

**Voorspellen van stagnaties in de vaardigheidsgroei op spelling op basis van
achtergrondkenmerken van leerlingen**

Naam: Jelinke (Jennie Klaaske) Nicolai (S3977617)

Begeleider: dr. N. Frans

Tweede beoordelaar: dr. J. Brouwer

Bachelorwerkstuk Academische Opleiding Leerkracht Basisonderwijs (AOLB)

Rijksuniversiteit Groningen

Faculteit Gedrag- en Maatschappijwetenschappen

Juni 2023

Abstract

Schools in the Netherlands are obliged to use a LOVS. Most schools use the Cito LOVS. Using Cito LOVS the spelling ability and the development of spelling can be monitored. More than 50% of the students show, at least once between kindergarten and third grade, a progress interruption in their spelling ability. We call those interruptions stagnation. They can be quickly identified using Cito LOVS. When a student shows a stagnation, this does not mean that students actually stagnate in terms of skill. A stagnation in Cito LOVS scores can also be due to measurement errors. Within this research, it was examined whether stagnations in spelling ability can be predicted based on background characteristics of the students. The following background characteristics have been included as variables: gender, age, birth month, migration background, grade retention, education level of the parents, number of tests and starting level. A correlational retrospective study ($N = 1402$) was conducted using logistic regression analysis to predict the risk of stagnation based on background characteristics of the involved students. Results show that the probability of stagnations can be predicted based on the starting level and the number of taken tests. This model predicts 76,9% of the stagnations correct. It is unlikely that stagnations stagnate in terms of skill. The measurement errors are therefore likely to have an impact. It is important for teachers to look further than just the Cito LOVS scores.

Inleiding

Alle scholen in Nederland moeten sinds schooljaar 2014/2015 verplicht een leerlingvolgsysteem gebruiken (Ministerie van OCW, 2022). In Nederland zijn verschillende leerlingvolgsystemen op de markt. Ruim 80% van de scholen gebruikt (een deel van) het Cito LOVS (Vlug 1997). Leerlingen maken één of twee keer per leerjaar een Cito toets voor verschillende vakken zoals rekenen-wiskunde, spelling en begrijpend lezen (Vlug, 1997; Glas et al., 2009). Deze toetsen hebben twee doelen: het bepalen van het niveau op dat moment en het bepalen van de progressie, de groei van de leerling. Het bepalen van de progressie is één van de voornaamste doelen van het Cito LOVS (Tomesen, 2016). We spreken van groei in vaardigheid wanneer een leerling een hogere vaardigheidsscore behaalt ten opzichte van een eerder meetmoment (Keuning et al., 2019). Omdat de toetsen zich bevinden op dezelfde gestandaardiseerde schaal kunnen *de vaardigheidsscores* hiervan met elkaar worden vergeleken. Het Cito LOVS maakt het mogelijk dat leerkrachten systematisch naar de resultaten van hun leerlingen kunnen kijken (Vlug, 1997; Reezigt, 2012) en daardoor hun onderwijs beter kunnen afstemmen op het (individuele) niveau van de leerlingen (Glas et al., 2009). Doordat het mogelijk is om de vordering van de leerlingen over een langere periode te volgen, kunnen ook stagnaties in groei worden gesignaleerd (Vlug, 1997; Reezigt, 2012). We spreken van een stagnatie wanneer er eenzelfde of lagere vaardigheidsscore te zien is dan bij een eerdere toetsafname.

Het leren correct spellen van woorden is een cruciale factor is in de algemene taal- en leesontwikkeling van kinderen. Ook helpt het goed kunnen spellen bij het schrijven van tekst en het kiezen van gepaste woorden (Hutcheon et al., 2012). Het is taalkundig van belang om woorden foutloos te kunnen schrijven. Spel- en taalfouten kunnen in formele schrijfsituaties zwaar aangerekend worden en worden in de schrijftaal in de maatschappij niet getolereerd (Bonset, 2010). Ook kan het slecht spellen invloed hebben later in de schoolcarrière

(Hutcheon et al., 2012). Het is dus belangrijk om vroegtijdig hulp te bieden aan kinderen die stagneren in vaardigheid op spelling.

Uit onderzoek van Frans et al. (2020) blijkt dat bij meer dan de helft van de leerlingen wel eens een stagnatie voorkomt. In de meeste gevallen kan een stagnatie gezien worden als een tijdelijke daling, die vaak binnen twee toetsmomenten weer is opgelost. Er is dan dus geen sprake van structureel beneden gemiddelde groei. Een stagnatie in toetsscore kan verschillende redenen hebben. Dit omdat er meerdere variabelen (kunnen) zijn die ook kunnen meespelen in de groei dan wel stagnatie van de leerling (Glas, et al., 2009). Een stagnatie hoeft dus niet altijd een daadwerkelijke stagnatie op vaardigheid zijn. Er kan ook sprake zijn van willekeurige meetfouten. Deze zijn statistisch te verklaren. Er mist echter nog kennis over wanneer een stagnatie willekeurig is en wanneer dit te verklaren is uit (één of meerdere) achtergrondkenmerken van de leerling. Deze informatie kan leerkrachten helpen bij vroegtijdige signalering van risicoleerlingen. Hierdoor kunnen beter verklaringen gevonden worden voor toetsresultaten en wordt voorkomen dat er hulp wordt ingezet bij leerlingen waarbij sprake is van meetfouten.

Jongens hebben meer moeite met taal en spelling en vallen hier ook eerder op uit dan meisjes (Nelson et al., 2006; Schjølberg et al., (2011)). Op basis hiervan worden stagnaties vaker verwacht bij jongens dan bij meisjes. In algemene zin scoren meisjes vaak hoger op spelling (en taal) dan jongens (Van de Gear et al., 2002). Ook kan een migratieachtergrond mogelijk effect hebben op het aantal stagnaties wat een leerling laat zien, deze leerlingen moeten namelijk een nieuwe taal leren. Het proces van het leren van een tweede taal is vergelijkbaar met het proces van het leren van de eerste taal, echter bij NT2 leerlingen is een lagere woordenschat en een lager syntactisch bewustzijn waarneembaar. Deze spelen een rol bij de spellingsontwikkeling (Harrison et al., 2016). Dit maakt dat deze leerlingen een groter risico hebben op een vertraagde ontwikkeling op taal en spelling en dus gezien kunnen

worden als risicoleerlingen (Schjølberg et al., 2011; Harrison et al., 2016). Verder wordt voor NT2 leerlingen het spellen van Nederlandse woorden steeds lastiger wanneer de woorden complexer worden (Beijerling-Oldenhof, 1996; Verhoeven, 2000). Leerlingen die Nederlands als moedertaal hebben, hebben minder moeite met spelling dan NT2 leerlingen (Verhoeven, 2000). Spelling is voor NT2 leerlingen in algemene zin lastiger (Kormos, 2020). Daarom worden stagnaties vaker verwacht bij leerlingen met een migratieachtergrond, met name in de hogere groepen van de basisschool omdat de complexiteit van de woorden die gespeld moeten worden dan toeneemt.

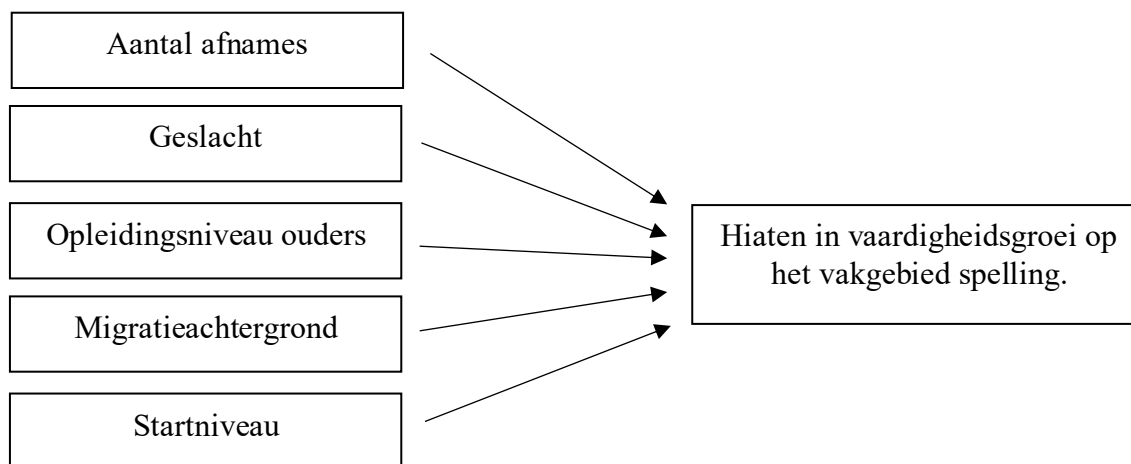
Het opleidingsniveau van ouders kan ook gezien worden als risicofactor voor stagnaties in de ontwikkeling van het vakgebied taal en spelling. Wanneer ouders zelf minder kennis hebben omdat ze minder hoog opgeleid zijn, kunnen zij dit in de thuissituatie ook niet overbrengen op hun kinderen (Nelson et al., 2006). Leerlingen van ouders met een hogere opleiding krijgen waarschijnlijk meer mee vanuit huis en zullen daardoor waarschijnlijk minder snel stagneren. Ook kan het startniveau van de leerling in groep 3 worden gezien als voorspeller van stagnaties. Bij leerlingen met een hoog startniveau worden meer stagnaties verwacht dan bij leerlingen met een laag startniveau, op basis van regression to the mean. Hierbij wordt gezien dat iemand die heel hoog scoort op een toets op moment één vaak vervolgens lager scoort op moment twee; dichterbij het gemiddelde (Smith & Smith, 2005). Regression to the mean hangt verder ook samen willekeurige meetfouten. Verder kan er sprake zijn van kanskapitalisatie bij meer afnames van de toets. Hoe vaker de toets is afgenomen, hoe waarschijnlijker dat één van deze resultaten een foutief resultaat is en er dus sprake is van een meetfout.

Binnen dit onderzoek staat de volgende vraagstelling centraal: In hoeverre kunnen voorspellingen worden gedaan over hiaten in vaardigheidsgroei op het vakgebied spelling tussen groep 3 en groep 5 op basis van achtergrondkenmerken van leerlingen? Er wordt een

verband verwacht tussen het geslacht van de leerling en mogelijke stagnaties op het vakgebied spelling. De verwachting is dat jongens vaker een stagnatie laten zien dan meisjes. Verder wordt ook verwacht dat leerlingen met een migratieachtergrond vaker een stagnatie laten zien in vergelijking met leerlingen zonder migratieachtergrond. Ten slotte wordt verwacht dat leerlingen met hoogopgeleide ouders minder stagnaties laten zien dan leerlingen met laagopgeleide ouders. Ten slotte is de verwachting dat hoe hoger het startniveau en hoe meer afnames, hoe groter de kans op het laten zien van een stagnatie. Deze variabelen zijn meegenomen in het conceptuele model (Figuur 1). Binnen het onderzoek zal ook nog exploratief gekeken worden naar een mogelijk verband tussen stagnaties op het vakgebied spelling en doubleurs, eenoudergezin, leeftijd en startniveau van de leerling.

Figuur 1

Conceptueel model



Methode

Design

Tijdens dit onderzoek is de onderzoeksvraag beantwoord aan de hand van een correlatieel retrospectief onderzoeksdesign. Er is gekeken naar de relaties tussen de verschillende variabelen. De gebruikte data voor dit onderzoek zijn verzameld voor het

onderzoek van Frans et al (2019), waarin onderzoek gedaan werd naar de stabiliteit van kleutertoetsen.

Populatie en Steekproef

De doelpopulatie binnen het onderzoek waren leerlingen in het basisonderwijs uit de groepen 3 tot en met 5 in Nederland. De oorspronkelijke steekproef bestond uit 1402 leerlingen van 59 scholen die allemaal gebruik maakten van minimaal één van de kleutertoetsen van het Cito LOVS. Alleen leerlingen die in 2014 startten in groep 6 en minimaal twee keer zijn getest op de Cito Spelling in de groepen 3 tot en met 5, zijn opgenomen in de steekproef. Om leereffecten bij herhaaldelijk toetsen te voorkomen zat er tussen de twee testmomenten een periode van minimaal twee maanden. Indien er minder tijd tussen lag is alleen de eerste afname behouden. Ook zijn leerlingen met een leerlinggebonden financiering uit de steekproef verwijderd omdat deze groep te klein was om conclusies over te kunnen trekken. De uiteindelijke steekproef bestond uit 1326 leerlingen.

Onderzoeksvariabelen en Instrumenten

Binnen dit onderzoek is gebruik gemaakt van de Cito LOVS toetsen voor spelling. Het doel van deze toetsen is om het spellingsniveau van de leerlingen vast te stellen en de ontwikkeling van de spellingvaardigheid meten. De toetsen worden tweemaal per schooljaar afgenomen door de leerkracht of intern begeleider in een groepssetting. Dit kan zowel digitaal als op papier zijn, met uitzondering van de toets M3. Deze is alleen op papier. De spellingvaardigheid van de leerlingen wordt getoetst in drie typen opgaven: woorddictee, zinsdictee en meerkeuzeopgaven. De toets bestaat in de groepen 3, 4 en 5 uit twee modules van elk 25 vragen. Cito heeft 48 LOVS-spellingcategorieën opgesteld. Slechts een deel hiervan komt voor tijdens de toets. Het aantal categorieën dat aangeboden moet worden neemt in de loop van de basisschoolperiode toe (De Wijs et al., 2010).

Het meetmodel van Cito maakt het mogelijk om scores van verschillende afnamemomenten en verschillende toetsen onderling te vergelijken. Hiervoor wordt de ruwe score omgezet tot een vaardigheidsscore op basis van een model uit de Item Response Theorie. De betrouwbaarheid van de toetsen zit tussen de .90 en .94 (De Wijs et al., 2010). Op basis van de criteria van de COTAN wordt deze gezien als goed voor toetsen voor belangrijke beslissingen op individueel niveau (Evers et al., 2010). De inhoudsvaliditeit van de toetsen is gegarandeerd door de manier waarop deze toetsen ontwikkeld zijn. De begripsvaliditeit wordt ook gezien als voldoende groot. De inhoud van de toets sluit aan bij de kerndoelen, leerlijnen en referentiekaders. Ook sluit de toets aan bij de meeste lesmethodes en heeft een gewenste verdeling van spellingscategorieën. Samenhang is zichtbaar tussen de verschillende toetsen voor spelling met een correlatie tussen de .49 en .96. Op basis hiervan kon de convergente validiteit ook als voldoende worden gezien (De Wijs et al., 2010).

De afhankelijke variabele binnen dit onderzoek was het aantal stagnaties op het vakgebied spelling. Wanneer een leerling één of meerdere stagnaties liet zien betekent dit dat deze leerling een daling of geen vooruitgang liet zien in zijn of haar vaardigheidsscore (Frans et al., 2020). Binnen dit onderzoek waren verschillende onafhankelijke variabelen: geslacht van de leerling (jongen = 0/meisje = 1), leeftijd (in jaren), geboortemaand (1-12), migratieachtergrond (ja = 1/nee = 0), doublures (ja = 1/nee = 0), opleiding ouders (zeer laag/laag/gemiddeld), toetsafnames (in aantal) en startniveau (score). De variabele opleiding ouders is gebaseerd op de oude gewichtenregeling. Op basis hiervan zijn de leerlingen ingedeeld in één van de drie categorieën: zeer laag opgeleid (maximaal één ouder diploma voortgezet onderwijs van maximaal VMBO-k, 1,2), laag opgeleid (beide ouders diploma voortgezet onderwijs van maximaal vmbo-k, 0,3) en gemiddeld opgeleid (alles boven een diploma VMBO-k, 0) (CBS, 2017).

Procedure

Via de e-mail zijn voor het oorspronkelijke onderzoek 1116 besturen van scholen uit heel Nederland benaderd. Bij de benadering zijn de deelnemers geïnformeerd over de doelstelling van het promotietraject waarvoor de data zijn verzameld. Ook is na twee weken een herinnering gestuurd naar de scholen. Scholen hebben zelf de keuze gemaakt om mee te doen aan het onderzoek. Uiteindelijk hebben 84 scholen positief gereageerd op een mogelijke deelname aan het onderzoek. De dataverzameling heeft plaatsgevonden tussen september 2014 en november 2015. Vervolgens hebben 25 van de 84 scholen de gegevens niet binnen de dataverzamelingstermijn ingeleverd. Dit resulteerde in een uiteindelijk totaal van 59 deelnemende scholen. De scholen hebben toestemming gegeven voor het gebruik van de data voor verwante onderzoeken. Voor dit onderzoek zijn de scores uit de leerlingvolgsystemen van de scholen gebruikt. De testgegevens werden met terugwerkende kracht vanaf groep 6 verzameld tot aan groep 1. Dit werd gedaan door de scholen zelf of in samenwerking met de onderzoekers. De overige informatie (geslacht, leeftijd, geboortemaand, opleiding ouders, migratieachtergrond, eenoudergezin, doubleur) is verkregen uit de administratie van de school. Door geen namen en exacte geboortedata te verzamelen maar de leeftijd af te ronden op de maand werd de anonimiteit van de respondenten gegarandeerd (Frans, 2019). Voor de dataverzameling op deze manier werd toestemming gegeven door de ethische commissie van PedOn.

Analyseplan

Alle analyses werden uitgevoerd in SPSS versie 28.0.1.0. Eerst is de steekproef beschreven door gebruik te maken van frequentietabellen van de categorische variabelen (geslacht, geboortemaand, opleiding ouders, migratieachtergrond, doubleures). Ook zijn histogrammen gemaakt om een uitspraak te kunnen doen over de verdeling van de steekproef. Daarna is de afhankelijke variabele (wel stagnatie = 1, geen stagnatie = 0) dichotoom gemaakt. Ook is een onderscheid gemaakt binnen de variabele geboortemaand op het gebied

van herfstkinderen. Kinderen geboren tussen 1 oktober en 1 januari zijn herfstkinderen (1), de overige maanden niet (0). Deze kinderen doubleren relatief vaak in groep 2 (Driessen et al., 2014) waardoor ze 2,5 jaar kleuteren. Dit levert deze kinderen vaak een voorsprong op (Resing et al., 2013). Verder zijn kinderen die een klas hebben overgeslagen (-1 keer doubleren) gehercodeerd naar geen doubleren (0), omdat deze groep te klein is om een uitspraak over te doen. Het variabele opleidingsniveau van ouders is gehercodeerd, waarbij zeer laag en laag opleidingsniveau is samengevoegd tot één categorie; (zeer) laag opgeleid (1), de categorie gemiddeld bleef ongewijzigd (0). Vervolgens zijn de bivariate relaties beschreven tussen de afhankelijke variabele en de categorische onafhankelijke variabelen. Hiervoor is een Chi Kwadraat test uitgevoerd met bijbehorende kruistabel. Ook is een *t*-toets uitgevoerd tussen de afhankelijke variabele en de continue onafhankelijke variabelen. Aan de hand van deze toetsen is een uitspraak gedaan over een mogelijke significante samenhang ($\alpha = .05$).

Ten slotte is een logistische regressieanalyse uitgevoerd om te onderzoeken of de stagnaties in score op spelling voorspeld kan worden door de onafhankelijke variabelen. Het model is gebouwd met forward selection, waarbij gestart is met de variabele startniveaus. Significante variabelen ($p < .05$) werden opgenomen in het model indien deze positieve invloed hadden op het percentage voorspelde variantie. De assumptie van lineariteit werd gecontroleerd door plots waarin de logit van de afhankelijke variabele werd afgezet tegen de continue onafhankelijke variabelen uit het model. De assumptie van multicollineariteit is gecontroleerd door de VIF ($> 5 - 10$) (Craney et al., 2002). De invloedrijke punten zijn gecontroleerd door gebruik te maken van de Cook's distance ($< 4/n$). Wanneer de waarde boven de 4 kwam werd deze gezien als outlier (Hoekstra, z.d.). Ten slotte is het belangrijk dat de resultaten onafhankelijk van elkaar verkregen zijn.

Resultaten

In totaal bestond de steekproef uit 1326 participanten. Deze voldeden allemaal aan de inclusiecriteria. In Tabel 1 is de verdeling van de steekproef op de verschillende categorische variabelen weergegeven. Het grootste deel van de steekproef zijn leerlingen met meer dan één ouder (88,6%) met een gemiddeld opleidingsniveau (88,2%), zonder migratieachtergrond (90,7%) die niet gedoubleerd zijn (92,0%) en één stagnatie lieten zien (74,4%).

Tabel 1

Beschrijving van de Steekproef voor de Categorische Variabelen (N = 1326)

Variabele	Categorieën	<i>n</i>	%
Doublure	Doublure	102	7,7%
	Geen doublure	1220	92,0
	Klas overgeslagen	3	0,2%
	Missing	1	0,1%
Eenouder	Eenoudergezin	149	11,2%
	Geen eenoudergezin	1175	88,6%
	Missing	2	0,2%
Geslacht	Jongen	653	49,2%
	Meisje	673	50,8%
Migratie	Migratieachtergrond	123	9,3%
	Geen migratieachtergrond	1203	90,7%
Opleidingsniveau ouders	(Zeer) laag	113	8,5%
	Gemiddeld	1169	88,2%
	Missing	44	3,3%

De steekproef lijkt verder redelijk gelijk te zijn verdeeld op de variabele geboortemaand. De maand die het meeste voorkomt is september. De maand die het minste

voorkomt is januari. Van de steekproef is 24,7% een herfstkind en van 0,1% is geen data met betrekking tot de geboortemaand. Van deze participant is bekend dat die niet stagneert. Uit de analyse van de variabele stagnaties blijkt dat 25,6% van de participanten geen stagnatie liet zien. De overige 74,4% liet minimaal één keer een stagnatie op het vakgebied spelling zien. Van het aantal leerlingen dat stagneert, stageert 62,6% één keer, 32,9% twee keer, 4,1% drie keer en 0,4% vier keer.

Tabel 2

Beschrijving van de Steekproef voor de Continue Variabelen (N = 1326)

Variabele	Missing (%)	Gem.	95%-bhi	SD	Mediaan	Min.	Max.
Afnames	0 (0,0%)	6,12	6,08; 6,16	0,68	6,00	2,00	10,00
Leeftijd	1 (0,1%)	9,80	9,77; 9,84	0,65	9,58	8,08	11,92
Startniveau	112 (8,4%)	109,56	109,17; 109,94	6,79	109,00	83,00	124,00

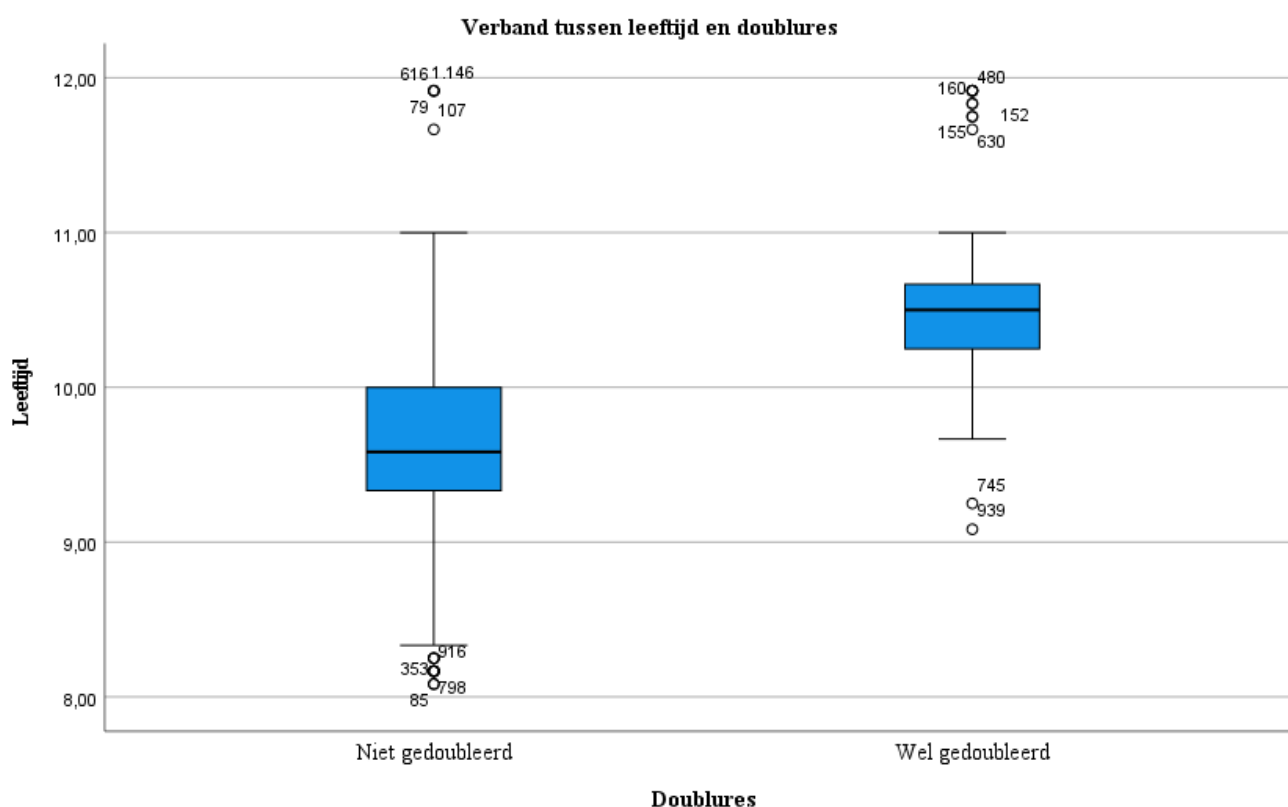
Tabel 2 laat een beschrijving van de continue variabelen zien. Hieruit blijkt dat de meeste leerlingen zes afnames hebben gehad. Dit komt overeen met de verwachtingen, twee keer per jaar een toets. Het startniveau heeft een vrij normale verdeling, tussen de 83 en 124. Het gemiddelde is 109,56 met een standaarddeviatie van 6,79. Bij 112 participanten (8,4%) is sprake van missing data. Veel van deze leerlingen komen van dezelfde scholen. Leerlingen waarbij het startniveau mist laten significant minder stagnaties zien en hebben significant minder afnames ($M = 3,9$) dan de groep die wel een startniveau heeft ($M = 6,1$). Verder hebben leerlingen waarbij sprake is van doublures ($M = 104,53$) een significant lager startniveau dan leerlingen zonder doublures ($M = 109,96$), $t(1214) = 7,458$, $p < .001$.

De gemiddelde leeftijd van de leerlingen is 9 jaar (verdeling tussen de 8 en 12 jaar). De variabele leeftijd is bimodaal verdeeld wanneer deze wordt weergegeven in een histogram met een piek bij ongeveer 9,4 en bij ongeveer 10,8 jaar. Deze verdeling heeft vier uitbijters aan

de bovenkant van de verdeling. Al deze leerlingen zijn geboren in de maand november. Het zijn dus allemaal herfstkinderen. Van één van deze leerlingen is bekend dat zij een migratieachtergrond heeft. Van haar is geen startscore bekend en zijn er maar twee vervolgmetingen. Waarschijnlijk is ze later ingestroomd in het Nederlandse onderwijs. Dit kan ook haar hogere leeftijd verklaren. De bimodale verdeling van leeftijd kan verklaard worden door de variabele doublures.

Figuur 2

Samenhang tussen Leeftijd en Doublures.



Figuur 2 geeft het verband tussen leeftijd en doublures weer. De leeftijd van de gedoubleerde leerlingen ligt significant hoger dan van de niet gedoubleerde leerlingen met $t(1324) = -12.5, p < .001$. De mediaan voor de niet gedoubleerde leerlingen is 9,5 jaar en voor de wel gedoubleerde leerlingen is dit 10,5.

Tabel 3*Chi-kwadraattoetsen tussen Stagnaties en de Categorische Variabelen.*

	Stagnatie <i>n</i> (%)			χ^2 -waarde	<i>p</i> -waarde
	Niet	Wel	Totaal		
Doublures (<i>N</i> = 1325)					
Wel	15 (4,4%)	87 (8,8%)	102 (7,7%)	6,788	,009
Niet	323 (95,6%)	900 (91,2%)	1223 (92,3%)		
Eenoudergezin (<i>N</i> = 1324)					
Wel	47 (13,9%)	102 (10,4%)	149 (11,3%)	3,109	,078
Niet	292 (86,1%)	883 (89,6%)	1175 (88,7%)		
Geslacht (<i>N</i> = 1326)					
Jongen	184 (54,3%)	469 (47,5%)	643 (49,2%)	4,613	,032
Meisje	155 (45,7%)	518 (52,5%)	673 (50,8%)		
Herfstkinderen (<i>N</i> = 1325)					
Wel	81 (24,0%)	246 (24,9%)	327 (24,7%)	0,125	,724
Niet	257 (76,0%)	741 (75,1%)	998 (75%)		
Migratieachtergrond (<i>N</i> = 1326)					
Wel	47 (13,9%)	76 (7,7%)	123 (9,3%)	11,393	<,001
Niet	292 (86,1%)	911 (92,3%)	1203 (90,7%)		
Opleidingsniveau ouders (<i>N</i> = 1282)					
(Zeer) laag	36 (11,1%)	77 (8,0%)	113 (8,8%)	2,773	,096
Gemiddeld	289 (88,9%)	880 (92,0%)	1169 (91,2%)		

Uit tabel 3 blijkt dat een significant verband tussen stagnaties en doublures. Leerlingen die gedoubleerd zijn stagneren vaker dan leerlingen die niet gedoubleerd zijn. Verder is het

verband tussen stagnaties en migratieachtergrond ook significant. Leerlingen zonder migratieachtergrond stagneren vaker dan leerlingen met migratieachtergrond. Ook het verband tussen stagnaties en geslacht is significant. Van de leerlingen die stagneren is 52,5% meisje tegenover 47,5% jongen. Het verband tussen eenoudergezin en stagnaties is net niet significant. Van de leerlingen die stagneren komt 10,4% uit een eenoudergezin en 89,6% niet.

Tabel 4

T-toetsen tussen Stagnaties en de Continue Variabelen.

	Stagnatie, gemiddelde (SD)		<i>df</i>	<i>t</i> -waarde	<i>p</i> -waarde
	Wel	Niet			
Afnames (<i>N</i> = 1326)	6,08 (0,79)	5,53 (1,27)	1324	-9,067	<,001
Leeftijd (<i>N</i> = 1325)	9,81 (0,65)	9,82 (0,68)	1323	0,361	0,718
Startniveau (<i>N</i> = 1214)	110,03 (7,07)	107,97 (5,47)	1212	-4,495	<,001

Uit tabel 4 blijkt een significant verband tussen stagnaties en het aantal afnames en tussen stagnaties en startniveau. Te zien is dat leerlingen die stagneren gemiddeld genomen meer afnames hebben gehad dan leerlingen die niet gestagneerd zijn; 6,08 afnames tegenover 5,53 afnames. Ook hangen stagnaties en startniveau significant samen. Het startniveau van leerlingen die stagneren is hoger dan dat van leerlingen die niet stagneren.

In tabel 5 is de modelbouw van de logistische regressie weergegeven. Uit de verschillende bivariate relaties komt uiteindelijk naar voren dat stagnaties een significant verband hebben met de variabelen doublure(s), geslacht, migratie, aantal afnames en startniveau. In tabel 5 is te zien dat gestart is met een model waarin alleen het startniveau is meegenomen. Binnen dit model was de voorspeller significant. Vervolgens is ook in model 2 het aantal afnames toegevoegd. Binnen dit model waren de beide voorspellers significant. Vervolgens is ook de variabele geslacht toegevoegd aan het model. Deze voorspeller was

echter niet significant. Om te ontdekken of het model ook nog op een andere manier vormgegeven kon worden, is vervolgens de variabele startniveau verwijderd. Dit had als gevolg dat de variabele geslacht hierdoor wel significant werd maar de voorspelde proportie werd minder groot. Hierna zijn nog verschillende andere modellen geprobeerd maar deze hadden geen significante voorspellers en/of een lager percentage voorspelde variantie. Er is uiteindelijk gekozen voor model 2 met als voorspellers de variabelen startniveau en afnames.

Tabel 5

Modelbouw van de Logistische Regressie.

	B	SE	<i>p</i>	Exp (B)	Nagelkerke R ²
Model 1					
Constante	-3,77	1,121	<,001	0,023	0,025
Startniveau	0,05	0,010	<,001	1,047	
Model 2					
Constante	-7,43	1,414	<,001	0,001	0,052
Startniveau	0,05	0,011	<,001	1,053	
Afnames	,50	0,113	<,001	1,643	
Model 3					
Constante	-7,33	1,415	<,001	0,001	0,055
Startniveau	0,05	0,011	<0,01	1,051	
Afnames	0,49	0,114	<0,01	1,632	
Geslacht	0,22	0,140	,109	1,251	
Model 4					
Constante	-1,66	0,691	,016	0,190	0,028
Afnames	0,57	0,071	<,001	1,567	
Geslacht	0,27	0,130	,041	1,324	

Aan de hand van dit model is een logistische regressievergelijking opgesteld waarmee de kans op stagneren aan de hand van het startniveau en het aantal afnames kan worden bepaald.

$$P(\text{stagneren}) = \frac{e^{-7,43 + 0,05 * \text{startniveau} + 0,50 * \text{aantal afnames}}}{1 + e^{-7,43 + 0,05 * \text{startniveau} + 0,50 * \text{aantal afnames}}}$$

De odds voor het startniveau binnen dit model is 1,06. De kans op stagneren bij een gemiddeld startniveau ($M = 109,56$) en een gemiddeld aantal afnames ($M = 6$) op basis van dit model is 74,04%. Een leerling met een gemiddeld startniveau en zeven afnames heeft een kans van 82,46% op stagneren. Een leerling met een gemiddeld aantal afnames en een startniveau van één standaarddeviatie ($SD = 6,79$) heeft een kans van 80,02% op stagneren. De kans op stagneren wordt vergroot wanneer het startniveau hoger is en wanneer er meer afnames zijn geweest.

Tabel 6

(On)correct voorspelde stagnaties

Geobserveerde stagnatie	Voorspelde stagnatie		% correct voorspeld
	Niet	Wel	
Niet	8	273	2,8 %
Wel	8	925	99,1 %
Totaal			76,9 %

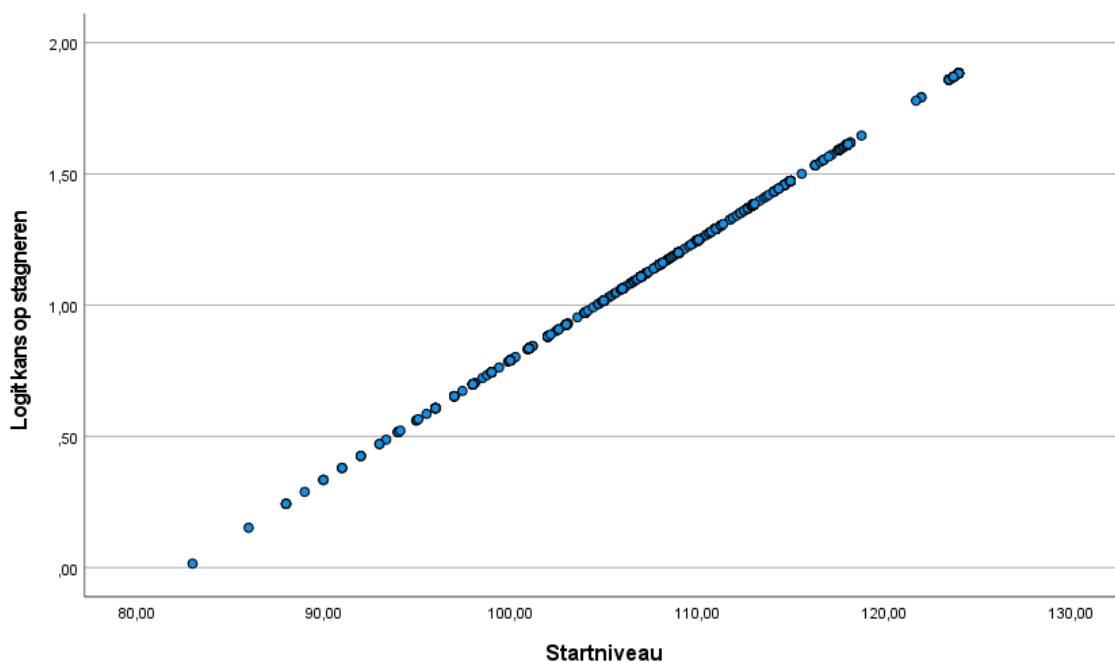
Uit tabel 6 blijkt dat 76,9% van het stagneren van leerlingen correct wordt voorspeld door het model. Voor leerlingen die wel stagneren is 99,1% juist voorspeld en voor leerlingen die niet stagneren wordt dit slechts in 2,8% van de gevallen juist voorspeld.

Binnen dit model is de VIF-waarde voor startniveau en afnames gelijk aan 1,000. Deze is kleiner dan de vooraf bepaalde waarde van 4. Er wordt dus voldaan aan de assumpties voor

multicollineariteit. De maximale waarde van Cook's Distance is 0,013, de outliers hebben dus waarschijnlijk weinig invloed gehad. Er wordt ook voldaan aan de assumptie van lineariteit, dit is te zien in figuur 3 en 4.

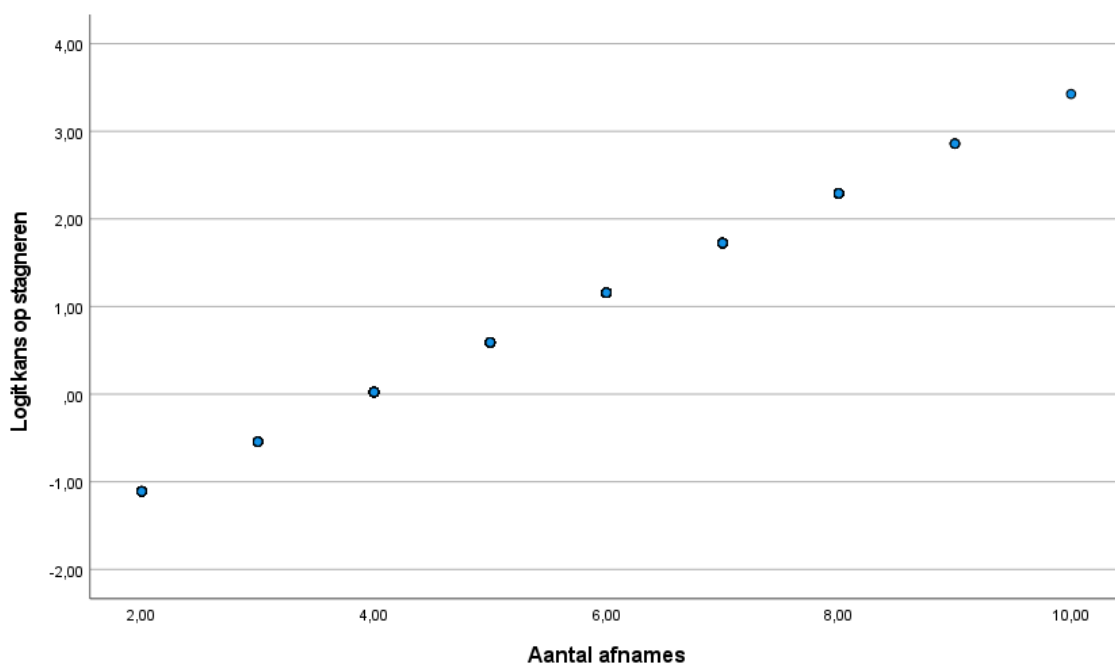
Figuur 3

Verband tussen Logit Kans op Stagneren en het Startniveau (N = 1214).



Figuur 4

Verband tussen Logit Kans op Stagneren en het Aantal Afnames (N = 1214).



Discussie

Het doel van het onderzoek was om een uitspraak te kunnen doen over in hoeverre hiaten voorkomen en voorspelt kunnen worden bij leerlingen op het vakgebied spelling op basis van achtergrondkenmerken van leerlingen. Uit dit onderzoek blijkt dat driekwart van de participanten minimaal één stagnatie laat zien. Voor leerlingen met een hoog startniveau en meer afnames is de kans op stagneren significant groter. Verder stagneren meisjes vaker dan jongens en spelen doublures een rol; leerlingen die gedoubleerd zijn stagneren vaker dan leerlingen die dit niet hebben gedaan. Leerlingen zonder migratieachtergrond stagneren verder vaker dan leerlingen met migratieachtergrond. Het uiteindelijke model bestaat uit de variabelen startniveau en aantal afnames. De overige achtergrondkenmerken hingen niet significant samen met stagnaties.

Op basis van eerder onderzoek waren hypothesen gesteld over de samenhang tussen achtergrondkenmerken van de leerling en stagnaties op spelling. Verwacht werd dat bij een hoger startniveau er een grotere kans was op stagneren. Deze hypothese is bevestigd. Dit zou, zoals verwacht, verklaard kunnen worden door regression to the mean. De leerlingen met een (extreem) hoog startniveau keren terug naar het gemiddelde; terug naar een evenwicht. Ook werd verwacht dat de kans op stagnaties werd vergroot bij meer afnames. Ook dit is bevestigd. Een mogelijke verklaring hiervoor is dat scholen vaker gaan toetsen bij leerlingen die stagneren, om zo de vorderingen beter in kaart te brengen. Een andere verklaring zou het significante verband tussen afnames en doublures kunnen zijn. Leerlingen die gedoubleerd zijn hebben significant meer afnames gehad en stagneren significant vaker dan leerlingen die niet gedoubleerd zijn. De overige variabelen zijn niet meegenomen in het uiteindelijke model omdat deze in het model niet meer significant samenhangen en/of de voorspelde variantie lager werd. Dit was onder andere het geval bij het geslacht. Er is een significante samenhang te zien tussen het geslacht en stagnaties; meisjes stagneren vaker dan jongens. Dit verband

kan echter verklaard worden door het significante verband tussen startniveau en geslacht. Het startniveau van meisjes ligt hoger dan van jongens en kan dus een verklaring zijn voor het vaker stagneren van meisjes. Ook werden er verbanden verwacht tussen stagnaties en de variabelen migratieachtergrond en eenoudergezin. Verwacht werd dat leerlingen met een migratieachtergrond vaker stagneren dan zonder migratieachtergrond. Dit onderzoek laat het tegenovergestelde zien. Een verklaring hiervoor kan het significante verband tussen startniveau en migratieachtergrond zijn. Leerlingen zonder migratieachtergrond scoren significant hoger dan leerlingen met migratieachtergrond, waardoor deze leerlingen eerder stagneren. Verder blijkt uit dit onderzoek dat leerlingen van ouders met een gemiddeld opleidingsniveau vaker stagneren, het tegenovergestelde werd verwacht. Ook dit zou verklaard kunnen worden doordat het startniveau van deze leerlingen significant hoger ligt dan dat van ouders met een (zeer) lage opleiding.

Binnen dit onderzoek is gebruik gemaakt van een ruime steekproef met voldoende participanten. Dit verhoogt de power en verlaagt de kans op een type-2 fout (Agresti et al., 2018). De steekproef was representatief voor het aantal afnames: zes keer aan begin groep 6. Het gemiddelde startniveau binnen de steekproef is 109,56. Dit is iets hoger dan het gemiddelde uit het normeringsonderzoek ($M = 106,83$). Dit verschilt minder dan één standaardafwijking (De Wijs et al., 2010) en wordt daarom niet als problematisch gezien. De steekproef is verder representatief op de variabelen dubblures (Van Vuuren & Van der Wiel, 2015) en geslacht. Uit gegevens van 2019 blijkt dat 21% van de bevolking laagopgeleid is, hierbij wordt wel aangegeven dat er een trend gaande is, waarin te zien is dat steeds minder mensen laagopgeleid zijn (Maslowski, 2020). De steekproef is voor nu niet helemaal representatief voor de populatie op deze variabele, maar, ervan uit gaande dat deze trend aanhoudt, in de loop van de tijd wel steeds meer. Binnen de populatie is 23% van de gezinnen een eenoudergezin (CBS, 2022), in de steekproef was dit maar 11,2%. Ook bevat de

steekproef minder leerlingen met een migratieachtergrond (9,3%) dan er daadwerkelijk in de populatie zijn (27,7%) (CBS, 2023). Door de significante samenhang tussen startniveau en de variabelen eenoudergezin en migratieachtergrond heeft deze niet-representativiteit waarschijnlijk geen invloed op de conclusie wanneer er gecorrigeerd wordt voor het startniveau. Een vervolgonderzoek zou gericht kunnen worden op een representatieve steekproef op alle variabelen, waardoor de externe validiteit verhoogd wordt en het onderzoek beter gegeneraliseerd kan worden.

Binnen het uiteindelijke model zijn slechts twee onafhankelijke variabelen meegenomen als voorspeller van de afhankelijke variabele die gezamenlijk slechts 5,2% van de proportie voorspelde variantie vormen. De rest van de variantie wordt dus niet verklaard door het model en wordt gezien als error. Bij een vervolgonderzoek kan gekeken worden of de error (deels) verklaard kan worden door andere variabelen. Hierdoor wordt het percentage verklaarde variantie vergroot. Verder komen de gegevens van de afhankelijke variabele van de resultaten van de leerlingen op de “oude” spellingtoets van Cito. Tegenwoordig maken scholen gebruik van de vernieuwde spelling 3.0 toetsen van Cito. Dit maakt dat er een risico zit in het generaliseren van dit onderzoek naar scholen die gebruik maken van de nieuwe toets. De betrouwbaarheid van deze toetsen zit tussen de .90 en .94, de toets is dus voldoende betrouwbaar voor het meten van de spellingvaardigheid en groei. Er is echter nog geen informatie over hoe betrouwbaar de toets is voor het vaststellen van stagnaties. Ook de validiteit van de Cito-toetsen is voldoende en een sterk punt van dit onderzoek.

Binnen het onderzoek wordt voldaan aan de assumpties van lineariteit en multicollineariteit. Ook zijn er geen punten die veel invloed hebben gehad op de resultaten. De assumptie van onafhankelijkheid van resultaten is wel geschonden. Een leerkracht en de school hebben namelijk veel invloed op de onderwijskwaliteit en daarmee op de scores van de leerlingen (Centraal Planbureau, 2011). Binnen het onderzoek is gebruik gemaakt van data

van 59 verschillende scholen. Er zijn verschillen te zien in de verhouding wel stageren versus niet stageren per school. Zo zijn er scholen waarbij 100% van de leerlingen niet stagneert maar ook scholen waarbij 100% wel stagneert. Er zou dus verder onderzoek gedaan kunnen worden naar de invloed van de school op hoeveel leerlingen stageren. Het gebruik van data van meerdere scholen versterkt wel de externe validiteit van het onderzoek.

Voor de onderwijspraktijk is het waardevol om te weten dat stagnaties deels verklaard zou kunnen worden door het startniveau en het aantal afnames. Dit betekent dus niet direct dat zij stagneren in vaardigheid op spelling. Het is belangrijk dat leerkrachten er zich van bewust zijn dat er ook andere verklaringen kunnen zijn voor de stagnatie, bv. op basis van achtergrondkenmerken van de leerling. Hierdoor zal een leerkracht niet alleen kijken naar de resultaten van de toetsen maar bijvoorbeeld ook naar de praktijk van de klas.

Referenties

- Agresti, A., & Finlay, B. (2018). *Statistical Methods for the Social Sciences, Global Edition* (5th ed.). Pearson.
- Beijerling-Oldenhof, E. (1996). Tussen klank en spelling; De leerbaarheid van de nieuwe regels voor de spelling van de tussen-e(n) getoetst. *Levende Talen Magazine*, 83(512), 407–410. Geraadpleegd van <https://lt-tijdschriften.nl/ojs/index.php/lm/article/view/1069>
- Bonset, H. (2010). Spelling in het onderwijs: hoe staat het ermee, en hoe kan het beter? *Levende Talen Tijdschrift*, 11 (3), 3–17. Geraadpleegd van <https://lt-tijdschriften.nl/ojs/index.php/lm/article/view/83>
- Centraal Bureau voor de Statistiek. (2017). Herziening gewichtenregeling primair onderwijs: Hoofdlijnenrapport. *Centraal Bureau Voor De Statistiek*. https://www.cbs.nl/-/media/_pdf/2017/06/gewichtenregeling-hoofdlijnenrapport.pdf

Centraal Bureau voor de Statistiek. (2022). *Alleenstaanden en gezinnen per gemeente*.

Centraal Bureau Voor De Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/regionaal/alleenstaanden>

Centraal Bureau voor de Statistiek. (2023). *Leerlingen (speciaal) basisonderwijs; migratieachtergrond 2011-2021*. Centraal Bureau Voor De Statistiek.

<https://www.cbs.nl/nl-nl/cijfers/detail/83295NED#shortTableDescription>

Centraal Planbureau. (2011). Evaluatie pilot investeren in kwaliteit leraren. In *CPB*.

<https://www.cpb.nl/sites/default/files/publicaties/download/cpb-achtergronddocument-evaluatie-pilot-investeren-kwaliteit-leraren.pdf>

Craney, T. A., & Surlles, J. G. (2002). Model-Dependent Variance Inflation Factor Cutoff Values. *Quality Engineering*, 14(3), 391–403. <https://doi-org.proxy-ub.rug.nl/10.1081/QEN-120001878>

De Wijs, A., Kamphuis, F., Kleintjes, F., & Tomesen, M. (2010). *Wetenschappelijke verantwoording: Spelling voor groep 3 tot en met groep 6*. Cito B.V.

Driessen, G., Mulder, L., Leest, B., & Verrijt, T. (2014). Zittenblijven in Nederland: Een probleem? *Tijdschrift voor Orthopedagogiek*, 53, 297-311.

Evers, A. W. M., Lucassen, W., Meijer, R. R., & Sijtsma, K. (2010). COTAN Beoordelingssysteem voor de kwaliteit van tests. *Nederlands Instituut Van Psychologen*, 978-90-6999-015–6. <https://psynip.nl/wp-content/uploads/2021/11/COTAN-Beoordelingssysteem-2010.pdf>

Frans, N. (2019). *A Captivating Snapshot of Standardized Testing in Early Childhood: On the stability and utility of the Cito preschool/kindergarten tests*. Ridderprint BV.

Frans, N., Post, W., Oenema-Mostert, C., & Minnaert, A. (2020). Signalering met de Cito kleutertoetsen: Ondergemiddeld is niet gelijk aan problematisch . *Tijdschrift voor Orthopedagogiek*, 59(2), 20-27. Geraadpleegd van

<https://research.rug.nl/en/publications/identification-with-the-cito-preschoolkindergarten-tests-below-av>

Glas, C. A., & Geerlings, H. (2009). Psychometric aspects of pupil monitoring systems.

Studies in Educational Evaluation, 35 (2–3), 83–88.

<https://doi.org/10.1016/j.stueduc.2009.10.006>

Hoekstra, R. (z.d.). *Statistische modellen 2: College 1*.

Harrison, G. L., Goegan, L. D., Jalbert, R., McManus, K., Sinclair, K. M., & Spurling, J.

(2016). Predictors of spelling and writing skills in first- and second-language learners.

Reading and Writing, 29(1), 69–89. <https://doi.org/10.1007/s11145-015-9580-1>

Hutcheon, G., Campbell, M., & Stewart, J. C. (2012). Spelling Instruction through

Etymology--A Method of Developing Spelling Lists for Older Students. *Australian Journal of Educational and Developmental Psychology*, 12, 60–70.

<http://files.eric.ed.gov/fulltext/EJ1002247.pdf>

Keuning, T., van Geel, M., Visscher, A., & Fox, J.-P. (2019). Assessing and validating effects

of a data-based decision-making intervention on student growth for mathematics and spelling. *Journal of Educational Measurement*, 56 (4), 757–792.

Kormos, J. (2020). Specific learning difficulties in second language learning and teaching.

Language Teaching, 53(2), 129–143. <https://doi.org/10.1017/s0261444819000442>

Maslowski, R. (2020). Onderwijs. In: *De sociale staat van Nederland: 2020*. Geraadpleegd

op 15 mei 2023 via <https://digitaal.scp.nl/ssn2020/onderwijs>.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2022, 12 juli). *Toetsen in het*

basisonderwijs. Onderwerp - Inspectie van het onderwijs. Geraadpleegd van

<https://www.onderwijsinspectie.nl/onderwerpen/toetsen-in-po>

- Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). Screening for speech and language delay in preschool children: systematic evidence review for the us preventive services task force. *Pediatrics*, *117*(2), 319. <https://doi.org/10.1542/peds.2005-1467>
- Reezigt, G. (2012). De kwaliteit van het basisonderwijs: ontwikkelingen in het laatste decennium. *Dijkstra, AB & Janssens, FJG, (red). Om de kwaliteit van het onderwijs. Kwaliteitsbepaling en kwaliteitsbevordering*, 25-42.
- Resing, W. & Smeets, J. (2013). De overgang van najaarsleerlingen naar groep 3 nader onderzocht. *Tijdschrift voor Orthopedagogiek*, *52*, 442-453
- Schjølberg, S., Eadie, P., Zachrisson, H. D., Oyen, A.-S., & Prior, M. (2011). Predicting language development at age 18 months: data from the norwegian mother and child cohort study. *Journal of Developmental and Behavioral Pediatrics : Jdbp*, *32*(5), 375–83. <https://doi.org/10.1097/DBP.0b013e31821bd1dd>
- Smith, G. A., & Smith, J. (2005). Regression to the Mean in Average Test Scores. *Educational Assessment*, *10*(4), 377–399. https://doi.org/10.1207/s15326977ea1004_4
- Tomesen, M., Wouda, J., & Horsels, L. (2016). *Wetenschappelijke verantwoording van de LVS-toetsen: Spelling 3.0 groep 5*. Cito B.V. <https://www.cito.nl/-/media/files/kennisbank/cito-bv/132-cito-lvs-spelling-3-0-gr-5-wet-verantwoording.pdf?la=nl-nl>
- Van de Gear, E., Van Damme, J., & De Munter, A. (2002). *Het verschil in schools presteren tussen jongens en meisjes: Eindrapport van het KULeuven-luik*. Katholieke Universiteit Leuven.
- Van Vuuren, D., & Van der Wiel, K. (2015). Zittenblijven in het primair en voortgezet onderwijs: Een inventarisatie van de voor- en nadelen. In *CPB* (No. 978-90-5833-674–3). Centraal Planbureau.

<https://www.cpb.nl/sites/default/files/publicaties/download/cpb-policy-brief-2015-01-zittenblijven-het-primair-en-voortgezet-onderwijs.pdf>

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, 4(4), 313–30.

Vlug, K. F. M. (1997). Because every pupil counts: the success of the pupil monitoring system in the netherlands. *Education and Information Technologies*, 2 (4), 287–306.

<https://doi.org/10.1023/A:1018629701040>