. Uit hetzelfde onderzoek bleek bovendien dat deze daling in het speciaal basisonderwijs (sbo) sterker is dan in het reguliere basisonderwijs (bo). In het sbo behaalt 27% van de leerlingen niet het fundamentele niveau 1F voor technisch lezen. Ook bleek dat sbo-leerlingen met hetzelfde schooladvies als bo-leerlingen, minder vaak het 1F-niveau behalen.

De dalende trend in het leesniveau van kinderen heeft wereldwijd geleid tot verschillende onderzoeken en interventies gericht op het verbeteren van de leesvaardigheid. In Nederland werd bijvoorbeeld de methodiek *begeleid hardop lezen* onderzocht ter verbetering van het technisch leesonderwijs (Oostdam, Blok, & Boendermaker, 2011). Ook werd de interventie *Bouw!* ontwikkeld, met als doel de preventie van ernstige leesproblematiek (Zijlstra, 2015). Deze digitale interventie zet in op het ontwikkelen van inzicht in alfabetische principes en het fonemisch bewustzijn. *Bouw!* is zowel wetenschappelijk als in de praktijk bewezen effectief, en wordt inmiddels op ruim 5800 scholen toegepast.

In Finland werd een digitaal programma ontwikkeld dat risicolezers ondersteunt bij het leren van technisch lezen. Het adaptieve spel kreeg de naam GraphoGame en liet veelbelovende resultaten zien in de Finse taal (Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007). In 2018 werd het programma vertaald en aangepast naar het Nederlands. Het spel werd getest op het regulier basisonderwijs waarbij positieve effecten gevonden voor onder andere letterkennis en letterherkenning (Glatz, 2018).

GraphoGame werd niet getest op het speciaal basisonderwijs in Nederland. Vele kenmerken van het programma lijken echter goed aan te sluiten op het speciaal basisonderwijs. Het programma is bijvoorbeeld gebaseerd op het *multimodal nature of*

learning, waarbij informatie op verschillende manieren wordt aangeboden. Ook biedt de software adaptieve oefeningen op het niveau van elke leering. Dit onderzoek richt zich dan ook op de effectiviteit van GraphoGame binnen het speciaal basisonderwijs.

Hierna zal eerst aandacht worden besteed aan de theoretische basis het van technisch lezen in het Nederlands, en de onderliggende cognitieve processen van het woordlezen. Vervolgens wordt er gekeken naar GraphoGame en de Nederlandse adaptatie daarvan, en hoe dit programma aansluit op de leerlingpopulatie en de onderwijspraktijk van het sbo.

1.1 Het meervoudige route model van technisch lezen

Voor het lezen van woorden zijn er in verschillende theoretische modellen ontwikkeld. Een zeer bekend voorbeeld hiervan is het zogeheten *Dual Route Cascaded (DRC) model*, waarbinnen twee manieren van lezen worden onderscheiden (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001): (1) een lexicale route, waarbij indirecte woordherkenning plaatsvindt. De unieke volgorde van de letters legt direct een link naar het mentale woordenboek van de lezer; (2) een indirecte, sublexicale route waarbij alle letters worden gedecodeerd.

In het meervoudige route model van Struiksmā, Van der Leij en Vieijra (2012), dat duidelijk aansluit op het DRC-model, worden ook de directe en indirecte woordherkenning van elkaar onderscheiden. Deze auteurs beschrijven daarbij twee expliciete verschillende wegen binnen de indirecte woordherkenning. Bij de meest basale weg moet de lezer het woord volledig decoderen. De eerste stap is het opdelen van het woord in losse grafemen, ook wel visuele analyse genoemd. Grafemen zijn schrifttekens, soms bestaand uit een enkele letter en soms uit meerdere letters aan elkaar (bijvoorbeeld *ie*). De grafemen worden herkend door middel van visuele discriminatie. De lezer ziet aan de vorm van het grafeem het verschil tussen bijvoorbeeld de *d* en *b*. De lezer koppelt de grafemen in het woord aan de juiste fonemen. Een foneem is de kleinste klankeenheid die invloed heeft op de betekenis van het woord. Dit proces heet de grafeem-foneemkoppeling, maar wordt ook vaak aangeduid als klank-tekenkoppeling. Tenslotte worden de klanken in een ‘klankenstroom’ bij elkaar gebracht, en herkend als één klankvorm: het gesproken woord. Dit heet auditieve synthese. Temporele ordewaarneming, het onthouden van auditieve informatie in de juiste volgorde, speelt hierbij ook nog een belangrijke rol. Deze deelvaardigheid zorgt ervoor dat de lezer de klanken niet op een verkeerde plek in het woord plaatst.

Bij de tweede ‘tussenweg’ van indirecte woordherkenning herkent de lezer grotere woorddelen. Het woorddeel bestaat uit een specifieke combinatie van letters binnen een

woord. Naarmate de lezer een lettercombinatie vaak tegenkomt, wordt het woorddeel geautomatiseerde kennis. De lezer herkent het woorddeel in plaats van alle losse letters, waardoor de leeshandeling wordt verkort. Weg 2 maakt het lezen van niet-klankzuivere woorden mogelijk, waarbij het volledig verklanken van weg 1 zich juist beperkt tot klankzuivere woorden. Niet-klankzuivere woorden wijken af van het fonologische principe. Eén grafeem kent dan meerdere fonemen. De letter *-d* kan bijvoorbeeld ook klinken als een *-t* zoals in het woord *paard*. Bij niet-klankzuivere woorden is de grafeem-foneemkoppeling dus niet betrouwbaar (Nijakowska, 2010). Een taal met zeer consistente grafeem-foneemkoppelingen wordt aangeduid als transparante orthografie (Seymour, Aro, & Erskine, 2003). De Finse taal is zeer consistent en is dan ook een voorbeeld van een transparante orthografie. Daar tegenover staat het Engels, dat zeer veranderlijke grafeem-foneemkoppelingen kent. De Nederlandse taal is gemiddeld consistent en daarmee minder transparant dan het Fins, maar transparanter dan het Engels (Helland & Kaasa, 2005). De transparantie van de taal is bepalend in hoeverre weg 1 of 2 van het lezen wordt toegepast. Hierbij geldt: hoe minder transparant de taal, hoe meer er een beroep wordt gedaan op de indirecte woordherkenning.

1.2 De cognitieve basis van woordlezen

In de literatuur worden er drie typen cognitieve vaardigheden beschreven die onderliggend zijn aan het technisch lezen: fonemische analyse, continue benoemsnelheid en letterkennis (Van Bergen, De Jong, Plakas, Maassen, & van der Leij, 2012; Verhoeven, De Jong, & Wijnen, 2014; De Groot, Van den Bos, Minnaert, & Van der Meulen, 2015).

Fonemische analyse, of in het Engels *Phonemic* of *Phonological Awareness* (PA), wordt beschreven als het vermogen om de onderdelen van gesproken taal te herkennen en manipuleren, en dit toepassen op alfabetische kennis en kennis van sublexicale units van woorden (De Groot, 2015). Uit onderzoek blijkt PA vooral een sterke voorspeller voor het lezen van pseudowoorden (Van den Boer, Van Bergen, & De Jong, 2015). Continue benoemsnelheid, in het Engels *Rapid Automated Naming* (RAN) genoemd, is het vlug ophalen van de naam van bekende stimuli (De Groot, 2015). RAN beïnvloedt ook voornamelijk de woordleesvloeiendheid (Wolf & Bowers, 1999). Zowel voor het lezen van woorden als pseudowoorden is RAN een sterke voorspeller (Van den Boer et al., 2015). Bij verminderde RAN-vaardigheid is het activeren en ophalen van fonologische informatie tijdens het oplezen van woorden verstoord (Bexkens et al., 2014). Tenslotte wordt letterkennis (LK),

ofwel de kennis van grafeem-foneemkoppelingen, gezien als sterke voorspeller van het leesniveau (Van Bergen, De Jong, Plakas, Maassen, & van der Leij, 2012).

Wanneer er bij een individu sprake is van een ernstig én hardnekkig tekort in lees- en spellingsvaardigheden, en dat niet kan worden verklaard door het cognitief vermogen of het onderwijs, wordt dit dyslexie genoemd (Bexkens, Van den Wildenberg, & Tijms, 2014). Hoewel dyslexie zowel wordt gedefinieerd als specifieke leerstoornis (Lyon, Shaywitz, & Shaywitz, 2003) en als neurobiologische stoornis (Blomert, 2006), bestaat er consensus in de wetenschap over het feit dat tekorten in de fonologische verwerking ten grondslag liggen aan leesproblematiek (Brandenburg et al., 2017; Moll et al., 2016).

1.3 GraphoGame

Om risicolezers preventief te ondersteunen in het leren lezen, werd in Finland een digitaal spel ontwikkeld dat de naam GraphoGame kreeg. De ontwikkeling van het programma werd onder andere gebaseerd op kennis die voortkwam uit het *Jyväskylä Longitudinal study of Dyslexia* project (JLD), waarbij kinderen met familiair risico op dyslexie vanaf geboorte tot adolescentie werden gevolgd (Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009). Dit onderzoek wees uit dat Finse kinderen met leesproblematiek moeite hebben met het koppelen van klanken en letters (letterkennis) en met het manipuleren van spraakklanken. GraphoGame werd dan ook een interactief en adaptief spel dat zich focust op het aanleren van grafeem-foneemkoppelingen in de Finse taal (Richardson & Lyytinen, 2014). Interventieonderzoek liet verbetering in leesgerelateerde taken zien als gevolg van de toepassing van GraphoGame (Lyytinen et al., 2007). Hoewel deze aanpak dus goed leek te werken voor de transparante orthografie van het Fins, waren er aanpassingen nodig om de methode ook toepasbaar te maken op het Nederlands. Bij het leren lezen van minder transparante talen is er namelijk een grotere rol weggelegd voor het fonologisch bewustzijn en continue benoemsnelheid (Glatz, 2018). Aan het programma werden dan ook onderdelen toegevoegd die letterkennis, vloeiendheid en fonemische analyse trainen. Deze oefeningen bevatten de vaardigheden spellen en blenden (het samenvoegen van grafemen tot een woord of woorddeel). Glatz (2018) slaagde erin om een Nederlandse adaptatie van GraphoGame (GG-NL) te ontwikkelen.

Graphogame is geen vervanger van klassikale instructie, maar dient juist als aanvulling om de blootstelling aan geschreven tekst, en de onderwijstijd voor het technisch lezen te vergroten (Glatz, 2018). Het programma is een vorm van *digital game-based learning* (DGBL). Een van de belangrijkste kenmerken van DGBL is dat het past binnen de

multimodal nature of learning (Lampropoulos, Anastasiadis, & Siakas, 2017). Hierbij wordt informatie in verschillende vormen gerepresenteerd, bijvoorbeeld afbeeldingen, tekst, getallen, gebaren of haptische feedback (Taylor & Leung, 2019). Daarnaast is GraphoGame adaptief. Het niveau van het spel past zich aan op het niveau van de individuele leerling, waardoor het mogelijk wordt voor ieder om op eigen tempo te leren (Glatz, 2018). De speelsessies van GraphoGame duren niet lang, zo'n 10 tot 15 minuten (Glatz, 2018). Kortere sessies hebben een positief effect op de concentratie en interesse van de leerling, wat vervolgens indirect effect heeft op de leeropbrengsten van het spel (Richardson & Lyytinen, 2014).

1.4 Speciaal basisonderwijs

Het speciaal basisonderwijs (sbo) in Nederland bestaat voor leerlingen met extra behoefte aan zorg een aandacht, waaraan het reguliere basisonderwijs niet kan voldoen. In het jaar 2023 gingen 33.919 kinderen naar het sbo (Ministerie van OCW, 2024). Leerlingen met een extra ondersteuningsbehoefte, ook wel zorgleerlingen genoemd, zijn leerlingen die kampen met leerproblemen, sociaal-emotionele problemen en/of gedragsproblemen en daarom speciale aanpassingen of zorg vragen binnen het onderwijsaanbod (Ledoux et al., 2015). Sinds de invoering van Passend Onderwijs worden zorgleerlingen voor zover mogelijk bediend in het regulier onderwijs (Inspectie van het Onderwijs, 2024). Een belangrijk verschil tussen zorgleerlingen op het regulier en speciaal basisonderwijs is de omvang van de problematiek. Sbo-leerlingen typeren zich door complexere en zwaardere problematiek (Ledoux, Roeleveld, Van Langen, & Smeets, 2012). Uit onderzoek blijkt ook dat zorgleerlingen in het sbo lager scoren op cognitieve toetsen dan zorgleerlingen op het regulier onderwijs (Ledoux et al., 2015).

Vaak hebben sbo-leerlingen disharmonische profielen. Dit houdt in dat er significante verschillen zijn in de groei van ontwikkelingsdomeinen: het ene domein ontwikkelt zich goed, terwijl andere ontwikkelingsdomeinen achterblijven. Deze disharmonie kan ook voorkomen binnen de taal- en leesontwikkeling (Van Druenen, Scheltinga, Wentink & Verhoeven, 2017). Als voorbeeld noemen zij dat een leerling bijvoorbeeld een voldoende tempo heeft in letters verklanken, maar anderzijds een beperkte woordenschat heeft. De disharmonisch profiel van de leerling zorgt voor grote variatie binnen de sbo-populatie. Uit het recente peilonderzoek leesvaardigheid blijkt er in het speciaal basisonderwijs al veel werk wordt verzet voor het technisch lezen (Inspectie van het Onderwijs, 2022). Zo wordt er in het sbo gemiddeld 76 minuten besteed aan technisch lezen, waar dat in het reguliere

basisonderwijs 44 minuten zijn. Ook blijkt dat sbo-leerkrachten twee keer vaker van de methode voor technisch lezen afwijkt dan een bo-leerkracht. Ruim de helft van de sbo-leerkrachten differentieert het technisch lezen binnen niveau- of tempogroepen, met eventueel ook differentiatie bij de verwerking van de stof. Deze cijfers ondersteunen wat uit de literatuur al blijkt: de variatie binnen de populatie op het sbo vraagt extra inzet van de leerkracht, en extra aandacht voor differentiatie.

De karakteristieken van GraphoGame zijn aantrekkelijk voor meerdere onderwijsvormen, maar sluiten op het oog uitzonderlijk goed aan bij het speciaal basisonderwijs. De adaptieve software van het programma biedt ingebouwde differentiatie, zonder extra inspanning van de leerkracht. Vooral bij de diverse groepen in het sbo is dit waardevol. De korte sessies lijken goed passend bij de beperkte belastbaarheid van sbo-leerlingen en is bevorderlijk voor hun concentratie. Daarnaast geeft de korte duur van sessies de leerkracht flexibiliteit om GraphoGame in te zetten in de klas.

Desondanks is er nog geen onderzoek gedaan naar de effectiviteit van GraphoGame binnen het sbo in Nederland. Glatz (2018) deed dit wel voor het reguliere basisonderwijs in Nederland. Deze studie bekeek de effecten van GraphoGame bij 286 leerlingen in Nederland en België. Na een interventieperiode van zes weken vond hij op korte termijn verbeterde letterkennis voor kinderen met zwakke pre-testresultaten, en versnelling van de letterherkenning voor kinderen met sterke pre-testresultaten. Een systematische review naar de effecten van GraphoGame in diverse contexten vond geen betekenisvol effect op het woordlezen (McTigue, Solheim, Zimmer, & Uppstad, 2019). Wel leidde het onderzoek tot de vraag naar onderzoek van de effecten van GraphoGame in relatie tot de leercontext. In 2023 werd in Nederland en Vlaanderen een studie uitgevoerd ($n = 247$) waarbij de effecten van GraphoGame onder de loep werden genomen. Na afloop van de interventieperiode van 7 weken werden positieve effecten gerapporteerd voor letterkennis en vloeiend lezen (Glatz et al., 2023).

In het huidige onderzoek zullen de kenmerken van de huidige steekproef in kaart worden gebracht. De resultaten van deze verkennende analyse zullen worden ingezet bij het onderzoek naar de effecten van de inzet van GraphoGame. In het huidige onderzoek staan twee hoofdvragen met een aantal deelvragen centraal:

- (1) Wat zijn de kenmerken van de middenbouw sbo-leerlingen in de steekproef op het gebied van woordlezen en leesgerelateerde cognitieve vaardigheden?

- a. Wat zijn de kenmerken van deze leerlingen op het fonemisch bewustzijn en de continue benoemsnelheid?
- b. Wat is het niveau van woordlezen en pseudowoordlezen van deze leerlingen?
- c. In hoeverre hangen de leesgerelateerde cognitieve vaardigheden van deze leerlingen samen met hun niveau van (pseudo-)woordlezen?
- d. In hoeverre is er sprake van vooruitgang op het (pseudo-)woordlezen en de leesgerelateerde cognitieve vaardigheden van deze leerlingen op de nul- en voormeting?

(2) In hoeverre zijn de middenbouw sbo-leerlingen in de experimentele en controlegroep binnen de steekproef vergelijkbaar?

- a. In hoeverre zijn de experimentele en controlegroep vergelijkbaar op het fonemisch bewustzijn en de continue benoemsnelheid?
- b. In hoeverre zijn de experimentele en controlegroep vergelijkbaar op het woordlezen en pseudowoordlezen?
- c. In hoeverre verschillen de experimentele en controlegroep van leerjaar 4 in termen van vooruitgang op deze vaardigheid tussen de nul- en voormeting?

De populatie leerlingen op het speciaal basisonderwijs verschilt van de populatie op het reguliere onderwijs. Onderzoek duidt ook op niveauverschillen op cognitieve toetsen tussen deze leerlingen. Hypothese 1a is dan ook dat de gemiddelde scores van de sbo-leerlingen in de steekproef op de gemeten variabelen niet vergelijkbaar zijn met de normgroep in het reguliere onderwijs. Daarnaast wordt er verwacht (Hypothese 1b) dat de continue benoemsnelheid en het fonemisch bewustzijn van de leerlingen in de steekproef positief samenhangen met het niveau van (pseudo-)woordlezen. Tevens wordt verwacht (Hypothese 1c) dat, door het onderwijsaanbod in de verstreken tijd tussen de nul- en voormeting, de leerlingen in de steekproef groei laten zien op het fonemisch bewustzijn, de continue benoemsnelheid en het (pseudo-)woordlezen.

Wat betreft de tweede onderzoeksvraag wordt verwacht (Hypothese 2a) dat de experimentele en controlegroep vergelijkbaar zijn op alle gemeten variabelen. Ook wordt verwacht (Hypothese 2b) dat beide groepen eenzelfde groei laten zien op deze variabelen tussen de nul- en voormeting. De leerlingen zijn namelijk door de school op willekeurige basis verdeeld over verschillende groepen 4 en 5. Tevens wordt verwacht (Hypothese 2c) dat de leerlingen in de experimentele en controlegroep een vergelijkbare vooruitgang laten zien op het fonemisch bewustzijn, de continue benoemsnelheid en het (pseudo-)woordlezen.

Methode

Om de onderzoeksvragen van het huidige onderzoek te beantwoorden werd, ter voorbereiding op het interventieonderzoek naar de effecten van GraphoGame binnen het speciaal basisonderwijs, een descriptief en hypothese toetsend onderzoek uitgevoerd met een nul- en voormeting.

Populatie en steekproef

De beoogde populatie voor dit onderzoek bestaat uit leerlingen van het speciaal basisonderwijs in Nederland. Eén sbo-school in de provincie Utrecht was bereid deel te nemen aan het onderzoek. Eerder werd bij 41 leerlingen een nulmeting uitgevoerd. Hieronder waren 14 kinderen uit groep 3, 15 leerlingen uit twee groepen 4, en 12 leerlingen uit groep 5. Eén groep 4, bestaande uit 6 leerlingen, vormde bij de nulmeting de controlegroep. De voormeting in het kader van het huidige onderzoek vond zeven maanden na de nulmeting plaats. Deze werd uitgevoerd bij dezelfde 41 leerlingen en werd tevens aangevuld met een controlegroep van 10 leerlingen uit een tweede groep 5.

Onderzoeksinstrumenten en variabelen

In dit onderzoek zijn vier variabelen gemeten: woordlezen, pseudowoordlezen, continue benoemsnelheid en fonemische analysevaardigheid.

Woordlezen

De EMT-Klepel is een duo-test, bestaande uit de *Een-Minuut-Test* (EMT) van Brus en Voeten (1973) en de *Klepel-R* (Van den Bos, De Groot, & De Vries, 2019). De test meet de technische leesvaardigheid van kinderen van 7 tot en met 14 jaar oud. De leerling leest in één minuut zoveel mogelijk woorden hardop. De EMT bestaat uit bestaande woorden uit de Nederlandse taal. Deze test meet de decodeervaardigheid van beginnende lezers. Bij gevorderde lezers is er sprake van directe woordherkenning, en meet de test het prestatieniveau daarvan. De Klepel bestaat uit pseudowoorden: uitspreekbare woorden zonder betekenis. Voorbeelden zijn *vas* of *briekem*. Deze test meet bij zowel beginnende als gevorderde lezers de decodeervaardigheid, omdat er geen sprake is van directe woordherkenning.

De EMT als de Klepel bestaan elk uit 116 items. De tests worden altijd samen en in één sessie afgenomen, te beginnen met de EMT. De testleider houdt de incorrect gelezen woorden bij en noteert waar de leerling is geëindigd met lezen na één minuut. Het aantal

correct gelezen woorden binnen de tijd vormt de ruwe score, die wordt omgezet naar een normscore ($M = 50$, $SD = 10$). De normgroep is gebaseerd op leerlingen binnen het reguliere basisonderwijs. De normscores van de EMT en de Klepel vormen samen een indexscore voor de algemene woordleesprestatie.

De betrouwbaarheid van de EMT-Klepel is op verschillende manieren onderzocht en goed bevonden (Van den Bos et al., 2019). De paralleltestbetrouwbaarheid vond voor alle leeftijden een correlatiecoëfficiënt groter dan 0,81. De test-hertestbetrouwbaarheid vond ook correlaties boven 0,80. Eveneens was de interbeoordelaarsbetrouwbaarheid hoog. De begripsvaliditeit en criteriumvaliditeit zijn beide tests goed bevonden (COTAN, 1981; COTAN, 2019).

Benoemsnelheid

Om de benoemsnelheid te meten, werd de *Continu Benoemen & Woorden Lezen* (CB&WL) afgenomen (Van den Bos et al., 2007). De test is geschikt voor kinderen van 5 tot en met 16 jaar en bestaat uit twee delen. Het Continu Benoemen omvat vier onderdelen van elk 50 items: kleuren benoemen, cijfers benoemen, plaatjes benoemen en letters benoemen. Bij het onderdeel Monosyllabische woorden lezen (Monosyl) moet de leerling 50 woorden van één lettergreep zo snel en accuraat mogelijk hardop lezen. Er wordt gemeten hoeveel tijd de leerling nodig heeft om te benoemen wat er wordt getoond. De ruwe score bestaat uit de tijd in seconden, en wordt omgezet in een normscore ($M=50$, $SD=10$). De normering is gebaseerd op leerlingen binnen het reguliere basisonderwijs. De betrouwbaarheid en validiteit van de CB&WL is positief beoordeeld (Egberink, De Leng, & Vermeulen, 2010).

Fonemische vaardigheid

Tenslotte werd de Fonemische Analyse Test (FAT-R) afgenomen (De Groot, Van den Bos, & Van der Meulen, 2014). De test beoogt het meten van de fonemische analysevaardigheid van kinderen van 7 tot en met 14 jaar, van groep 3 in het basisonderwijs tot en met het tweede jaar van het voortgezet onderwijs. De FAT-R bestaat uit twee subtests. Een daarvan is foneemweglating (FW) waarbij de leerling een woord herhaalt en daarbij een foneem niet uitspreekt, zoals *voetbal* zonder *bal*, wat *voet* wordt. De andere subtest is foneemverwisseling (FV), waarbij de leerling twee namen voorleest waarvan de eerste letters zijn verwisseld. Bij de namen *Kees Bos* leest de leerling *Bees Kos* voor.

Beide subtests bestaan uit 12 items met daarnaast elk 3 voorbeelditems. De itemscore bestaat uit de ruwe tijdscore, of uit een penaltyscore bij een incorrect gegeven antwoord. De

penalty- en tijdscores worden samengevoegd tot een ruwe vaardigheidsscore ($M = 50$, $SD = 10$). De normgroep is gebaseerd op leerlingen binnen het reguliere basisonderwijs. Uit de ruwe scores volgen normscores met een 95%-betrouwbaarheidsinterval. De subschalen vormen samen één normscore, de FAT-R-index.

De FAT-R wordt digitaal afgenomen door een testleider. De instructies en items worden door de software voorgelezen. De testleider stuurt het programma en geeft aan of de items correct of incorrect zijn. Hierbij wordt gebruik gemaakt van de afnamehandleiding.

De betrouwbaarheid van de FAT-R kan beschouwd worden als goed (Kerkmeer & Leemans, 2016). De inter-itemrelaties toonden betrouwbaarheid aan voor zowel de subtests ($\lambda > 0,75$) als de FAT-R-index ($\lambda > 0,82$). Uit het test-hertestonderzoek volgden hoge correlaties tussen de eerste en tweede meting voor FW, FV en de FAT-R-index ($r = 0,81$; $r = 0,90$; $r = 0,89$). Ook de interbeoordelaarsbetrouwbaarheid was goed voor de subtests (FW: $\kappa = 0,76$; FV: $\kappa = 0,92$). De validiteit van de test kan ook worden beschouwd als goed, en wordt ondersteund de resultaten uit onderzoek met neventests, klinische groepen en via voorspellingen van verschillende criteria (Kerkmeer & Leemans, 2016).

Procedure

Dit onderzoek is uitgevoerd met toestemming van de Ethische Commissie van Pedagogische en Onderwijswetenschappen van de Rijksuniversiteit Groningen. Voorafgaand aan het onderzoek zijn er informatiebrieven gestuurd naar ouders van de kinderen in de beoogde steekproef. In de brief werd het doel van het onderzoek beschreven en werd er toestemming gevraagd. Ook werd benoemd dat op elk gewenst moment het onderzoekstraject kan worden afgebroken door kind of ouders. Hierna hebben de onderzoekers kennis gemaakt met de leerlingen door een dagdeel mee te lopen in de klas. Het beoogde doel hiervan was het verlagen van mogelijke spanning bij de leerlingen, ten gunste van de meting. De tests werden tijdens individuele sessies afgenomen. De verzamelde data werd geanonimiseerd en opgeslagen op een beveiligde server van de Rijksuniversiteit Groningen.

Analyse

Voor de statistische analyses is gebruik gemaakt van IBM SPSS Statistics 28. Ter voorbereiding op de analyse zijn de ruwe scores in de dataset omgezet naar normscores. Voor de missende data is gekozen voor *pairwise deletion*. Tijdens de dataverzameling bleek namelijk dat sommige leerlingen een beperkte testbaarheid kenden. De data voor de tests waartoe zij wel in staat waren is dus wel meegenomen in dit onderzoek.

Ten eerste zijn voor de gemeten variabelen de beschrijvende statistieken berekend. Dit betrof de gemiddelde normscores, standaarddeviaties en de mediaan. Deze gegevens zijn voor de nul- en voormeting berekend, per leerjaar en onderzoeksgroep. Vervolgens zijn hiërarchische regressieanalyses uitgevoerd om de samenhang van het (pseudo-)woordlezen met de FAT-R en de benoemtaken van de CB&WL te onderzoeken. Tevens werd er ANOVA voor herhaalde metingen uitgevoerd, om de vooruitgang van de leerlingen tussen de nul- en voormeting in kaart te brengen. Ook werd het verschil in groei tussen de experimentele en controlegroep van leerjaar 4 hiermee bekeken. Tenslotte is een niet-parametrische test, de Mann-Whitney test, uitgevoerd om de experimentele en controlegroepen te vergelijken op de gemeten variabelen. Er is gekozen voor een niet-parametrische test, omdat er niet wordt voldaan aan de assumpties voor een parametrische test.

Resultaten

In Tabel 1 zijn de beschrijvende statistieken van de normscores weergegeven voor groep 3. De bevindingen laten consistent lage scores zien ten opzichte van de normgroep ($M = 50$, $SD = 10$), waarop de schaalscores zijn gebaseerd. Dit suggereert dat de groep-3-leerlingen uit de steekproef gemiddeld gezien beschikken over een tragere benoemsnelheid en een lager fonemisch bewustzijn dan leerlingen uit regulier onderwijs van dezelfde leeftijd. Voor de variabelen Kleuren Benoemen, Cijfers Benoemen en Letters Benoemen wordt in de steekproef een standaarddeviatie gerapporteerd die aanzienlijk hoger lijkt dan in de normgroep. Dit duidt op een hogere variabiliteit in scores in de steekproef, ten opzichte van leeftijdsgenoten in de normgroep. Hetzelfde geldt voor de FAT-R-normscores bij de voormeting.

Tabel 1

Beschrijvende statistiek van de normscores op de onderzoeksvariabelen bij de nul- (T0) en voormeting (T1), weergegeven voor Groep 3

Variabele	Meetmoment	N	Gemiddelde	Mediaan	SD
Kleuren benoemen	T0	14	41.36	36.0	16.20
	T1	14	41.64	40.0	15.67
Cijfers benoemen	T0	14	33.00	29.0	12.60
	T1	14	38.29	34.5	14.36
Plaatjes benoemen	T0	14	39.64	39.5	11.19
	T1	14	40,29	38.0	10.83
Letters benoemen	T0	13	40.85	40.0	12.69
	T1	14	41.86	38.5	12.35
MONOSYL	T0	0			
	T1	5	45.20	39.0	12.05
EMT	T0	0			
	T1	0			
Klepel-R	T0	0			
	T1	0			
FAT-R	T0	14	24.36	20.0	8.75
	T1	11	33.82	32.0	13.56

In Tabel 2 is de beschrijvende statistiek van de normscores weergegeven voor groep 4. Ook voor deze groep geldt dat de gemiddelde normscores aanzienlijk onder het gemiddelde liggen. Deze groep scoort dus gemiddeld lager ten opzichte van hun leeftijdsgenoten in het regulier basisonderwijs. Dit geldt voor zowel de experimentele als de

controlegroep. Vooral het fonemische bewustzijn is bij de groep-4-leerlingen in de steekproef fors lager dan in de normgroep. Tevens is de standaarddeviatie binnen de steekproef op de variabelen Cijfers Benoemen (nameting) en MONOSYL hoger dan bij de normgroep in het regulier onderwijs.

Tabel 2

Beschrijvende statistiek van de normscores op de onderzoeksvariabelen bij de nul- (T0) en voormeting (T1) voor de experimentele en controlegroep, weergegeven voor Groep 4

Variabele	Meet-moment	Onderzoeks-groep	N	Gemiddelde	Mediaan	SD
Kleuren benoemen	T0	Controle	6	41.67	41.0	12.47
		Experimenteel	8	40.75	41.0	7.59
		Totaal	14	41.14	41.0	9.54
	T1	Controle	6	38.67	42.0	13.77
		Experimenteel	8	44.38	44.0	6.99
		Totaal	14	41.93	43.0	10.38
Cijfers benoemen	T0	Controle	6	34.67	37.0	8.02
		Experimenteel	8	40.38	40.0	10.31
		Totaal	14	37.93	39.0	9.52
	T1	Controle	6	44.50	50.0	16.45
		Experimenteel	8	45.88	46.0	11.45
		Totaal	14	45.29	46.0	13.23
Plaatjes benoemen	T0	Controle	6	41.00	43.5	8.32
		Experimenteel	8	38.25	40.0	8.58
		Totaal	14	39.43	40.5	8.26
	T1	Controle	6	37.33	39.0	7.15
		Experimenteel	8	40.38	40.0	8.28
		Totaal	14	39.07	39.5	7.68
Letters benoemen	T0	Controle	6	36.83	36.5	6.56
		Experimenteel	8	41.13	41.5	13.17
		Totaal	14	39.29	36.5	10.72
	T1	Controle	6	40.50	39.0	9.71
		Experimenteel	8	35.50	31.5	12.02
		Totaal	14	37.64	36.0	10.99
MONOSYL	T0	Controle	5	30.83	32.5	5.85
		Experimenteel	8	35.63	38.5	13.85
		Totaal	13	33.57	34.0	11.06
	T1	Controle	5	43.80	47.0	7.79
		Experimenteel	8	41.63	46.0	15.96
		Totaal	13	42.46	47.0	13.04
EMT	T0	Controle	5	36.00	36.0	5.00

	T1	Experimenteel	7	31.13	31.5	9.92
		Totaal	12	33.00	35.0	8.48
		Controle	5	44.00	44.0	3.61
		Experimenteel	7	42.86	42.0	10.76
		Totaal	12	43.33	43.5	8.26
Klepel-R	T0	Controle	5	37.60	34.0	7.83
		Experimenteel	4	36.38	35.0	10.42
		Totaal	9	36.85	34.0	9.17
	T1	Controle	5	45.20	41.0	11.35
		Experimenteel	4	48.50	48.0	9.98
		Totaal	9	46.67	41.0	10.24
FAT-R	T0	Controle	6	25.00	20.5	7.90
		Experimenteel	7	28.22	20.0	10.01
		Totaal	13	26.93	20.0	9.07
	T1	Controle	6	31.00	32.5	10.39
		Experimenteel	7	22.00	22.0	2.08
		Totaal	13	26.15	23.0	8.31

De beschrijvende statistiek van de normscores van de leerlingen uit groep 5 is weergegeven in Tabel 3. Uit de tabel is af te lezen dat bij de nulmeting enkel de experimentele groep is onderzocht. Bij de voormeting is de controlegroep toegevoegd aan de steekproef. Net als de groepen 3 en 4 uit de steekproef, laat ook groep 5 gemiddeld lagere scores zien dan de normgroep uit het reguliere onderwijs, waarop de normering van de tests is gebaseerd. Vooral voor de taak Letters Benoemen van de CB&WL, en de FAT-R, valt op dat de groep-5-leerlingen uit de steekproef gemiddeld lager scoren dan hun leeftijdsgenoten op het regulier onderwijs. Voor de taken Cijfers Benoemen en Monosyl van de CB&WL, de EMT en de Klepel valt op dat de standaarddeviatie hoger is dan bij de normgroep. Dit duidt op een grotere variabiliteit binnen de groepen 5 van steekproef.

Tabel 3

Beschrijvende statistiek van de normscores op de onderzoeksvariabelen bij de nul- (T0) en voormeting (T1) voor de experimentele en controlegroep, weergegeven voor Groep 5

Variabele	Meet-moment	Onderzoeks-groep	N	Gemiddelde	Mediaan	SD
Kleuren benoemen	T0	Experimenteel	12	37.83	41.0	10.64
	T1	Controle	10	48.00	46.5	9.50
		Experimenteel	11	45.09	46.0	9.73
		Totaal	21	46.48	46.0	9.50

Cijfers benoemen	T0	Experimenteel	12	35.67	34.0	11.65
	T1	Controle	10	43.10	42.0	14.34
		Experimenteel	11	43.91	43.0	14.69
		Totaal	21	43.52	43.0	14.17
Plaatjes benoemen	T0	Experimenteel	12	36.83	36.0	9.25
	T1	Controle	10	43.80	41.5	7.50
		Experimenteel	11	46.91	43.0	8.47
		Totaal	21	45.43	43.0	7.98
Letters benoemen	T0	Experimenteel	12	35.92	34.0	11.26
	T1	Controle	10	32.20	32.0	9.20
		Experimenteel	11	31.27	26.0	12.49
		Totaal	21	31.71	31.0	10.79
MONOSYL	T0	Experimenteel	12	35.83	37.0	12.83
	T1	Controle	10	36.50	33.0	14.62
		Experimenteel	11	39.36	40.0	13.74
		Totaal	21	38.00	39.0	13.88
EMT	T0	Experimenteel	12	37.15	37.0	12.55
	T1	Controle	9	37.00	42.0	14.30
		Experimenteel	10	39.60	39.0	10.86
		Totaal	19	38.37	40.0	12.31
Klepel-R	T0	Experimenteel	12	38.92	37.5	17.33
	T1	Controle	8	42.13	47.5	15.55
		Experimenteel	9	44.78	42.0	12.50
		Totaal	17	43.53	45.0	13.63
FAT-R	T0	Experimenteel	12	28.83	23.0	11.77
	T1	Controle	10	29.40	29.5	8.70
		Experimenteel	9	31.11	31.0	11.66
		Totaal	19	30.21	31.0	9.95

Om te bepalen wat mogelijke voorspellers zijn voor het woordlezen, is vervolgens een hiërarchische regressieanalyse uitgevoerd, waarbij de normscore van de taak MONOSYL van de CB&WL als afhankelijke variabelen is genomen. De resultaten zijn af te lezen in Tabel 4. De MONOSYL, waarbij éénlettergrepige woorden gelezen worden, werd namelijk vaker afgenomen dan de EMT. Uit de analyse blijkt dat de leeftijd in maanden (model 1) geen significante voorspeller is van de score op de MONOSYL. Wanneer de normscore op de FAT-R wordt toegevoegd (model 2), blijkt de voorspellende waarde wel significant. Het woordleesniveau van éénlettergrepige woorden wordt voor 46.1% verklaard door de leeftijd in maanden en het fonemisch bewustzijn. Vervolgens werden ook de normscores op de benoemtaken van de CB&WL toegevoegd (model 3). De combinatie van leeftijd in maanden,

het fonemisch bewustzijn en de continue benoemsnelheid verklaart 73.2% van de variantie in het lezen van éénlettergrepige woorden.

In Tabel 5 wordt weergegeven hoeveel unieke variantie wordt verklaard door de afzonderlijke voorspellers in het model. De leeftijd in maanden blijkt in geen enkel model een significante voorspeller te zijn. In model 2 is het fonemisch bewustzijn een significante voorspeller van MONOSYL, waarbij 36.2% uniek wordt verklaard door deze variabele. In model 3, waarbij de continue benoemsnelheid is toegevoegd, is het fonemisch bewustzijn niet langer een significante voorspeller van het lezen van éénlettergrepige woorden. In dit model zijn enkel nog de taken Kleuren Benoemen en Cijfers Benoemen van de CB&WL significante voorspellers. De unieke bijdragen van deze variabelen aan de verklaarde variantie bedragen respectievelijk 13.9% en 20.7%.

Tabel 4

Resultaten regressieanalyse met woordlezen (éénlettergrepig) als afhankelijke variabele, en respectievelijk leeftijd in maanden, FB en CB als voorspellers

Model	R	R ²	ΔR^2	p-waarde
1	.307	.094	.094	.069
2	.679	.461	.367	<.001
3	.856	.732	.271	<.001

Tabel 5

Coëfficiënten, t-, p-waarden en unieke verklaarde varianties (sr^2) regressieanalyse met woordlezen (éénlettergrepig) als afhankelijke variabele

Model	Voorspeller	B	β	t	p-waarde	sr^2
1	Constante	75.452		2.705	.012	
	Leeftijd	-.353	-.242	-1.223	.233	.059
2	Constante	26.593		1.031	.313	
	Leeftijd	-.064	-.044	-.264	.794	.002
	FAT-R	.685	.634	3.796	.001	.362
3	Constante	11.652		.522	.608	
	Leeftijd	-.070	-.048	-.315	.756	.001
	FAT-R	.261	.242	1.565	.134	.028
	Kleuren Benoemen	-.584	-.541	-3.500	.002	.139

Cijfers Benoemen	.596	.644	4.268	.000	.207
Plaatjes Benoemen	.460	.325	1.995	.061	.045
Letters Benoemen	.187	.197	1.128	.273	.014

Ook voor het pseudowoordlezen werd een hiërarchische regressieanalyse uitgevoerd (Tabel 6). De normscores van de Klepel vormden daarbij de afhankelijke variabele. De leeftijd in maanden (model 1) blijkt geen significante voorspeller te zijn. Model 2, waarbij de FAT-R score is toegevoegd als voorspeller, is wel significant. De combinatie van de leeftijd in maanden en het fonemisch bewustzijn verklaart 27.4% van de variantie van het pseudowoordlezen. De toevoeging van het fonemisch bewustzijn leidt tot een extra verklaarde variantie van 17.8% ten opzichte van model 1. Wanneer de benoemtaken van de CB&WL worden meegenomen (model 3), blijkt het model niet meer significant.

Uit de gegevens in Tabel 7 blijkt, dat enkel binnen model 2 de FAT-R-score een significante voorspeller is. Dit duidt erop dat van de totaal verklaarde variantie ($R^2 = .274$) van het pseudowoordlezen, 17.8% uniek wordt verklaard door het fonemisch bewustzijn. De benoemtaken van de CB&WL, meegenomen in model 3, leveren geen significante unieke bijdrage aan de verklaarde variantie.

Tabel 6

Resultaten regressieanalyse met pseudowoordlezen als afhankelijke variabele, en respectievelijk leeftijd in maanden, FB en CB als mogelijke voorspellers

Model	R	R ²	ΔR^2	p-waarde
1	.309	.096	.096	.141
2	.523	.274	.178	.034
3	.728	.531	.257	.098

Tabel 7

Coëfficiënten regressieanalyse met pseudowoordlezen als afhankelijke variabele

Model	Voorspeller	B	β	t	p-waarde	sr^2
1	Constante	102.599		2.676	.014	
	Leeftijd	-.534	-.309	-1.526	.141	.095
2	Constante	75.446		2.031	.055	
	Leeftijd	-.446	-.258	-1.377	.183	.066
	FAT-R	.561	.425	2.268	.034	.178

3	Constante	18.257		.407	.689	
	Leeftijd	-.157	-.091	-.412	.685	.005
	FAT-R	.215	.163	.784	.444	.017
	Kleuren Benoemen	.462	.347	1.529	.145	.065
	Cijfers Benoemen	.263	.289	1.067	.301	.031
	Plaatjes Benoemen	-.089	-.058	-.254	.802	.002
	Letters Benoemen	.179	.154	.596	.559	.010

Omdat bij de voorgaande regressieanalyse er geen significante bijdrage van de Continue Benoemsnelheid bleek te zijn, werd er nogmaals een regressieanalyse uitgevoerd waarbij deze variabele in volgorde werd gewisseld met het fonemisch bewustzijn. Uit de resultaten van de nieuwe analyse blijkt dat model 2, waarin leeftijd in maanden en de benoemtaken van de CB&WL zijn meegenomen, 51.4% van de variantie van het pseudowoordlezen verklaart. Echter bleek er voor geen van de variabelen in deze hiërarchische regressieanalyse een significante unieke bijdrage.

Tabel 8

Resultaten regressieanalyse met pseudowoordlezen als afhankelijke variabele, en respectievelijk leeftijd in maanden, CB en FB als mogelijke voorspellers

Model	R	R ²	ΔR^2	p-waarde
1	.309	.096	.096	.141
2	.717	.514	.418	.019
3	.728	.531	.017	.444

Om de groei tussen de nul- en voormeting in kaart te brengen, is er gebruik gemaakt van *Repeated Measures ANOVA*. Ook werd hierbij gekeken naar mogelijke verschillen in groei tussen de experimentele en controlegroep van leerjaar 4. Voor de assumpties van deze test is voorafgaand aan de uitvoering gecheckt. De assumpties bleken niet allemaal op te gaan. Echter bleek Levene's test voor de meeste variabelen niet significant. Dit geeft aan dat de aanname van gelijke varianties niet is geschonden. Voor de variabelen Letters Benoemen en FAT-R was Levene's test wél significant. Hierdoor zijn die variabelen niet meegenomen in dit onderzoek.

Enkel voor de benoemtaak Kleuren Benoemen werd een interactie-effect gevonden tussen het meetmoment (nul- en voormeting) en de onderzoeksgroep (experimenteel of

controle-) voor groep 4 ($F = 3.45$, $p = .090$, $\eta p^2 = .239$). Hoewel dit interactie-effect niet statistisch significant is, maar er wel een relatief grote effectgrootte werd gevonden en er sprake is van een kleine steekproef, kan het interactie-effect mogelijk relevant zijn.

Tenslotte is de Mann-Whitney test uitgevoerd om te onderzoeken of de experimentele en controlegroepen van elkaar verschillen op de normscores. De resultaten hiervan zijn weergegeven in Tabel 9. De test bleek voor geen van de variabelen significant. Dit duidt erop dat de onderzoeksgroepen van de leerjaren 4 en 5 niet significant van elkaar verschillen op de gemeten variabelen.

Tabel 9

Resultaten van de Mann-Whitney test op basis van normscores

Groep	Variabele	<i>n</i>	U	p-waarde
4	Kleuren benoemen	14	33.5	.228
	Cijfers benoemen	14	24.5	1.00
	Plaatjes benoemen	14	30.0	.491
	Letters benoemen	14	17.0	.414
	MONOSYL	13	20.5	1.00
	EMT	12	15.5	.755
	Klepel-R	9	11.5	.730
	FAT-R	13	11.0	.181
5	Kleuren benoemen	21	47.0	.605
	Cijfers benoemen	21	55.5	1.00
	Plaatjes benoemen	21	70.5	.282
	Letters benoemen	21	49.5	.705
	MONOSYL	21	64.5	.512
	EMT	19	47.0	.905
	Klepel	17	34.0	.888
	FAT-R	19	47.0	.905

Discussie

In het huidige onderzoek is beoogd om de steekproef in kaart te brengen in een voorbereidende stap in het kader van het effectonderzoek naar GraphoGame in het speciaal basisonderwijs. Hierbij zijn verschillende eigenschappen van de leerlingen op het gebied van leesvaardigheid onder de loep genomen. Tevens zijn de kenmerken van de experimentele en controlegroep naast elkaar gelegd om te bepalen in hoeverre deze groepen vergelijkbaar zijn. In deze discussie zullen eerst de resultaten worden besproken. Vervolgens komen de inhoudelijke en methodologische beperkingen van het onderzoek aan bod.

Uit de resultaten is gebleken dat de sbo-leerlingen in de steekproef op vrijwel alle vlakken onder het gemiddelde scoren. De normgroepen van de meetinstrumenten die zijn gebruikt in het huidige onderzoeken bestaan allen uit leerlingen die regulier onderwijs volgen. De resultaten suggereren dus dat de sbo-leerlingen in de steekproef over een lager niveau van (pseudo-)woordlezen en de daarvan onderliggende cognitieve vaardigheden beschikken dan hun leeftijdsgenoten binnen het reguliere onderwijs. Dit is in lijn met eerder onderzoek, waarin werd aangetoond dat zorgleerlingen in het speciaal basisonderwijs lagere scores behalen op cognitieve toetsen (Ledoux et al., 2015) en ondersteunen hypothese 1a. De lage gemiddelden op de benoemtaken van de CB&WL houden in dat de leerlingen uit deze sbo-steekproef moeite hebben met het snel ophalen van bekende stimuli. Een lagere continue benoemsnelheid verstoort het activeren en ophalen van fonologische informatie tijdens het lezen (Bexkens et al., 2014). Dit heeft een negatieve invloed op zowel de vaardigheid en de vloeïendheid van het woordlezen (Wolf & Bowers, 1999). De verminderde continue benoemsnelheid van de leerlingen in de steekproef vormt dan ook een risicofactor voor hun leesvaardigheid. Daarnaast scoort de steekproef lager dan gemiddeld op de FAT-R. Dit houdt in dat zij meer dan gemiddeld moeite hebben met het herkennen en manipuleren van gesproken taal (De Groot, 2015). Ook dit vormt een risicofactor voor de leesvaardigheid.

In de huidige steekproef werden bij allerlei variabelen een grotere standaarddeviatie gevonden dan gemiddeld is binnen de normgroep van het regulier onderwijs. Binnen de steekproef is er dus sprake van een hogere variabiliteit in scores. Dit sluit aan op de bevindingen van Van Druenen et al. (2017). Zij beschreven hoe de disharmonische profielen van sbo-leerlingen, die ook binnen de taal- en leesontwikkeling bestaan, een grotere variatie in niveau binnen deze populatie veroorzaken. De grotere variatie kan binnen dit onderzoek hebben geleid tot de verhoogde gevonden standaarddeviaties. De variabiliteit van de

populatie kan mogelijk invloed hebben op de effectiviteit van GraphoGame en vormt daardoor een aandachtspunt voor effectonderzoek hiernaar.

Verder werd in het huidige onderzoek gekeken naar mogelijke samenhang tussen het (pseudo-)woordlezen en de leesgerelateerde cognitieve vaardigheden. Ook de leeftijd in maanden van de leerling werd hierin meegenomen. Er bleek echter dat de leeftijd geen unieke significante voorspeller is van het lezen. Andere voorspellers werden wel gevonden. Voor het van éénlettergrepige woorden bleek het fonemisch bewustzijn een significante voorspeller, met een unieke verklaarde variantie van 36.2%. De combinatie van de leeftijd en het fonemisch bewustzijn verklaarde 46.1% van de variatie op het woordlezen. Daarnaast werd ook de continue benoemsnelheid meegenomen in het model. De combinatie van deze drie variabelen resulteerde in een verklaarde variantie van 73.2%. Deze gevonden samenhang sluit aan op eerdere bevindingen, die uitwezen dat de continue benoemsnelheid een voorspeller is van het woordlezen (Van den Boer et al., 2015). Opvallend is dat niet alle benoemtaken van de CB&WL een significante, unieke bijdrage leverden aan de verklaarde variantie. Enkel de taken Kleuren Benoemen en Cijfers Benoemen bleken significante voorspellers van het woordlezen. De andere twee, Plaatjes Benoemen en Letters Benoemen, lijken dus niet samen te hangen met het niveau van woordlezen.

Ook voor het lezen van pseudowoorden werd mogelijke samenhang met het fonemisch bewustzijn, het continu benoemen en de leeftijd in maanden onderzocht. Ook hierbij had de leeftijd in maanden geen unieke bijdrage aan het verklaren van de variantie. De leeftijd en het fonemisch bewustzijn verklaren gezamenlijk 27.4% van de variantie binnen het pseudowoordlezen. Binnen dit model bestaat een unieke bijdrage van 17.8% van het fonemisch bewustzijn. Deze bevindingen worden ondersteund door eerdere literatuur, waarin wordt beweerd dat het fonemisch bewustzijn een sterke voorspeller is van het lezen van pseudowoorden (Van den Boer et al., 2015). Opvallend is dat het toevoegen van de continue benoemsnelheid ertoe leidde dat het model niet meer significant was. De combinatie van de drie variabelen blijkt dus geen samenhang te kennen met het lezen van pseudowoorden. Een model bestaande uit de leeftijd in maanden en het continu benoemen bleek wél significant. De combinatie van deze twee variabelen verklaart 51.4% van de variantie van het pseudowoordlezen. Ook deze samenhang werd eerder al beschreven door Van den Boer et al. (2015).

Verder is in het huidige onderzoek getracht de groei in ruwe scores van de leerlingen tussen de nul- en voormeting in kaart te brengen. Er werd echter geen significante vooruitgang gerapporteerd. Mogelijke verklaringen hiervoor worden verderop in deze sectie

bekeken. Tevens werd de vooruitgang van de experimentele en controlegroep met elkaar vergeleken. Eén interactie-effect bleef net buiten de grenzen van significantie, maar kende weliswaar een hoge effectgrootte ($F = 3.45$, $p = .090$, $\eta p^2 = .239$). Dit interactie-effect duidt op een verschil in vooruitgang tussen de experimentele en controlegroep van leerjaar 4. Gezien de methodologische beperkingen van het huidige onderzoek kan het waardevol zijn om deze gegevens in toekomstige analyses in overweging te nemen.

Tenslotte zijn de experimentele en controlegroepen vergeleken. Hierbij werd beoogd te onderzoeken of deze groepen vergelijkbaar zijn op de leesgerelateerde cognitieve vaardigheden, en op het niveau van (pseudo-)woordlezen. De resultaten van deze groepen werden voor de leerjaren 4 en 5 met elkaar vergeleken. De uitkomsten van de non-parametrische tests duiden op geen significant onderscheid tussen beide onderzoeksgroepen, zowel voor het (pseudo-)woordlezen als voor de cognitieve vaardigheden. De vergelijkbaarheid van de onderzoeksgroepen is van grote waarde voor het vervolgonderzoek naar de effecten van GraphoGame.

Hoewel het huidige onderzoek sterke punten kent, zijn er ook een aantal beperkingen die de interpretatie van de resultaten van het onderzoek bemoeilijkt. De voornaamste beperking is de steekproefgrootte. Voor het huidige onderzoek werd één school gevonden die bereid was om mee te werken. Hierdoor werd het aantal participanten beperkt tot de leerlingen op deze school. Verder bleken niet alle leerlingen testbaar op alle onderdelen van dit onderzoek. Dit resulteerde in missende data op verschillende tests. Vooral de afname van de Klepel-R bleek voor meerdere leerlingen niet haalbaar. Voor leerlingen uit groep 3 bleken ook de EMT en de MONOSYL van de CB&WL vaak van een te hoge moeilijkheidsgraad. Hierdoor ontstond een systematische uitval van data aan de ondergrens.

De kleine steekproefgrootte heeft de uitkomsten van het onderzoeken op verschillende manieren negatief beïnvloed. Zo is bekend dat een kleinere steekproef de statistische power van analyses vermindert en kan leiden tot toevallige afwijkingen (Garwood, Varghese & Vernon-Feagans, 2017). Hierdoor zijn wellicht resultaten gerapporteerd die bij een grotere steekproef anders zouden zijn, of zijn mogelijke significante resultaten gemist. Dit gaat ten koste van de betrouwbaarheid van het onderzoek. Daarnaast gingen niet voor alle tests de assumpties op. De steekproef bleek te klein voor een parametrische tests, waardoor de non-parametrische Mann-Whitney test werd ingezet. Tevens kon voor twee variabelen geen *Repeated Measures ANOVA* worden uitgevoerd, omdat Levene's test uitwees dat voor deze variabelen de assumptie van gelijke varianties werd geschonden.

Een andere beperking van het onderzoek was het ontbreken van een controlegroep van leerjaar 3. Hoewel er voor de leerjaren 4 en 5 wel een experimentele én controlegroep deelnamen aan het onderzoek, bestond er binnen de deelnemende school maar één groep 3. Voor het huidige onderzoek houdt dit in dat er geen vergelijking kon worden gemaakt tussen groep-3-leerlingen in verschillende onderzoeksgroepen.

Het huidige onderzoek kent ook inhoudelijke beperkingen. De gestandaardiseerde meetinstrumenten die werden gebruikt, zijn elk gebaseerd op normgroepen uit het regulier onderwijs. Hierdoor is het onduidelijk of deze tests even betrouwbaar zijn wanneer ze worden afgenomen bij sbo-leerlingen. Deze populatie wordt gekenmerkt door hun extra ondersteuningsbehoeften op sociaal-emotioneel en cognitief vlak (Ledoux et al., 2015). Men kan zich afvragen of deze leerlingkenmerken invloed hebben op de afname en resultaten van de meetinstrumenten. Wellicht speelt de algemene intelligentie een rol bij bijvoorbeeld het begrijpen van de instructies, of heeft gedragsproblematiek mogelijk invloed op de concentratie en taakgerichtheid van de leerling tijdens de afname.

Het huidige onderzoek dient als voorbereiding op het effectonderzoek naar GraphoGame binnen het speciaal basisonderwijs. De resultaten van dit onderzoek geven richting aan de analyse en interpretatie binnen dit toekomstige studie. Er blijken een aantal aandachtspunten te zijn die van belang zijn voor het effectonderzoek. Zo blijkt het gemiddelde niveau van de steekproef op vrijwel alle variabelen lager dan dat van de normgroep uit het regulier onderwijs. Daarnaast werd een bijna significant interactie-effect gevonden tussen de experimentele en controlegroep uit leerjaar 4. Het in acht nemen van deze, en andere belangrijke bevindingen, zal ten goede komen aan de resultaten van dit toekomstige onderzoek.

Het zou bevorderlijk zijn voor de statistische power om onderzoek uit te voeren met een grotere steekproef van sbo-leerlingen. Dit is essentieel om de resultaten te kunnen extrapoleren naar de gehele populatie binnen het speciaal basisonderwijs. Daarnaast wordt het aanbevolen om scholen uit andere regio's toe te voegen aan de steekproef, wegens de generaliseerbaarheid van de uitkomsten van het onderzoek.

In dit onderzoek is aangetoond dat de sbo-leerlingen verschillen op de leesgerelateerde cognitieve vaardigheden en het (pseudo-)woordlezen, ten opzichte van hun leeftijdsgenoten binnen het regulier onderwijs. Derhalve zou het van waarde zijn om de effectiviteit van GraphoGame te vergelijken voor het reguliere en het speciale basisonderwijs.

Dit kan inzicht geven in mogelijke verschillen in de effecten van de interventie tussen leerlingen uit beide onderwijsvormen.

Het gemiddelde gemeten niveau van het lezen van de sbo-leerlingen onderschrijft nogmaals het belang van goede interventies voor de leesontwikkeling binnen deze populatie. Het effectonderzoek naar GraphoGame binnen het speciaal basisonderwijs is dan ook zeer wenselijk. De bevindingen van het huidige onderzoek zullen hieraan een waardevolle bijdrage leveren.

Literatuurlijst

- Bexkens, A., van den Wildenberg, W. P. M., & Tijms, J. (2014). Rapid automatized naming in children with dyslexia: Is inhibitory control involved? *Dyslexia*, 21(3), 212-234.
<https://doi.org/10.1002/dys.1487>
- Brandenburg, J., Kleszczewski, J., Schuchardt, K., Fischbach, A., Büttner Gerhard, & Hasselhorn, M. (2017). Phonological processing in children with specific reading disorder versus typical learners: factor structure and measurement invariance in a transparent orthography. *Journal of Educational Psychology*, 109(5), 709–726.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.
- Commissie Testaangelegenheden Nederland (1981). *Een-Minuut-Test*. Van
<http://www.cotandocumentatie.nl/beoordelingen>
- Commissie Testaangelegenheden Nederland (2019). *De Klepel-R*. Van
<http://www.cotandocumentatie.nl/beoordelingen>
- De Groot, B. J. (2015). *Neurocognitive Profiling of Children with Specific or Comorbid Reading Disabilities*. Garant Publishers.
- De Groot, B. J. A., van den Bos, K. P., Minnaert, A. E. M. G., & van der Meulen, B. F. (2015). Phonological Processing and Word Reading in Typically Developing and Reading Disabled Children: Severity Matters. *Scientific Studies of Reading*, 19(2), 166-181. <https://doi.org/10.1080/10888438.2014.973028>
- Egberink, I. J. L. De Leng, W. E., & Vermeulen, C. S. M. (2010). COTAN beoordeling 2010,

CB&WL [COTAN review 2010, CB&WL]. Opgevraagd van
www.cotandocumentatie.nl

Garwood, J. D., Varghese, C., & Vernon-Feagans, L. (2017). Internalizing Behaviors and Hyperactivity/Inattention: Consequences for Young Struggling Readers, and Especially Boys. *Journal Of Early Intervention*, 39(3), 218–235.
<https://doi.org/10.1177/1053815117706524>

Glatz, T. K. (2018). *Serious games as a level playing field for early literacy: A behavioural and neurophysiological evaluation*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.

Glatz, T., Tops, W., Borleffs, E., Richardson, U., Maurits, N., Desoete, A., & Maassen, B. (2023). Dynamic assessment of the effectiveness of digital game-based literacy training in beginning readers: a cluster randomised controlled trial. *PeerJ*, 11.
<https://doi.org/10.7717/peerj.15499>

Helland, T., & Kaasa, R. (2005). Dyslexia in English as a second language. *Dyslexia*, 11(1), 41–60. <https://doi.org/10.1002/dys.286>

Inspectie van het Onderwijs. (2022). *Peil.Leesvaardigheid: Einde (speciaal) basisonderwijs 2020-2021*. Geraadpleegd op 11 november 2024, van <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/themarapporten/2022/12/13/peil-leesvaardigheid-einde-basisonderwijs-en-speciaal-basisonderwijs-2020-2021/Rapport+Peil+Leesvaardigheid+einde+speciaal+basisonderwijs+2020-2021.pdf>

Inspectie van het Onderwijs. (2024). *Naar het gespecialiseerd onderwijs of niet?* Geraadpleegd op 11 november 2024, van <https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2024/04/17/rapport-naar-het->

gespecialiseerd-onderwijs-of-niet/Rapport+Naar+het+gespecialiseerd+onderwijs+of-
+niet.pdf

Kerkmeer, M., & Leemans, G. I. (2016). *FAT-R Psychometrische eigenschappen*. Pearson.

Geraadpleegd op 14 januari 2025, van [https://www.pearsonclinical.nl/pub/me-
dia/whitepapers/FAT-R_2-Deel-2-psychometrische-eigenschappen.pdf](https://www.pearsonclinical.nl/pub/media/whitepapers/FAT-R_2-Deel-2-psychometrische-eigenschappen.pdf)

Lampropoulos, G., Anastasiadis, T., & Siakas, K. V. (2017). Digital Game-based Learning in Education: Significance of Motivating, Engaging and Interactive Learning Environments. In *Global Connectivity and Learning across the Generations* (pp. 117–127). Solent University.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., . . .

Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, 54(6), 686–694. <https://doi.org/10.1111/jcpp.12029>

Ledoux, G., Roeleveld, J., Van Langen, A., & Smeets, E. (2012). *Cool Speciaal: Inhoudelijk rapport*. Kohnstamm Instituut.

Ledoux, G., Van Langen, A., Regtvoort, A., Smeets, E., Roeleveld, J., & Paas, T. (2015).

Onderwijsloopbanen, cognitieve en sociaal-emotionele ontwikkeling van leerlingen in het speciaal basisonderwijs en speciaal onderwijs: COOL Speciaal, tweede meting. Kohnstamm Instituut.

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a

science-based application: A learning tool for reading acquisition. *Scandinavian Journal Of Psychology*, 50(6), 668–675. [http://inclusiveeducationpdresources.com/-
documents/lyytinen_computer_game_for_ef_2009.pdf](http://inclusiveeducationpdresources.com/-documents/lyytinen_computer_game_for_ef_2009.pdf)

- Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A., & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology*, 59(2), 109–126. <https://doi.org/10.1027/-1901-2276.59.2.109>
- McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2019). Critically Reviewing GraphoGame Across the World: Recommendations and Cautions for Research and Implementation of Computer-Assisted Instruction for Word-Reading Acquisition. *Reading Research Quarterly*, 55(1), 45–73. <https://doi.org/10.1002-/rrq.256>
- Meelissen, M. R., Maassen, N. A., Gubbels, J., Van Langen, A. M., Valk, J., Dood, C., Derks, I., Zandt, M. I., & Wolbers, M. (2023). *Results PISA-2022 - An overview*. <https://doi.org/10.3990/1.9789036559461>
- Ministerie van OCW. (2024, maart). *Ontwikkeling van het aantal leerlingen in het primair onderwijs*. OCW in Cijfers. [https://www.ocwincijfers.nl/sectoren/primair-onderwijs/leerlingen/aantallen-ontwikkeling-aantal-leerlingen#:~:text=In%20het%20speciaal%20basisonderwijs%20\(sbao,644%20meer%20dan%20in%202022.](https://www.ocwincijfers.nl/sectoren/primair-onderwijs/leerlingen/aantallen-ontwikkeling-aantal-leerlingen#:~:text=In%20het%20speciaal%20basisonderwijs%20(sbao,644%20meer%20dan%20in%202022.)
- Moll, K., Göbel, S. M., Gooch, D., Landerl, K., & Snowling, M. J. (2016). Cognitive risk factors for specific learning disorder: processing speed, temporal processing, and working memory. *Journal of Learning Disabilities*, 49(3), 272–281. <https://doi.org/10.1177/0022219414547221>
- Nijakowska, J. (2010). Dyslexia in the Foreign Language Classroom. In *Multilingual Matters eBooks*. <https://doi.org/10.21832/9781847692818>

- Oostdam, R., Blok, H., & Boendermaker, C. (2011). Zwakke lezers effectief helpen met begeleid hardop lezen. *Tijdschrift Voor Orthopedagogiek*, 50(11), 574–583.
https://www.researchgate.net/profile/Ron_Oostdam/publication/265965104_Zwakke_lezers_effectief_helpen_met_begeleid_hardop_lezen/links/542d58040cf29bbc126d29d8.pdf
- Richardson, U., & Lyytinen, H. (2014). The GraphoGame Method: The Theoretical and Methodological Background of the Technology-Enhanced Learning Environment for Learning to Read. *Human Technology*, 10(1), 39–60. <https://doi.org/10.17011/ht-urn.201405281859>
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal Of Psychology*, 94(2), 143–174.
<https://doi.org/10.1348/000712603321661859>
- Sheriston, L., Critten, S., & Jones, E. (2016). Routes to Reading and Spelling: Testing the Predictions of Dual-Route Theory. *Reading Research Quarterly*, 51(4), 403–417.
<https://doi.org/10.1002/rrq.143>
- Struiksma, A. J., Van Der Leij, A., & Viejra, J. P. (2012). *Diagnostiek van technisch lezen en aanvankelijk spellen*. VU Uitgeverij.
- Taylor, S. V., & Leung, C. B. (2019). Multimodal Literacy and Social Interaction: Young Children's Literacy Learning. *Early Childhood Education Journal*, 48(1), 1–10.
<https://doi.org/10.1007/s10643-019-00974-0>
- Van Bergen, E., De Jong, P. F., Plakas, A., Maassen, B., & van der Leij, A. (2012). Child and parental literacy levels within families with a history of dyslexia. *Journal of Child Psychology and Psychiatry*, 53(1), 28-36.

- Van den Boer, M., Van Bergen, E., & De Jong, P. F. (2015). The specific relation of visual attention span with reading and spelling in Dutch. *Learning And Individual Differences*, 39, 141–149. <https://doi.org/10.1016/j.lindif.2015.03.017>
- Van den Bos, K. P., De Groot, B. J. A., & De Vries, J. R. (2019). *Klepel-R Handleiding*. Pearson Benelux.
- Van Druenen, M., Scheltinga, F., Wentink, H., & Verhoeven, L. (2017). Preventie van leesproblemen. Groep 1 en 2. Nijmegen: Expertisecentrum Nederlands.
- Verhoeven, L., De Jong, P., & Wijnen, F. (2014). *Dyslexie 2.0*. Garant.
- Vernooy, K. (2006). Het LISBO- en VLOT-PROJECT: leerkrachtgestuurde leesinterventieprojecten voor het speciaal basisonderwijs. *Tijdschrift Voor Orthopedagogiek*, 45, 162–169.
- Verschueren, K., & Koomen, H. (2016). *Handboek Diagnostiek in de leerlingenbegeleiding*. Garant Uitgevers.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological functions*. Cambridge, MA: Harvard University Press.
- Zijlstra, A. (2015). *Early grade learning: The role of teacher-child interaction and tutor assisted intervention* [PhD-proefschrift, Universiteit van Amsterdam].
https://pure.uva.nl/ws/files/2615928/166609_DEF_met_correctie_Proefschrift_Early_grade_learning_H_Zijlstra_compleet.pdf

Bijlage

Resultaten van de Mann-Whitney tests.





