

Op de grens van diagnose: de effecten van diagnoseprotocolwijzigingen voor dyslexieprevalentie

Student: N.H. Wiekens (s4477421)

Begeleider en eerste beoordelaar: dr. B.J.A. de Groot

Tweede beoordelaar: dr. A.M. Sluiter-Oerlemans

Rijksuniversiteit Groningen

Faculteit der Gedrags- en Maatschappijwetenschappen

Bachelorwerkstuk Pedagogische Wetenschappen

Juli 2025

Aantal Woorden: 7399

Abstract

The criteria for the formal diagnosis of the neurologically based learning disability dyslexia in the Netherlands have changed over the past 20 years. Over that same timespan the reported prevalence increased, this raises the question as to what extent this may be due to the changes in diagnostic protocols. Therefore, the goal of present study was to investigate possible differences, in terms of prevalence and cognitive profiles of dyslexia, resulting from the application of two different sets of diagnostic criteria, cf. PDDB 1.0 (2006) and PDDB 3.0 (2021), to one and the same group of Dutch primary school children. It was hypothesised that prevalence would be lower in PDDB 3.0 and that phonemic awareness would better predict reading ability in PDDB 3.0, whereas rapid automatised naming would better predict reading ability in PDDB 1.0.

A dual cross-sectional quantitative comparative design was applied to test our hypotheses for a study sample ($n = 235$) of primary school children, that was composed of an existing dataset from 2007, which was extended with current collected data. Included participants completed standardized Dutch reading(-related) tests, i.e., EMT, de Klepel-R, CB&WL, and FAT-R, in order to (1) gain insight in their word and pseudoword reading abilities, rapid automatised naming, and phonemic awareness, respectively, and (2) classify them into PDDB version-based subgroups. As to the latter, and based on the resulting test scores, 12 (46.2% of diagnosed) participants met the criteria of both protocols, 9 (34.6% of diagnosed) participants only fulfilled the criteria of PDDB 1.0, and 5 (19.2% of diagnosed) children only met those for the (current) PDDB 3.0 protocol. A McNemar test showed that the apparent difference in group size was not significant, however. Furthermore, it was found that the actual PDDB 3.0 protocol results in diagnoses where reading of existing words is better predicted by rapid automatised naming and reading of pseudowords is better predicted by phonemic awareness.

Due to methodological limitations, the resulting small numbers of children meeting the PDDB criteria in particular, interpretations of the results necessarily must remain tentative. However, the present study did show a possible effect of diagnostic criteria with regard to the prevalence rates, and the derived reading-related cognitive profiles of diagnosed students. Follow-up research, with a larger sample, is needed to further examine these effects.

Inleiding

De leesvaardigheid van Nederlandse kinderen daalt (PISA, 2022; PISA, 2018). Leesvaardigheid, als onderdeel van geletterdheid, is een belangrijke vaardigheid in de huidige maatschappij. Lezen is niet alleen een manier voor kinderen om de wereld om hen heen beter te begrijpen, maar dat wat anderen schrijven levert informatie, adviseert bij het maken van keuzes, en vormt de basis voor navigatie door een politiek systeem (Heath, 1991; Rintaningrum, 2009). Aan de basis van een sterke geletterdheid ligt het snel en accuraat kunnen lezen van woorden.

Bij het gros van de kinderen verloopt het leren lezen goed. Er is echter een groep bij wie dit om verschillende redenen niet goed gaat. Dyslexie is een leesstoornis, waarbij, ondanks instructie van hoge kwaliteit, de leesontwikkeling ernstig en hardnekkig stagneert. Het Nederlands Kwaliteitsinstituut Dyslexie definieert deze momenteel als volgt: “Dyslexie is een specifieke en hardnekkige lees- en spellingstoornis met een basis in de neurobiologische ontwikkeling, die niet verklaard kan worden door een algemeen leerprobleem, inadequaat onderwijs of sensorische beperkingen” (Tijms et al., 2021, p.5). Dyslexie verschilt daarmee (differentiaal-diagnostisch) van een onderwijsachterstand, want dit laatste is in principe te remediëren met behulp van goed onderwijs.

Centraal in het huidige onderzoek staan de veranderingen in het Protocol Dyslexie Diagnose en Behandeling (PDDB), oorspronkelijk uitgebracht in 2006 (PDDB 1.0) (Blomert, 2006) en recent voor de tweede keer herzien voor een derde versie (PDDB 3.0) (Tijms et al., 2021). In 2019 rapporteerde de Inspectie van het Onderwijs (2019) dat 7,5% van de leerlingen de basisschool met een dyslexieverklaring verlaat. Dit is hoger dan het prevalentiecijfer van 3,6% uit 2005, dat aan de basis ligt van PDDB 1.0 (Blomert, 2005). Hiernaast laat het CBS (2016) zien dat het percentage kinderen met dyslexie tussen de 7 en 12 jaar van 5,9% in de periode 2001-2008 gestegen is tot 7,7% in de periode van 2009-2015. Zowel de Inspectie van Onderwijs als het CBS laten hiermee dus een groei zien ten aanzien van het aantal gestelde dyslexiediagnoses. Dit roept de vraag op in hoeverre deze groei verklaard kan worden door de veranderingen in de formeel gehanteerde diagnostische protocollen en welke theoretisch-inhoudelijke en praktische implicaties dit mogelijk heeft in termen van de signalering en begeleiding van leerlingen met dyslexie. In de huidige studie wordt specifiek onderzocht in hoeverre deze veranderingen invloed hebben op het aantal gestelde dyslexiediagnoses en de voortvloeiende cognitieve leesgerelateerde profielen van gediagnostiseerde leerlingen.

In deze inleiding zal hierna eerst een leestheoretisch kader worden geschetst, waarmee kan worden beschreven hoe een kind leest en leert lezen. Daarbij wordt het voor de huidige

studie centrale *Dual Route Cascaded Model* (Coltheart et al., 2001) met betrekking tot de woordleesvaardigheid uitgelegd. Ook worden twee belangrijke onderliggende, leesgerelateerde cognitieve processen van de woordleesvaardigheid, i.c. benoemsnelheid en fonemisch bewustzijn, behandeld. Voorts wordt nog verder ingegaan op wat dyslexie is en welke theoretische verklaringsmodellen hiervoor bestaan. Vervolgens wordt gekeken naar hoe de criteria van het Protocol Dyslexie Diagnose & Behandeling (Blomert, 2006; Tijms et al., 2021) zijn veranderd en wordt deze verandering gekoppeld aan het theoretisch kader. Ten slotte zullen na een uitwerking van de probleemstelling twee onderzoeksvragen en bijbehorende hypothesen worden gepresenteerd.

Om te weten hoe een kind leert lezen is het van belang onderscheid te maken tussen vier verschillende manieren van woordlezen. Ehri (1999) beschrijft drie manieren waarop onbekende woorden gelezen kunnen worden. De eerste manier is fonologisch recoderen. Hierbij gebruikt de lezer kennis van grafeem-foneemkoppelingen om een woord opnieuw op te bouwen. Een foneem is de kleinste betekenisvolle klank eenheid in onze taal. Een grafeem is een manier om dit kleinste klankelement op te schrijven (Peterson & Pennington, 2012). Een kind kijkt naar een woord, ziet grafemen, koppelt dit aan fonemen, en bepaalt welke uitspraak van het woord hierbij hoort. Het kan voorkomen dat dit proces meermaals doorlopen moet worden, voordat deze koppeling beklijft. Bij de andere twee manieren van het lezen van onbekende woorden maakt de lezer gebruik van bekende woorden of bekende spellingspatronen en relateert deze aan het onbekende woord (Ehri, 1991). Hiervoor is een gedegen kennis van het alfabet nodig (Ehri, 2005). De vierde en meest gevorderde manier van woordlezen is het zogeheten *sight word reading*. Hierbij wordt een geheel geleerd woord (direct) herkend vanuit het langetermijngeheugen. Er hoeft niet meer gekeken te worden naar onderdelen binnen het woord. Wanneer elk grafeem met een bekend foneem correspondeert, is het makkelijk om deze snel op te halen. Bij oefening zullen woorden memoriseerd worden en het lezen ervan geautomatiseerd verlopen.

Coltheart en collega's (2001) beschreven het *Dual Route Cascaded model* voor woordlezen dat goed aansluit bij het werk van Ehri (2005) en beschrijft hoe het lezen van woorden met de ontwikkeling verschuift van een analytisch proces naar een herkenningproces. Deze theorie stelt dat bij het woordlezen twee processen in principe gelijktijdig bijdragen aan de mogelijkheid tot woordlezen: de lexicale niet-semantic route en de grafeem-foneemkoppelingroute. Beide routes beginnen bij het zien van (visuele eenheden van) een woord, waarbij de onderdelen worden herkend als letters, hierna splitsen de routes af. De eerste, grafeem-foneemkoppelingroute gaat de letters die herkend zijn

afzonderlijk af, deze worden individueel gekoppeld aan klanken die, in de context en volgens uitspraakregels, juist corresponderen. Wanneer dit voor alle letters is gebeurd, zijn de fonemen vastgesteld en kan een uitspraak bedacht worden. De tweede, lexicale, niet-semantic route vergelijkt de lettercombinatie met de volledige ‘mentale lijst’ aan bekende woordbeelden. Is hier een overeenstemming gevonden, dan wordt deze gekoppeld aan de fonemen die erbij horen en wordt van hieruit de uitspraak bedacht. Deze routes binnen het proces worden gelijktijdig afgegaan. Zodra alle fonemen gevonden zijn, via welke route ook, fout of correct, is het proces ten einde. Op deze manier heeft de snelheid van afzonderlijke routes een effect op uiteindelijke accuratesse. Het model benoemt ook een derde route: de lexicale semantische route. Deze combineert en ondersteunt de lexicale route door middel van de betekenis van de te lezen woorden (Coltheart et al., 2001).

Zowel Ehri's (1991) manieren van lezen als Coltheart en collega's (2001) *Dual Route Cascaded model* leveren verschillende manieren van lezen van echte woorden in vergelijking met pseudowoorden (niet-bestaande woorden als ‘zeedkneep’). Echte woorden kunnen herkend worden (*sight word reading* / lexicale route), terwijl niet-bestaande gedecodeerd moeten worden (fonologisch recoderen / grafeem-foneemkoppelingsroute).

Bovenstaand leesmodel beschrijft direct observeerbaar leesgedrag. Onderliggend aan dit leesgedrag liggen cognitieve processen van leerlingen. Naar deze processen wordt gekeken door een lens van leesgerelateerde vaardigheden. Deze vaardigheden zijn een cognitieve afspiegeling van bij de leesroutes betrokken mentale processen. De lexicale route maakt gebruik van het snel herkennen van woorden. Deze vaardigheid wordt continu benoemen genoemd. De grafeem-foneemkoppelingsroute maakt gebruik van de vaardigheid fonemisch bewustzijn.

Dit onderzoek kijkt naar de werking van diagnoseprotocollen voor dyslexie. Hiervoor wordt gekeken naar deze onderliggende processen van woordlezen en hoe deze zich verhouden tot dyslexie en hoe leerlingen met dyslexie anders lezen dan leerlingen zonder. Deze processen zullen hieronder verder uitgelegd worden.

De continue benoemsnelheid heeft betrekking op hoe snel kinderen bekende stimuli kunnen benoemen, denk hierbij aan cijfers, kleuren en memoriseerde woorden. Continue benoemsnelheid speelt een belangrijke rol bij het leren lezen. Er zijn vier theorieën over de relatie van benoemsnelheid tot lezen. Torgesen en Wagner (1987) stellen dat benoemsnelheid overeenkomt met de snelheid waarop fonologische informatie opgehaald kan worden. Dit maakt het deel van het fonologisch proces (Georgiou et al., 2008). Wolf en Bowers (1999) stellen dat benoemsnelheid juist gaat over het identificeren van grafemen zelf, niet de grafeem-

foneemkoppeling. Wanneer identificatie van de grafemen te langzaam gaat, kunnen deze hierna niet snel genoeg gekoppeld worden aan fonemen en kan geen woord gededuceerd worden. Manis, Seidenberg en Doi (1999) stellen dat benoemsnelheid gaat over arbitraire connecties tussen teken en klank. De relatie tussen de letter 'k' en de klank 'k' is willekeurig. Als dit snel herkend wordt, zullen woorden met onregelmatige spellingen ook snel herkend worden. De lezer heeft dan namelijk weinig aan het recoderen van grafeem-foneemkoppelingen, de lezer moet bijvoorbeeld weten dat de 'j' in *'jalapeño'* niet klinkt als een 'j' maar als een 'h'. Kail, Hall & Caskey (1999) schrijven dat benoemsnelheid gaat om algemene cognitieve verwerkingssnelheid, als deze hoger ligt zal het lezen ook sneller gaan. In dit onderzoek wordt continu benoemen gezien als het identificeren van iets bekends. Het kan hierbij onder andere gaan om een letter, cijfer of memoriseerd woord. Leerlingen met dyslexie hebben moeite met deze vaardigheid. Een automatiseringsprobleem bij deze leerlingen zorgt ervoor dat herkenning minder snel en vaker verkeerd verloopt (Ehri & Saltmarsh, 1995).

Fonemisch bewustzijn heeft betrekking op in hoeverre kinderen weten hoe letters, grafemen, gekoppeld zijn aan klanken, fonemen. Hiernaast betreft het de mogelijkheid om deze fonemen te manipuleren (Peterson & Pennington, 2012). Het leren van deze grafeem-foneemkoppelingen is van groot belang voor het kunnen analyseren welke klanken in een woord zitten (Wolf & Bowers, 1999). Wanneer dit niet goed beheerst wordt, is het voor leerlingen lastiger om pseudowoorden te lezen (Rack, Snowling & Olson, 1992). Omdat kinderen deze woorden nog niet hebben gezien, kunnen ze niet op herinnering terugvallen en moeten ze deze woorden recoderen. Hetzelfde gaat op voor onbekende bestaande woorden (Wolf & Bowers, 1999). In dit onderzoek wordt fonemisch bewustzijn gezien als het hebben en toepassen van kennis over grafeem-foneemkoppelingen. Door minder automatisering en woordherkenning moeten dyslectische lezers meer gebruikmaken van grafeem-foneemkoppelingen voor het lezen van woorden (Ehri, 2005). Toch scoren leerlingen met dyslexie ook op dit vlak slechter (Wolf & Bowers, 1999).

Om de opzet van de diagnoseprotocollen te begrijpen wordt hieronder de aard van dyslexie uitgelegd en wordt beschreven welke verklaringsmodellen bestaan voor dyslexie. Dyslexie is een leesstoornis gekenmerkt door problemen met woordherkenning, decoderen en spellen. Mensen met dyslexie kunnen ook andere leesuitdagingen hebben (APA, 2013). De problemen met woordherkenning zijn het meest kenmerkend. Kinderen met dyslexie lezen bestaande woorden beduidend beter dan pseudowoorden (niet bestaande woorden als 'zeedknep' en 'bleimelneur') (Roberts et al., 2011).

Lage vaardigheid bij het fonologisch verwerken van grafemen is een probleem bij kinderen die moeite hebben met lezen. Dit wil zeggen dat het voor deze kinderen lastig is om een klank bij de letters te vinden en om deze klanken tot een woord te vormen. De ontwikkeling van leren lezen bij kinderen met dyslexie verschilt van kinderen zonder dyslexie, doordat op dit vlak meer oefening nodig is om goed te leren lezen. Door moeite met grafeem-foneemkoppelingen worden woorden niet goed gememoriseerd en zullen kinderen langer bezig zijn met het decoderen van woorden en blijft minder breincapaciteit over voor begrip (Roberts et al., 2011). Onderzoek heeft aangetoond dat dyslectische kinderen geen verminderde algemene automatiseringsvaardigheden hebben, maar wel een verminderde benoemsnelheid en fonologische vaardigheid (Wimmer et al., 1998). Dit wordt tegengesproken door Nicolson en collega's (2001), zij rapporteren een lagere vaardigheid bij automatiseringsvaardigheden. Hieruit is op te maken dat er enige overlap is tussen benoemsnelheid en fonologische vaardigheid, en automatisering specifiek op leesgebied.

Door de jaren heen zijn er verschillende verklaring modellen geweest voor dyslexie. Eén theorie stelt dat er één cognitieve zwakte, specifiek op gebied van lezen, moet zijn voor kinderen met dyslexie. Zwakke fonologische vaardigheden vormen de basis van dyslectisch leesgedrag: een *Phonological Deficit Hypothesis* (Stanovich, 1988; Ramus, 2003). Wolf & Bowers (1999) stellen met hun *Double Deficit Hypothesis (DDH)* dat fonologische gebreken en een lage continue benoemsnelheid twee losstaande zaken zijn in relatie tot leesachterstanden. Ze kunnen los van elkaar voorkomen en de leesvaardigheid verminderen. Wanneer ze beide aanwezig zijn kan dit zorgen voor de grootste leesachterstanden.

Latere theorieën benadrukken dat leren lezen een interactief proces is van vele cognitieve factoren. Deze staan in verband met elkaar en zijn als geheel verantwoordelijk voor dyslexie. Een sterk fonemisch bewustzijn kan zo bijvoorbeeld compenseren voor een gebrek in woordenschat (Snowling & Göbel, 2011). Ten slotte stelt Pennington (2006) een *multiple deficit model* voor. Dit model stelt dat het complexe gedrag van dyslexie ontstaat vanuit de interactie van, belemmerende én beschermende, genetische en omgevingsfactoren. Deze factoren veranderen de ontwikkeling van cognitieve vaardigheden, wat resulteert in gedragssymptomen. Een combinatie van factoren is nodig om een stoornis voort te brengen. Ten slotte is het goed om te benoemen dat er geen concrete grens is voor wél of géén stoornis. Dit maakt elke gestelde grens arbitrair.

Het PDDB is een leidraad opgesteld voor Nederlandse professionals voor het diagnostiseren, indiceren en behandelen van dyslexie. De eerste versie van het protocol (PDDB

1.0) verscheen in 2006, de derde en actuele versie (PDDDB 3.0) in 2021 (Blomert, 2006; Tijms et al., 2021).

In het eerste Protocol Dyslexie Diagnostiek & Behandeling (Blomert, 2006) staat beschreven aan welke criteria een dyslexiediagnose moest voldoen.

1. Er moet sprake zijn van ernstige lees- en spellingsproblemen. Dit is het geval bij een score in de laagste 10% van het normgemiddelde lezen of in de laagste 16% van lezen en de laagste 10% van het normgemiddelde spellen.
2. Er wordt gekeken of kinderen op andere schoolvakken ook uitvallen, gezien dyslexie een specifiek lees- en spellingsprobleem is. Ook moet er bij lezen stagnatie zijn aangetoond, die niet verbetert bij extra begeleiding.
3. Er moet sprake zijn van een dyslexietyperend cognitief profiel. Hierbij moeten kinderen bij twee of meer van de volgende indicatoren in de laagste 10% scoren: accuratesse en/of snelheid van fonologische verwerking, accuratesse en/of snelheid van grafeem-foneemassociatie, snel serieel benoemen van cijfers en letters
4. Er mag geen sprake zijn van alternatieve verklaringen.

In de derde versie van het Protocol Dyslexie Diagnostiek & Behandeling (Tijms et al., 2021) zijn de criteria van dyslexiediagnose veranderd:

1. Er moet sprake zijn van ernstige lees- en spellingproblemen. Bij drie soorten scores wordt dit vastgesteld:
 - a. Een score in de laagste 6,7% op woordlezen en de laagste 10% op pseudowoord lezen
 - b. Een score in de laagste 10% op pseudowoord lezen en de laagste 6,7% op woordlezen
 - c. Een score in de laagste 6,7% op spellen en de laagste 10% op woordlezen
2. Er moet sprake zijn van hardnekkigheid. Dit wordt vastgesteld wanneer op drie (halfjaarlijkse) meetmomenten in de laagste 10% wordt gescoord op woordlezen.
3. Er mag geen sprake zijn van een algemeen leerprobleem.

Gekeken naar deze twee criterialijsten vallen twee verschillen op. De eerste versie van het protocol is minder streng als het gaat om percentielscores van leestests en kijkt naast deze scores ook naar het cognitieve profiel van het kind. De vraag die dit oproept is of deze verschillen zorgen voor andere diagnoses. Zijn er leerlingen die volgens het huidige protocol geen diagnose zouden krijgen, terwijl die dit 20 jaar geleden wel hadden gekregen en vice versa?

PDDDB 1.0 benoemt niet specifiek welk verklaringsmodel ten grondslag ligt aan de diagnosecriteria. Het leunen op het cognitief-typerend profiel (Blomert, 2006) geeft de indruk dat het een model zoals dat van Snowling en Göbel (2011) volgt, waarbij vast wordt gehouden

aan cognitieve factoren die verantwoordelijk zijn voor dyslexie. Dit staat in contrast met PDDB 3.0, die dit profiel niet meeneemt. Het protocol neemt dit niet mee vanwege het multifactoriële karakter van dyslexie (Tijms, 2021). Dit komt overeen met het *multiple deficit model* van Pennington (2006). Het verwijderen van het ‘dyslexietyperend cognitief profiel’-criterium maakt PDDB 3.0 minder streng. Dit profiel zorgde er namelijk voor dat leerlingen een zwakke fonologische verwerking en benoemsnelheid nodig hadden voor een diagnose. Volgens het *Dual Route Cascaded Model* (Coltheart et al., 2021) zijn dit leerlingen met zwakkere lexicale en grafeem-foneemkoppelingsroutes en dus zwakkere lezers. Het zou kunnen dat door exclusie van dit criterium (een deel van) deze leerlingen in het PDDB 3.0 geen diagnose zouden krijgen. Om hiervoor te compenseren is gekozen om de scores op lezen zwaarder te laten tellen. Dit zou voorkomen dat de nieuwe criteria leiden tot een toename in dyslexiediagnoses. De focus verschuift door de criteriumverandering meer van de onderliggende mentale processen naar het observeerbare leesgedrag. De vraag is of deze verandering daadwerkelijk recht doet aan het diagnoseproces. In hoeverre verandert de diagnosepreventie en in hoeverre veranderen de cognitieve profielen van gediagnostiseerde leerlingen bij invoering van de nieuwe diagnosecriteria in PDDB 3.0?

Op basis van een cohort van $n = 2.642$ kinderen met dyslexie is dit bevestigd. Volgens de oude criteria zou 69,6% een diagnose gekregen hebben, volgens de nieuwe 59,9% (Tijms et al., 2021). Tijms en collega's (2021) gaan niet in op een verandering van het cognitieve profiel van gediagnostiseerde leerlingen.

Door het strenger maken van de leesvaardigheidseisen zijn het zwakkere lezers die een dyslexiediagnose krijgen. Zoals eerder beschreven leunen zwakke lezers meer op herkenning (continu benoemen) dan sterke lezers, die meer op fonemisch bewustzijn leunen (Ehri, 2005). De verandering van het PDDB zou hierom als gevolg kunnen hebben dat de leesstrategie van de gemiddelde leerling met dyslexie verandert: meer gefocust op fonemisch bewustzijn dan op continu benoemen.

Het huidige onderzoek

In de huidige studie wordt specifiek onderzocht in hoeverre deze veranderingen invloed hebben op het aantal gestelde dyslexiediagnoses en de voortvloeiende cognitieve leesgerelateerde profielen van gediagnostiseerde leerlingen. Voor de onderzoeksteekproef is nagegaan in hoeverre de deelnemers voldoen aan de diagnostische criteria van PDDB 1.0 en/of PDDB 3.0. Op deze manier is vergeleken hoe deze subgroepen zich tot elkaar verhouden en in hoeverre deze andere criteria zorgen voor dyslexiegroepen met verschillende kenmerken.

Kennis over de kenmerken van leerlingen met dyslexie biedt handelingsgerichte aanknopingspunten voor de onderwijspraktijk. Hiernaast is kennis theoretisch-inhoudelijk relevant, omdat het inzicht geeft in de onderlinge verhoudingen van betrokken cognitieve processen tot, al dan niet juiste, diagnoses. Het is hierbij van belang dat bij leerlingen met dyslexie specialistische hulp wordt aangeboden (Blomert, 2006). De inconsistentie van het diagnoseproces vormt een grote uitdaging hierbij en kan zorgen voor *false negative* of *false positive* diagnoses. Dit kan zorgen voor situaties waarin leerlingen die dyslexiehulp nodig hebben dit niet ontvangen (Daniel et al., 2024). Kennis over deze effecten is daarmee ook relevant voor de praktijk, omdat het inzicht geeft op de legitimiteit van diagnoses en handvatten geeft voor behandeling op maat.

De eerste onderzoeksvraag luidt: “In hoeverre verschilt de prevalentie van leerlingen die in aanmerking zouden komen voor een dyslexiediagnose volgens de eerste versie van het PDDB ten opzichte van de derde versie?”

Verwacht wordt dat, gezien de strengere leeseisen en gezien eerder onderzoek een lagere prevalentie bij de derde versie liet zien (Tijms et al., 2021), ook in huidig onderzoek een lagere dyslexieprevalentie gevonden wordt volgens de derde versie van het PDDB (Hypothese 1).

De tweede onderzoeksvraag luidt: “In hoeverre verschillen de voorspellende waarden van fonemisch bewustzijn en continue benoemsnelheid voor leesvaardigheid van pseudo- en echte woorden bij leerlingen die volgens PDDB 1.0 in aanmerking zouden komen voor een dyslexiediagnose met leerlingen die volgens PDDB 3.0 in aanmerking zouden komen voor een dyslexiediagnose?”

Door het strenger maken van de leesvaardigheidseisen, als gevolg van een verklaringsmodel gericht op de multifactoriële aard van dyslexie (Pennington, 2006), zijn het per definitie zwakkere lezers die een dyslexiediagnose krijgen. De eerder beschreven leestheorie stelt dat sterkere lezers meer herkennen dan decoderen, zij leunen relatief meer op continu benoemen. Zwakkere lezers leunen relatief meer op fonemisch bewustzijn (Ehri, 2005). Continu benoemen wordt gebruikt in de lexicale route van het *Dual Route Cascaded model* en betreft het herkennen van woorden, dit komt meer overeen met het lezen van echte woorden. Fonemisch bewustzijn wordt gebruikt in de grafeem-foneemkoppelingsroute en betreft het decoderen van woorden en komt meer overeen met het lezen van pseudowoorden (Coltheart et al., 2001). PDDB 3.0 diagnostiseert zwakkere lezers, deze verandering zou als gevolg kunnen hebben dat de leesstrategie van de gemiddelde leerling met dyslexie verandert: meer gefocust op fonemisch bewustzijn dan op continu benoemen.

Verwacht wordt daarom dat, gezien strengere eisen op woordlezen cf. PDDB 3.0, de voorspellende waarde van fonemisch bewustzijn op leesvaardigheid hoger is voor PDDB 3.0 dan voor PDDB 1.0 en dat de voorspellende waarde van continue benoemen op leesvaardigheid hoger is voor PDDB 1.0 dan voor PDDB 3.0 (Hypothese 2a). Verder wordt verwacht dat scores op leesvaardigheid lager zullen uitvallen in PDDB 3.0 gezien deze versie hierop strenger is en dat scores op fonemisch bewustzijn en continue benoemensnelheid lager zullen uitvallen bij PDDB 1.0 gezien deze versie het cognitief-typerend profiel meeneemt als criterium (Hypothese 2b). Volgens het *Dual Route Cascaded Model* zou verwacht worden dat lagere leesvaardigheid samenhangt met lagere scores op fonemisch bewustzijn en continue benoemensnelheid. Het *multiple deficit model* herinnert dat hiernaast meer factoren meespelen met leesvaardigheid. Hierom wordt verwacht dat een strenger criterium op dit cognitieve profiel (PDDB 1.0) op een directe manier zorgt voor lagere scores op fonemisch bewustzijn en continue benoemensnelheid dan dat strengere criteria op algemene leesvaardigheid (PDDB 3.0) op een indirecte manier doen.

Methode

Design

Dit onderzoek maakt gebruik van een duaal cross-sectioneel kwantitatief vergelijkend ontwerp. Er is onderzocht wat de verschillen in leesgerelateerde vaardigheden (leesvaardigheid van pseudo- en echte woorden, fonemisch bewustzijn en continue benoemensnelheid) zijn tussen leerlingen die wel of niet in aanmerking komen voor een dyslexiediagnose volgens de criteria van PDDB 1.0 en PDDB 3.0. In de inleiding is per versie beschreven welke criteria van het PDDB van toepassing zijn voor een dyslexiediagnose. Figuur 1 geeft nog een beknopt overzicht van de gehanteerde classificatiecriteria.

Op basis van de toepassing van de criteria uit Tabel 1 op de testresultaten is de participanten op categoriën “PDDB1.0dys” en “PDDB3.0dys” de waarde “0” (geen diagnose) of “1” (wel een diagnose) toebedeeld.. Hieruit ontstaan vier subgroepen: participanten zonder mogelijke diagnose, participanten met twee mogelijke diagnoses, en twee groepen participanten met één mogelijke diagnose. Zo kan onderzocht worden hoe de aantallen van diagnosestellingen tussen protocollen verschillen, maar ook hoe de groepen leerlingen die gediagnostiseerd worden van elkaar verschillen.

Figuur 1

Criteria voor Diagnose volgens de Protocollen Dyslexie Diagnose en Behandeling

PDDB 1.0:	PDDB 3.0:
Criterium 1: Een T-score van 37 of lager op lezen (pseudo- en regulier woordlezen)	Criterium 1: A1. Een T-score van 35 of lager op woordlezen EN Een T-score van 37 of lager op pseudowoordlezen
Criterium 2: Een T-score van 37 of lager op twee van de volgende variabelen	A2. Een T-score van 37 of lager op woordlezen EN Een T-score van 35 of lager op pseudowoordlezen
2.A1 Foneemweglating	
2.A2 Foneemverwisseling	
2.B1 Benoemsnelheid letters	
2.B2 Benoemsnelheid cijfers	

Steekproef

De doelpopulatie van dit onderzoek betreft leerlingen uit de groepen vier tot en met acht van een Nederlandse basisschool in het regulier onderwijs. De cutoff van groep vier is gebaseerd op het feit dat vanaf deze groep dyslexiediagnoses gesteld kunnen worden in verband met de aantoonbaarheid van hardnekkigheid.

De onderzoeksteekproef is gebaseerd op de gegevens van een bestaande dataset, die in het kader van het huidige onderzoek recentelijk is aangevuld. In de periode 2007-2012 zijn landelijk representatieve onderzoeken uitgevoerd ter hernormering van de EMT, de Klepel-R, CB&WL, en FAT-R (zie Testmaterialen voor detailbeschrijvingen). Deze gecombineerde dataset is beschikbaar gesteld door de thesisbegeleider, dr. B.J.A. de Groot, voor gebruik in dit onderzoek. Deze dataset bevatte bruikbare gegevens van 1.748 leerlingen uit groep vier tot en met acht. Deze dataset werd recent aangevuld met een gelegenheidssteekproef van 126 leerlingen, afkomstig van vijf scholen in de regio Noord-Nederland, wat het totaal bracht op 1.874 leerlingen. Echter, enkel participanten bij wie zowel de EMT, de Klepel-R, de CB&WL en de FAT-R zijn afgenomen zijn geïnccludeerd. Na verwijdering van leerlingen die minstens één test missen konden uiteindelijk 235 participanten worden geïnccludeerd in de definitieve steekproef.

Procedure

In aanvulling op een bestaande dataset zijn, in het kader van het bachelor-werkstuk Academische Opleiding Leerkracht Basisonderwijs (AOLB) van de Rijksuniversiteit Groningen, in de periode maart-april 2025 aanvullende gegevens verzameld door een vijftal studentonderzoekers. De toestemming voor het betreffende onderzoek is geregeld met een consentformulier en getoetst aan de Ethische Commissie Pedagogische Wetenschappen en Onderwijskunde van de Rijksuniversiteit Groningen (dossiernummer 100709847) als onderdeel van het lopende onderzoek van begeleider dr. B.J.A. de Groot.

De studentonderzoekers hebben elk bij de leerlingen van één klas van 18 tot 26 leerlingen de bovenstaande tests afgenomen op de locatie van de school. Voorafgaand aan afname is toestemming gevraagd aan zowel schoolleiding als ouders/verzorgers van deelnemende kinderen. Toestemming is in de vorm van een handtekening onder een informatie- en toestemmingsbrief verkregen. Ook is via de betrokken groepsleerkrachten informatie en data ontvangen aangaande schoolresultaten, eventuele meertaligheid en eventuele leesstoornisdiagnoses. Dit in het kader van andere deelstudies binnen het overkoepelende onderzoeksproject. Na toestemming ging de procedure van dataverzameling als volgt. De onderzoekers spraken momenten, waarop de tests afgenomen zouden worden, af met de groepsleerkrachten in kwestie. Leerlingen zijn op deze momenten één voor één uit de klas gehaald en hebben op een afgezonderde plek onder begeleiding van de onderzoeker de tests gemaakt. Dit betrof in een eerste ronde de CB&WL, EMT en de Klepel-R en in een tweede ronde de FAT-R. Hiervoor is gekozen om de werkdruk per meetmoment voor kinderen te verlichten. Dit proces nam per klas ongeveer zes dagdelen in beslag. Van de verzamelde data is een gecombineerde dataset gemaakt, waarin normscores zijn berekend op basis van de ruwe scores.

Testmaterialen

In dit onderzoek wordt gekeken naar de verschillen tussen leerlingen die wel en leerlingen die niet in aanmerking komen voor een specifieke dyslexiediagnose. Deze groepen zijn vergeleken op basis van hun leesvaardigheid van pseudowoorden en echte woorden, fonemisch bewustzijn en continue benoemsnelheid. Deze variabelen zijn tevens gebruikt om de deelnemers in te delen in de subgroepen, zoals toegelicht in de Design-paragraaf.

Woordlezen. Voor het vaststellen van leesvaardigheid is gebruikgemaakt van de Een-Minuut-Test (EMT) (Brus & Voeten, 1973). Kinderen krijgen een lijst woorden te zien die geleidelijk lastiger worden. De onderzoeker houdt bij hoeveel woorden binnen een minuut

correct voorgelezen worden. Uit ruwe scores worden gestandaardiseerde T-scores berekend. In de Klepel-R handleiding (Van den Bos, De Groot & De Vries, 2019) wordt de theoretische basis voor de EMT en de Klepel-R uitgelegd, hierbij wordt ook ingegaan op de inhouds-, begrips-, en constructvaliditeit. Er wordt overtuigend bewijs geleverd dat de EMT aan deze validiteitseisen voldoet. De betrouwbaarheidscoëfficiënt van de EMT is .93, dit is een hoge betrouwbaarheid (Van Den Bos et al., 2019).

Voor het vaststellen van decodeervaardigheid is gebruikgemaakt van de Klepel-R (van den Bos et al., 2019). Kinderen krijgen hierbij een lijst pseudowoorden. De onderzoeker houdt bij hoeveel woorden binnen twee minuten correct voorgelezen worden. Uit ruwe scores worden gestandaardiseerde T-scores berekend. In de Klepel-R handleiding worden zowel de inhouds- en begripsvaliditeit als criteriumvaliditeit besproken (Van Den Bos et al., 2019). Voor al deze vormen van validiteit wordt overtuigend bewijs geleverd dat voldoende validiteit van de Klepel-R ondersteunt. De betrouwbaarheidscoëfficiënt van de Klepel is .90 ook dit is een hoge betrouwbaarheid (Van Den Bos et al., 2019). De COTAN beoordeelt deze test ruim voldoende (Egberink & De Leng, 2021).

Benoemsnelheid. Voor het vaststellen van de benoemsnelheid is de Continu Benoemen & Woorden Lezen-test (CB&WL) gebruikt (Van Den Bos & Lutje Spelberg, 2010). Bij afname van deze test krijgen kinderen een blad te zien met hierop 50 items. Aan hen is de taak om deze zo snel mogelijk in volgorde te benoemen. De onderzoeker houdt de tijd en accuratesse bij. De benoemtaken beslaan kleuren, letters, cijfers, plaatjes en woorden met één lettergreep. Uit ruwe scores worden gestandaardiseerde T-scores berekend. In de verantwoording van de CB&WL wordt een beeld geschetst van de begrips- en criteriumvaliditeit van de test. Hieruit wordt geconcludeerd dat de test valide is. Hiernaast worden betrouwbaarheidscoëfficiënten genoemd, deze zitten tussen de .83 en .91 en zijn hiermee betrouwbaar (Van Den Bos & Lutje Spelberg, 2010). De COTAN beoordeelt deze test overwegend voldoende. De criteriumvaliditeit is onvoldoende beoordeeld (Egberink & De Leng, 2010).

De continue benoemsnelheid wordt als variabele berekend door de T-scores van cijfers en letters benoemen te middelen. Deze keuze is gemaakt in lijn met PDDb 1.0 (Blomert, 2006), dat deze aspecten gebruikt om de benoemsnelheid te operationaliseren. Wanneer de CB&WL in de verdere context van dit onderzoek wordt genoemd, wordt deze gemiddelde T-score bedoeld.

Fonemische vaardigheid. Voor het vaststellen van fonemisch bewustzijn is de Fonemische Analyse Test Herzien versie (FAT-R) (De Groot, Van den Bos & Van der Meulen, 2014) gebruikt. Deze test wordt digitaal afgenomen en bestaat uit twee onderdelen. Het eerste

onderdeel gaat over foneemweglating: *streek* zonder *r*. Het tweede deel gaat over foneemwisseling: *Simon Vinkenoog* wordt *Vimon Sinkenoog*. De onderzoeker houdt bij hoelang de leerling nodig heeft de bewerking uit te voeren en te verbaliseren. Uit ruwe scores worden gestandaardiseerde T-scores berekend, deze FAT-indexscore heeft een gemiddelde van 50 en standaardafwijking van 10. De FAT-R handleiding (De Groot et al., 2011) bespreekt onderzoek naar de inhouds-, begrips- en criteriumvaliditeit van de test. Deze validiteitskenmerken worden als voldoende bestempeld. Hiernaast stelt de handleiding een voldoende betrouwbaarheid vast (De Groot et al., 2011). De COTAN beoordeelt deze test ruim voldoende (Egberink & De Leng, 2016).

Analyses

De eerste onderzoeksvraag, over de prevalentieverschillen, wordt beantwoord door middel van een McNemar test (Lachenbruch, 2005). Deze wordt gebruikt om vast te stellen in hoeverre er een significant verschil in aantallen is per groep. Voor deze analyse kan wel de volledige dataset gebruikt worden, omdat de eis van onafhankelijke groepen niet geldt voor deze test, $n = 235$. De hypothese is dat PDDB 3.0 een hogere diagnoseprevalentie heeft dan PDDB 1.0, als dit het geval is en de test een significant resultaat laat zien wordt deze hypothese aangenomen ($\alpha = .05$).

De gemiddelde scores op de verschillende tests zijn per diagnosegroep berekend. Voor de beantwoording van de tweede onderzoeksvraag wordt enkel de data gebruikt van de groepen die wél een diagnose volgens het ene protocol hebben, maar géén diagnose volgens de andere, $n = 14$. Dit in verband met de assumptie van onafhankelijke groepen. De 12 leerlingen die bij beide protocollen voor een diagnose in aanmerking komen worden hierom niet meegenomen.

In het kader van relevante assumpties is de normaliteit van deze scores geanalyseerd door middel van een Shapiro-Wilk test. Indien deze scores normaal verdeeld waren zijn deze geanalyseerd middels een t-toets, zo niet, met een Mann-Whitney U-test. Op deze manier is getoetst of er significante verschillen zijn tussen deze groepen ($\alpha = .05$). De hypothese is dat de scores op de EMT en Klepel-R hoger zullen uitvallen voor PDDB 3.0 en dat de scores op de FAT-R en CB&WL hoger zullen uitvallen voor PDDB 1.0. Als dit significant het geval is wordt deze hypothese aangenomen.

Om te analyseren hoe de voorspellende waarde van de FAT-R en CB&WL voor de EMT en Klepel zich verhouden bij de twee protocollen worden per diagnosegroep multiële lineaire regressieanalyses uitgevoerd, waarbij de CB&WL en FAT-R als voorspellers zijn

gebruikt en de EMT en de Klepel als afhankelijke variabelen zijn gebruikt. Hierbij wordt zowel de verklaring van beide voorspellers samen berekend als de deelverklaringen.

Effectgroottes worden gegeven als R^2 , deeleffectgroottes worden gegeven als sr^2 . De waarden hiervan worden als volgt geïnterpreteerd: rond .01 begint een klein effect, rond .09 een medium effect, en rond .25 een groot effect (Cohen, 1988). De hypothese is dat de voorspellende waarde van fonemisch bewustzijn op leesvaardigheid hoger is voor PDDB 3.0 dan voor PDDB 1.0 en dat de voorspellende waarde van continu benoemen op leesvaardigheid hoger is voor PDDB 1.0 dan voor PDDB 3.0. Deze hypothese wordt aangenomen als de sr^2 van de FAT-R op de Klepel en EMT hoger is voor PDDB 3.0 en als de sr^2 van de CB&WL hoger is voor PDDB 1.0. Deze effectgroottes zijn significant bij een Cronbach's alpha van .05.

Resultaten

Hieronder zullen de resultaten van de analyses besproken worden. Eerst wordt ingegaan op de diagnoseprevalentie, naar de eerste onderzoeksvraag. Hierna wordt, naar de tweede onderzoeksvraag, gekeken naar de scores op de tests. Eerst naar de gemiddelde scores, hierna naar de voorspellende waarden van tests onderling.

Tabel 1 laat zien hoe deze groepen zich met elkaar verhouden. 209 leerlingen kregen geen diagnose, 12 leerlingen kregen beide diagnoses, 9 leerlingen kregen enkel een diagnose volgens PDDB 1.0, en 5 leerlingen kregen enkel een diagnose volgens PDDB 3.0. Van de 26 gediagnostiseerde leerlingen had 46.2% beide diagnoses, 34.6% enkel volgens PDDB 1.0, en 19.2% enkel volgens PDDB 3.0.

Tabel 1

Frequenties diagnoses naar versie en diagnostische status

	PDDB 3.0	
	Geen diagnose	Wel diagnose
PDDB 1.0		
Geen diagnose	209	5
Wel diagnose	9	12

Verwacht werd dat een lagere diagnoseprevalentie aanwezig zou zijn PDDB 3.0. Opgeteld resulteerde PDDB 1.0 in 21 diagnoses en PDDB 3.0 in 17 diagnoses. De uitgevoerde McNemar test liet geen significant verschil zien in het aantal kinderen dat volgens de verschillende diagnosecriteria een dyslexiediagnose kreeg. $N = 235$, $\chi^2(1) = 0,65$ $p = 0,424$. De

McNemar test geeft een niet-significante bevinding. Dit heeft te maken met de kleine steekproef. Hoewel het verschil in aantal tussen de criteria niet significant is, is het merkwaardig dat voor meer dan de helft van de participanten (53,2%) de diagnose verschillend is. Dit wijst er voorzichtig op dat de protocollen verschillen in welke leerlingen ze diagnostiseren.

De tweede onderzoeksvraag kijkt naar de verschillen in cognitief profiel van gediagnosticeerde leerlingen tussen de protocollen. Dit wordt gedaan door zowel resultaten op gekozen tests te vergelijken als de voorspellende waarde van deze tests onderling. In Tabel 2 staat de beschrijvende statistiek van de gemiddelde scores op de vier tests per diagnosegroep. Deze groepen bestaan uit participanten met een enkele diagnose uit de eerste of derde versie van het PDDB. De gemiddelde scores voor de EMT, CB&WL en FAT-R lijken redelijk gelijk te zijn. Wat opvalt is dat de gemiddelde scores op de Klepel-R verder uit elkaar liggen: 35.71 versus 29.31.

Middels een Shapiro-Wilk test is de normaliteit van de testscores berekend. Deze waren geen van allen normaal verdeeld. Hierom is geen t-test uitgevoerd, maar een Mann-Whitney U-test. Deze liet zien dat enkel de Klepel-R significant verschilde tussen de twee diagnosegroepen: $U = 0.50$, $Z = -2.94$, $p = .003$ (hogere score PDDB 1.0). Voor de EMT ($U = 21.00$, $p = .841$, hogere score PDDB 1.0), CB&WL ($U = 8.00$, $p = .052$, hogere score PDDB 3.0) en FAT-R ($U = 16.00$, $p = .385$, hogere score PDDB 3.0) werden geen significante verschillen gevonden. De score op de CB&WL is bijna significant. Tabel 2 laat zien dat de gemiddelde EMT- en Klepelscore bij de PDDB 1.0 hoger liggen dan bij PDDB 3.0. De gemiddelde CB&WL en FAT-R scores liggen juist bij PDDB 3.0 hoger dan bij PDDB 1.0. Dit is, hoewel voor de CB&WL slechts marginaal significant, in lijn met Hypothese 2b.

Tabel 2

Beschrijvende resultaten voor woordlezen, benoemsnelheid en pseudowoordlezen per PDDB versie

		PDDB 1.0	PDDB 3.0
EMT	Gem.	27.56	26.74
	SD	7.45	6.182
	N	9	5
Klepel	Gem.	35.71	29.31
	SD	10.845	6.096
	N	9	5
CB&WL	Gem.	33.76	35.41
	SD	8.564	7.847
	N	9	5
FAT	Gem.	28.86	29.12
	SD	9.324	8.824
	N	9	5

Verwacht werd dat de voorspellende waarde van de FAT-R hoger zou zijn voor PDDB 3.0 ten opzichte van PDDB 1.0 en dat de voorspellende waarde van de CB&WL hoger is voor PDDB 1.0 ten opzichte van PDDB 3.0. Geen van de regressiemodellen is significant, dit moet meegenomen worden in eventuele interpretatie. Tabel 3 laat zien dat de sr^2 van de FAT-R als voorspeller voor de EMT en Klepel voor PDDB 1.0 respectievelijk .122 (medium effect) en .065 (klein effect) zijn, voor PDDB 3.0 zijn dit .037 (klein effect) en .306 (groot effect). De sr^2 van de CB&WL als voorspeller voor de EMT en Klepel-R zijn voor PDDB 1.0 respectievelijk .174 (medium effect) en .362 (groot effect), voor PDDB 3.0 zijn dit .381 (groot effect) en .040 (klein effect). De voorspellende waardes van de FAT-R en CB&WL komen op gebied van de Klepel-R overeen met Hypothese 2a ($.065 < .306$ en $.362 > .040$). De voorspellende waardes van de FAT-R en CB&WL komen op gebied van de EMT niet overeen met Hypothese 2a ($.122 > .037$ en $.174 < .381$). Voor de Klepel-R is de voorspellende waarde van de FAT-R dus sterker voor PDDB 3.0 en die van de CB&WL zwakker. Voor de EMT is de voorspellende waarde van de FAT-R dus zwakker voor PDDB 3.0 en die van de CB&WL sterker.

Tabel 3*Voorspellende Waarden van de FAT-R en CB&WL op de EMT en Klepel*

Protocol	Test	R ²	F	p	β	p	sr ²	β	p	sr ²
					FAT- R	FAT- R	FAT- R	CB&WL	CB&WL	CB&WL
PDDDB1	EMT	.519	3.241	.111	.387	.263	.122	-.463	.191	.174
PDDDB1	Klepel	.362	1.701	.260	.281	.466	.065	.667	.115	.362
PDDDB3	EMT	.738	2.821	.262	.225	.648	.037	-.721	.230	.381
PDDDB3	Klepel	.316	0.462	.684	.646	.444	.306	.234	.765	.040

Discussie

In onderstaande discussie zal antwoord gegeven worden op beide onderzoeksvragen. Hierbij worden voorzichtige, tentatieve conclusies getrokken over de mogelijke prevalentieverschillen als gevolg van PDDDB 3.0 en over de voorspellende waarden van cognitieve tests voor de gediagnostiseerde populatie. Hierna wordt gekeken naar beperkingen van dit onderzoek. Ten slotte wordt de opbrengst besproken en worden aanbevelingen gedaan voor vervolgonderzoek.

De eerste onderzoeksvraag van dit onderzoek luidt: “In hoeverre verschilt de prevalentie van leerlingen die in aanmerking zouden komen voor een dyslexiediagnose volgens de eerste versie van het PDDDB ten opzichte van de derde versie?”

Verwacht werd dat, gezien de strengere leeseisen en gezien eerder onderzoek een lagere prevalentie bij PDDDB 3.0 liet zien (Tijms et al., 2021), ook in huidig onderzoek een lagere dyslexieprevalentie gevonden zou worden volgens PDDDB 3.0. Een belangrijke noot bij het eerdere onderzoek is echter dat deze kinderen al gediagnostiseerd zijn. Ten tijde van hun diagnose waren de oude criteria van kracht. Het is aannemelijk dat er leerlingen zijn die volgens de nieuwe criteria in aanmerking zouden komen voor een diagnose, die niet in dit onderzoek zijn meegenomen.

In totaal hebben 26 participanten een diagnose gekregen, 12 kregen bij beide protocollen een diagnose, 9 enkel bij de eerste en 5 enkel bij de derde. Meer dan de helft (14) van de gediagnostiseerde leerlingen krijgt dus niet dezelfde diagnose volgens de protocollen. Waar PDDDB 1.0 een verklaringsmodel gebruikt dat zich richt op cognitieve factoren (Snowling & Göbel, 2011), maakt PDDDB 3.0 gebruik van een verklaringsmodel dat overeenkomt met het *multiple deficit model* (Pennington, 2006). Het erkennen van de multifactoriële aard van

dyslexie heeft mogelijk geleid tot een lagere focus op afzonderlijke factoren (cognitief profiel, benoemsnelheid, fonemisch bewustzijn) en een hogere focus op uiteindelijke kenmerken (leesvaardigheid). Dit voorkomt dat de ene factor zwaarder weegt dan een andere. Deze verandering van focus heeft dus mogelijk een gevolg voor welke leerlingen een diagnose krijgen. PDDB 3.0 levert voor de huidige steekproef ogenschijnlijk een lagere diagnoseprevalentie dan PDDB 1.0. Deze conclusie heeft echter weinig kracht aangezien de McNemar test een niet-significant resultaat laat zien en de steekproefgrootte zeer klein is. Replicatieonderzoek zou met een grotere steekproef dit met meer statistische power kunnen controleren.

De tweede onderzoeksvraag luidt: “In hoeverre verschillen de voorspellende waarden van fonemisch bewustzijn en continue benoemsnelheid voor leesvaardigheid van pseudo- en echte woorden bij leerlingen die volgens PDDB 1.0 in aanmerking zouden komen voor een dyslexiediagnose met leerlingen die volgens PDDB 3.0 in aanmerking zouden komen voor een dyslexiediagnose?”. Verwacht werd dat de voorspellende waarde van fonemisch bewustzijn bij beide tests hoger zou zijn voor PDDB 3.0 dan voor PDDB 1.0 en dat de voorspellende waarde van continu benoemen op beide tests hoger zou zijn voor PDDB 1.0 dan voor PDDB 3.0.

Dit zou zo zijn wanneer de voorspellende waarde van de FAT-R op zowel de Klepel als de EMT hoger zou zijn voor PDDB 3.0 en dat de voorspellende waarde van de CB&WL op zowel de Klepel als de EMT hoger zou zijn voor PDDB 1.0.

De resultaten laten zien dat voor de Klepel-R de voorspellende waarde van de FAT-R sterker is bij PDDB 3.0 en die van de CB&WL zwakker. Voor de EMT is de voorspellende waarde van de FAT-R zwakker bij PDDB 3.0 en die van de CB&WL sterker. Dit komt overeen met de hypothese voor de Klepel-R, voor de EMT komt dit niet overeen met de hypothese.

Coltheart en collega's (2001) beschrijven twee routes voor woordlezen: de grafeem-foneemkoppelingsroute en de lexicale route. Aan deze eerste route is fonemisch bewustzijn onderliggend, aan de tweede continu benoemen. PDDB 3.0 heeft strengere eisen voor leesvaardigheid en classificeert dus zwakkere lezers (Tijms et al., 2021). Zwakke lezers maken meer gebruik van fonemisch bewustzijn (Ehri, 2005). Doordat de Klepel-R uit pseudowoorden bestaat, maken leerlingen noodzakelijk gebruik van fonemisch bewustzijn (ze kunnen niet herkennen/benoemen). De FAT-R test fonemisch bewustzijn en voorspelt de Klepel-R sterker voor PDDB 3.0 dan voor PDDB 1.0. Het is dus terug te zien dat deze leerlingen meer gebruikmaken van de grafeem-foneemkoppelingsroute, gezien de vaardigheid die daarbij hoort een sterkere voorspeller is.

PDDB 1.0 indiceert daarentegen relatief sterkere lezers. Deze kunnen beter woorden memoriseren, zij maken hiervan tijdens het lezen meer gebruik (Ehri & Saltmarsh, 1995). De EMT meet algemene woordleesvaardigheid, daarbij komt benoemsnelheid kijken. De CB&WL voorspelt de EMT meer voor PDDB 3.0 dan voor PDDB 1.0. Verwacht zou worden dat de CB&WL een sterkere voorspeller zou zijn voor de EMT bij de sterkere lezers van PDDB 1.0. Dit is niet het geval. Dit betekent dat de voorspellende waarde van deze test voor PDDB 1.0 lager is dan voor PDDB 3.0.

De bovenvermelde resultaten laten zien dat de Klepel op de FAT-R een sterkere voorspeller is voor leerlingen gediagnostiseerd volgens PDDB 3.0 ten opzichte van PDDB 1.0. Op de EMT is de CB&WL een sterkere voorspeller voor leerlingen gediagnostiseerd volgens PDDB 3.0 ten opzichte van PDDB 1.0. Dit wijst erop dat de voorspellers beter in lijn staan met de respectievelijke leestests bij PDDB 3.0. Dit zou betekenen dat in het derde protocol, ondanks het niet meenemen van het cognitieve profiel, een dyslexiediagnose meer voorspeld wordt door cognitieve factoren dan in het eerste protocol. Dit maakt het een beter diagnostisch raamwerk. Deze conclusie heeft weinig kracht aangezien de regressieanalyses niet-significante resultaten laten zien en de steekproef klein is, wel laten de (niet significante) effectgroottes die slaan op de FAT-R als voorspeller van de Klepel en op de CB&WL als voorspeller van de EMT grote effecten zien bij PDDB 3.0.

Ook werd verwacht dat scores op leesvaardigheid lager zouden uitvallen bij PDDB 3.0 en dat scores op fonemisch bewustzijn en continue benoemsnelheid lager zouden uitvallen bij PDDB 1.0. Hoewel dit enkel bij de Klepel en bijna bij de CB&WL significant was, komt dit overeen met Hypothese 2b. Dit wordt verklaard door de strengere eis van de derde versie van het protocol op leesvaardigheid en het ontbreken van eisen voor het cognitieve profiel die de eerste versie wel had (Blomert, 2006; Tijms et al., 2021).

Zoals hierboven meermaals vermeld is de kleine steekproef een grote methodologische beperking. Hoewel de oorspronkelijke dataset bijna 2000 respondenten had, was bij een groot deel van de leerlingen niet de volledige testbatterij afgenomen. De 235 leerlingen die overbleven vormden een prima steekproefgrootte, echter moet rekening gehouden worden met het feit dat slechts een klein deel van deze leerlingen een dyslexiediagnose kan krijgen.

Een andere beperking is dat het dyslexietyperend cognitief profiel, zoals beschreven in PDDB 1.0, niet volledig op dezelfde manier is vormgegeven in huidig onderzoek als in het protocol. In het PDDB worden de accuratesse en/of snelheid van fonologische verwerking, accuratesse en/of snelheid van grafeem-foneemassociatie, en snel serieel benoemen van cijfers en/of letters meegenomen. Dit zijn niet exact geteste variabelen in de tests. Om het dichtst bij

dit criterium te komen is gekozen voor de vaardigheid foneemweglating, de vaardigheid foneemwisseling en het snel serieel benoemen van cijfers en van letters. Deze andere uitwerking van het criterium heeft mogelijk gevolgen gehad voor de accuratesse van diagnosestelling en daarmee op de accuratesse van de gevonden prevalentie en gemiddelde scores op testcores. De reden dat dit profiel, met de gebruikte tests, niet volledig vast te stellen is, komt overeen met de reden waarom dit profiel niet meer wordt gebruikt in PDDB 3.0. Nieuwere inzichten ten aanzien van de multifactoriële aard van dyslexie laten zien dat een cognitief-typerend profiel niet goed vast te stellen is en dus ook niet effectief is voor diagnosestelling (Pennington, 2006; Peterson & Pennington, 2012; Tijms et al., 2021). Ook zijn snelheid en accuratesse niet goed van elkaar te scheiden aangezien deze niet losstaand zijn (Coltheart et al., 2001). De *Speed-Accuracy Tradeoff* stelt dat hogere snelheden vaak resulteren in lagere scores (Alfers et al., 2024). Bij het lezen is het voor de leerling van belang om snel te lezen, maar wel op een tempo dat niet in (te veel) fouten resulteert. Wanneer snelheid en accuratesse als verbonden aan elkaar worden gezien vermindert het effect van deze trade-off. Het PDDB 1.0 ziet deze bij het cognitieve-profielcriterium los van elkaar, het PDDB 3.0 ziet deze als verbonden aan elkaar.

In dezelfde trant van verschillen in diagnosecriteria is spelling niet meegenomen in dit onderzoek, terwijl het wel deel van de diagnosecriteria is.

Naast de criteria met betrekking op testcores neemt PDDB 1.0 een criterium mee dat PDDB 3.0 weglaat, dit is in hoeverre een leerling uitvalt op andere vakken (Blomert, 2006). Dit criterium is niet meegenomen in dit onderzoek. Het is aannemelijk dat bij invoering van dit criterium bepaalde leerlingen buiten de diagnosegroep zouden vallen. Dit zou een onbekend effect kunnen hebben op de reikwijdte van de resultaten van huidig onderzoek.

Gezien de kleine steekproefgrootte en een gebrek aan significante resultaten zijn de waarde en opbrengsten van huidig onderzoek beperkt. Echter, wel laat het zien dat de diagnoseprevalentie en het cognitief profiel van gediagnostiseerde leerlingen na invoer van PDDB 3.0 van het protocol mogelijk anders is. Hoewel insignificant, zijn de effectgroottes van de FAT-R op de Klepel-R en van de CB&WL op de EMT bij PDDB 3.0 sterke effecten, groter dan bij PDDB 1.0. Deze kennis is belangrijk om mee te nemen in toekomstig onderzoek en bij interpretatie en vergelijking van (toekomstige) prevalentiecijfers aangaande dyslexiediagnoses die dateren na de invoer van PDDB 3.0. Het wijst erop dat het nieuwe protocol resulteert in minder diagnoses en een cognitief profiel dat meer verklaard wordt door fonemisch bewustzijn en continue benoemsnelheid. In het bijzonder wordt pseudowoordlezen meer verklaard door fonemisch bewustzijn en het lezen van gewone woorden meer door continue benoemsnelheid.

Specifiekere conclusies zijn, in verband met de genoemde beperkingen, niet te trekken. Hiervoor zou grootschaliger onderzoek uitgevoerd moeten worden over dit onderwerp. Vervolgonderzoek zou gebruik moeten maken van een grote steekproef, waarbinnen voldoende data verzameld wordt om leerlingen volgens alle criteria van beide protocollen in te delen in substantiële diagnosegroepen. Alleen op deze manier is het mogelijk kennis te hebben van de effecten van de protocolwijziging (en in het verlengde hiervan de juistheid van deze wijziging).

Referenties

- Alfers, T., Gittler, G., Ulitzsch, E., & Pohl, S. (2024). Assessing the Speed–Accuracy tradeoff in psychological testing using experimental manipulations. *Educational and Psychological Measurement*. <https://doi.org/10.1177/00131644241271309>
- APA. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. American Psychiatric Publishing.
- Beech, J. R. (2005). Ehri's model of phases of learning to read: a brief critique. *Journal of Research in Reading*, 28(1), 50–58. <https://doi.org/10.1111/j.1467-9817.2005.00252.x>
- Blomert, L. (2006). *Protocol Dyslexie Diagnostiek en Behandeling*.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). LAWRENCE ERLBAUM ASSOCIATES.
<https://www.utstat.toronto.edu/brunner/oldclass/378f16/readings/CohenPower.pdf>
- 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(1), 204–256. <https://doi.org/10.1037/0033-295x.108.1.204>
- COTAN. (1981). *COTAN beoordeling Een-Minuut-Test, vorm A en B, EMT*.
<https://www.cotandocumentatie.nl/beoordelingen/b/13696/een-minuut-test--vorm-a-en-b/>
- COTAN. (2010). *COTAN beoordeling Continu Benoemen en Woorden Lezen, CB&WL*.
<https://www.cotandocumentatie.nl/beoordelingen/b/14599/continu-benoemen-en-woorden-lezen/>
- COTAN. (2015). *COTAN beoordeling Fonemische Analyse Test herziene versie, FAT-R*.
<https://www.cotandocumentatie.nl/beoordelingen/b/14812/fonemische-analyse-test-herziene-versie/>
- COTAN. (2021). *COTAN beoordeling Klepel-R*.
<https://www.cotandocumentatie.nl/beoordelingen/b/15848/klepel-r/>

- Daniel, J., Clucas, L., & Wang, H. (2024). Identifying students with dyslexia: exploration of current assessment methods. *Annals of Dyslexia*. <https://doi.org/10.1007/s11881-024-00313-y>
- De Groot, B. J. A., Van Den Bos, K. P., & Van Der Meulen, B. F. (2011). *Handleiding FAT-R*. Pearson.
- Ehri, L. C. (1991). Development of the Ability to Read Words. In *Handbook of Reading Research Volume II* (pp. 383–417). Longman Publishing Group.
- Ehri, L. C. (1999). Phases of development in learning to read words. In *Reading development and the teaching of reading: A psychological perspective* (pp. 79–108). Blackwell Science.
- Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, And Issues. In *Scientific Studies of Reading* (9th ed., Vol. 2). Taylor & Francis.
- Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing*, 7(3), 295–326.
<https://doi.org/10.1007/bf03162082>
- Heath, S. B. (1991). The Sense of Being Literate: Historical and Cross-Cultural Features. In *Handbook of Reading Research Volume II* (pp. 3–25). Longman Publishing Group.
- Lachenbruch, P. A. (2005). McNemar test. *Encyclopedia of Biostatistics*.
<https://doi.org/10.1002/0470011815.b2a10035>
- Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in Neurosciences*, 24(9), 508–511.
[https://doi.org/10.1016/s0166-2236\(00\)01896-8](https://doi.org/10.1016/s0166-2236(00)01896-8)
- OECD. (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*. PISA, OECD Publishing. <https://doi.org/10.1787/5f07c754-en>.

- OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. PISA, OECD Publishing. <https://doi.org/10.1787/53f23881-en>.
- Pennington, B. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, 101(2), 385–413. <https://doi.org/10.1016/j.cognition.2006.04.008>
- Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, 379(9830), 1997–2007. [https://doi.org/10.1016/s0140-6736\(12\)60198-6](https://doi.org/10.1016/s0140-6736(12)60198-6)
- Ramus, F. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841–865. <https://doi.org/10.1093/brain/awg076>
- Rintaningrum, R. (2009). Literacy: Its Importance and Changes in the Concept and Definition. *TEFLIN Journal*, 20(1).
- Roberts, T. A., Christo, C., & Shefelbine, J. A. (2011). Word Recognition. In *Handbook of Reading Research, Volume IV* (pp. 229–258). Taylor & Francis.
- Snowling, M. J., & Göbel, S. M. (2011). Reading Development and Dyslexia. In *Blackwell handbook of childhood cognitive development* (2nd ed., pp. 524–548). Wiley-Blackwell.
- Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the Garden-Variety poor reader: The Phonological-Core Variable-Difference Model. In *Journal of Learning Disabilities*.
http://www.keithstanovich.com/Site/Research_on_Reading_files/Stanovich_JLD_1988.pdf
- Stoodley, C. J., & Stein, J. F. (2012). Cerebellar function in developmental dyslexia. *The Cerebellum*, 12(2), 267–276. <https://doi.org/10.1007/s12311-012-0407-1>
- Tijms, J., De Bree, E. H., Bonte, M., De Jong, P. F., Loykens, E., & Reij, R. (2021). *Protocol Dyslexie Diagnostiek en Behandeling - versie 3.0*. Nederlands Kwaliteitsinstituut Dyslexie (NKD).

Van Den Bos, K. P., De Groot, B. J. A., & De Vries, J. R. (2019). *Klepel-R Handleiding*. Pearson.

Van Den Bos, K. P., & Lutje Spelberg, H. C. (2010). *CB&WL verantwoording*. Boom test uitgevers.

Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: a deficit in Skill-Automatization or a phonological deficit? *Scientific Studies of Reading*, 2(4), 321–340. https://doi.org/10.1207/s1532799xssr0204_2