

De invloed van leerkrachtverwachtingen op het onderpresteren van jongens in lezen

Winnaar van de Mens & Maatschappij Bachelorscriptieprijs 2021

Mayke Nollet

Universiteit Utrecht

m.a.y.nollet@students.uu.nl

Kim Stienstra

Universiteit Utrecht

l.k.stienstra@uu.nl

Twitter: [@KimStienstra](https://twitter.com/KimStienstra)

Abstract

The effect of teacher expectations on boys' underperformance in reading

Prior research indicates a gender gap in reading performance among primary school students in which boys underperform compared to girls. Based on theories on gender differences in teacher expectations and how low teacher expectations lead to low student reading performance, we, firstly, examined to what extent teacher expectations can explain this gender gap in reading performance. Secondly, by including theories on gender stereotype threat and stigmatization, we studied the possible 'buffering' effect of high teacher expectations against boys' underperformance in reading. The 2016 Dutch PIRLS data were used to examine the mediating and moderating effect of teacher expectations on the effect of student gender on students' actual reading performance. The conducted multilevel analyses show that, although the gender gap in reading performance indeed exists, teacher expectations do not significantly explain or buffer it. Limitations of the current research and implications for future research are discussed.

Keywords: gender gap, reading performance, teacher expectations, Pygmalion effect, gender stereotypes, PIRLS

1 Inleiding

Verskillende studies laten zien dat onder Nederlandse basisschoolleerlingen meisjes gemiddeld betere leesprestaties hebben dan jongens (Ma, 2008; OECD, 2014; Onderwijsraad, 2020). Dit genderverschil in lezen is groter dan voor andere schoolvakken (OECD, 2014) en neemt toe: de afgelopen jaren zijn de leesprestaties van jongens achteruitgegaan, terwijl de prestaties voor meisjes minder hard zijn gedaald of gelijk zijn gebleven (Inspectie van het Onderwijs, 2017; Gubbels, Netten & Verhoeven, 2017). Dit is problematisch omdat leesvaardigheid ook invloed heeft op hoe goed leerlingen presteren in andere schoolvakken, waaronder rekenen (Martin & Mullis, 2013). Toenemende genderverschillen in lezen zouden daarom kunnen leiden tot negatieve consequenties op de langere termijn waarbij jongens minder bereiken in het onderwijs (Martin & Mullis, 2013). Dit is schadelijk voor jongens zelf, maar ook voor de samenleving. Veel jongens krijgen dan namelijk niet de kans om hun potentiële kennis en capaciteiten volledig te ontwikkelen en bereiken daardoor minder in en voor de samenleving (Onderwijsraad, 2020).

Er worden verschillende verklaringen gegeven voor de lagere leesprestaties van jongens, waaronder verschillen in hoe jongens en meisjes thuis en op school gesocialiseerd worden, motivatie, leesplezier en het zelfconcept van leerlingen (OECD, 2019). De verklaring waar wij in dit onderzoek op focussen is het verschil in verwachtingen die leerkrachten hebben van de leerling. De verwachtingen ten aanzien van leesprestaties en daaraan gekoppelde gedragingen van leerkrachten zijn mogelijk lager voor jongens dan voor meisjes (Gentrup & Rjosk, 2018; Onderwijsraad, 2020; Robinson & Lubienski, 2011; Tsiplakides & Keramida, 2010).

Leerkrachtverwachtingen kunnen op verschillende manieren een rol spelen. Ten eerste kunnen leerkrachtverwachtingen het effect van gender op leesprestaties deels verklaren (d.w.z., een *mediatie-effect*). Volgens het Pygmalion effect (Rosenthal & Jacobsen, 1968) zijn de communicatie en gedragingen van leerkrachten richting leerlingen gebaseerd op hun (soms foutieve) verwachtingen. Dit heeft vervolgens invloed op leerlingen, zodat de initiële verwachtingen van leerkrachten over deze leerlingen uitkomen als een zelfvervullende voorspelling (Gentrup & Rjosk, 2018; Jones & Dindia, 2004; Trouilloud, Sarrazin, Martinek & Guillet, 2002). Wanneer leerkrachten lagere verwachtingen hebben van jongens dan van meisjes, bijvoorbeeld door heersende genderstereotypen, dan zouden leerkrachtverwachtingen een verklaring kunnen zijn voor het genderverschil in lezen (zie Gentrup & Rjosk, 2018). Het meeste onderzoek laat zien dat hogere verwachtingen

van leerkrachten inderdaad leiden tot betere prestaties van leerlingen (voor een review, zie Wang, Rubie-Davies & Meissel, 2018), maar voor genderverschillen in leerkrachtverwachtingen is het empirische bewijs niet zo eenduidig (Hinnant, O'Brien & Ghazarian, 2009; McKown & Weinstein, 2008; Van den Bergh, Denessen, Hornstra, Voeten & Holland, 2010). Het mediatie-effect is weinig onderzocht, zeker als het gaat om lezen (Gentrup & Rjosk, 2018). De enkele studies die er zijn naar het mediërende effect van leerkrachtverwachtingen laten gemengde resultaten zien (zie bijv. De Boer, Bosker & Van der Werf, 2010; Gentrup & Rjosk, 2018). In dit onderzoek bestuderen we daarom de volgende onderzoeksvraag: *In hoeverre kunnen genderverschillen in de leesprestaties van leerlingen verklaard worden door leerkrachtverwachtingen?*

Daarnaast kunnen leerkrachtverwachtingen een sterker effect hebben op de leesprestaties van jongens versus meisjes (d.w.z., een *moderatie-effect*). De leesprestaties van jongens worden mogelijk sterker beïnvloed door (en zijn dus meer afhankelijk van) de verwachtingen van leerkrachten dan die van meisjes door de ervaren 'stereotypedreiging' (*stereotype threat*) en stigmatisering met betrekking tot de leesvaardigheden van jongens (Gentrup & Rjosk, 2018; Hinnant e.a., 2009; Jussim & Harber, 2005; Pansu, Régner, Colé, Nezlek & Huguet, 2016; Robinson & Lubinski, 2011). Empirische studies geven geen eenduidig bewijs en zijn vaker gefocust op prestaties in het algemeen of rekenprestaties dan op leesprestaties (De Boer e.a., 2010; Gentrup & Rjosk, 2018; McKown & Weinstein, 2002; Wang e.a., 2018). Sommige studies vinden in het geval van rekenen dat juist meisjes – in dit geval de gestigmatiseerde gendergroep – vatbaarder waren voor leerkrachtverwachtingen dan jongens (zie McKown & Weinstein, 2002; Wang e.a., 2018). De vraag is of jongens vatbaarder zijn voor leerkrachtverwachtingen dan meisjes als het gaat om lezen, gezien de leerkrachtverwachtingen dat jongens slechtere leesprestaties hebben. We onderzoeken daarom ook de volgende vraag: *In hoeverre hebben leerkrachtverwachtingen een sterker effect op de leesprestaties van jongens dan op de leesprestaties van meisjes?*

Het is belangrijk om zowel het mediërende als modererende effect te onderzoeken, omdat ze beiden een andere invloed van leerkrachtverwachtingen op genderverschillen in lezen veronderstellen. Wanneer er een significant mediatie-effect is, zou beleid dat inspeelt op gelijkere verwachtingen van leerkrachten ten aanzien van jongens en meisjes zorgen voor een kleiner genderverschil in leesprestaties (Nilsen & Bergem, 2020). In het geval van een moderatie-effect, varieert het genderverschil in lezen afhankelijk van hoe hoog de leerkrachtverwachtingen zijn (Nilsen & Bergem, 2020). Ook als leerkrachten even hoge verwachtingen van jongens en meisjes hebben, zijn

verwachtingen in het geval van moderatie nog steeds gerelateerd aan het genderverschil in lezen. Als leerkrachtverwachtingen een sterker positief effect hebben voor jongens, dan zouden hoge leerkrachtverwachtingen werken als een ‘buffer’ tegen de lagere leesprestaties van jongens. Een voordeel van het tegelijkertijd onderzoeken van beide effecten is dat dezelfde operationalisatie gebruikt zal worden waardoor de mogelijk mediërende en modererende effecten van leerkrachtverwachtingen nauwkeuriger met elkaar vergeleken kunnen worden.

Voor het beantwoorden van beide onderzoeksvragen maken we gebruik van de Nederlandse steekproef van de *Progress in International Reading Literacy Study* (PIRLS) data (PIRLS 2016 International Database, 2018). Deze dataset bevat de leesprestatiescores van basisschoolleerlingen met verschillende achtergrondkenmerken waaronder gender. Daarnaast hebben leerlingen vragen beantwoord over hun interacties met de leerkracht (bijvoorbeeld, ‘de leerkracht laat mij zien wat ik heb geleerd’). Verschillen in leerkrachtverwachtingen vertalen zich naar verschillen in leerkrachtgedragingen (Jussim & Harber, 2005; Meece, Glienke en Burg, 2006). In deze studie gebruiken we daarom de vragen aan de leerlingen over leerkrachtgedrag als maat voor leerlingpercepties van de verwachtingen die leerkrachten van hen hebben.

2 Theorie

2.1 Het mediatie-effect van leerkrachtverwachtingen

2.1.1 *Het effect gender op leerkrachtverwachtingen*

Vanuit historisch oogpunt heerst er de perceptie dat jongens van nature meer academische potentie hebben dan meisjes (Epstein, 1998) en blijkt het dat jongens in vergelijking met meisjes meer tijd krijgen in de les (Beaman, Wheldall en Kemp, 2006). Maar deze perceptie lijkt veranderd te zijn over de tijd in het nadeel van jongens. Dit valt bijvoorbeeld te zien aan het feit dat leerkrachtverwachtingen met betrekking tot algemene onderwijsprestaties en leesprestaties nadeliger lijken te werken voor jongens dan voor meisjes. Hier zijn verschillende verklaringen voor. Ten eerste worden de verwachtingen van leerkrachten beïnvloed door verschillen in hun perceptie van gedragingen van jongens en meisjes (Gentrup & Rjosk, 2018). Leerkrachten lijken hogere verwachtingen te hebben van leerlingen die zij waarnemen als onafhankelijk en die in hun ogen beschikken over positieve werkgewoonten op school (Åhslund & Boström, 2018). Eerder onderzoek laat zien dat, gemiddeld

genomen, leerkrachten dit vaker ervaren bij meisjes dan bij jongens. Leerkrachten ervaren dat meisjes meer moeite doen op school, onafhankelijker zijn, betere communicatievaardigheden hebben, beter georganiseerd zijn en minder storend zijn dan jongens. Jongens daarentegen worden door hun leerkrachten vaker gezien als afhankelijk, lui en ongemotiveerd (Åhslund & Boström, 2018; Downey & Yuan, 2005). Dit ervaren gedragsverschil in de klas tussen jongens en meisjes zou kunnen verklaren waarom leerkrachten lagere verwachtingen hebben over de algemene onderwijsprestaties van jongens in vergelijking tot meisjes.

Naast waargenomen genderverschillen in gedragingen kunnen ook veronderstelde genderverschillen op basis van stereotypen een rol spelen. Volgens de *expectancy-value theory* hebben mensen in de gehele samenleving, inclusief leerkrachten, onbewust verschillende stereotype verwachtingen waarbij sommige demografische groepen bevoordeeld worden en andere niet. In het geval van leesprestaties zouden meisjes wellicht bevoordeeld worden en hogere verwachtingen ontvangen dan jongens door de bestaande culturele opvatting dat lezen een activiteit is waar meisjes van nature goed in zijn en motivatie voor hebben en jongens minder (Ma, 2008; Martin & Mullis, 2013; Retelsdorf, Schwartz & Asbrock, 2015; Wolter, Braun & Hannover, 2015). Het bestaan van dit stereotype zou kunnen leiden tot lagere verwachtingen van leerkrachten over de leesvaardigheid van jongens.

Met inachtneming van de verwachtingen van leerkrachten over de gedragskenmerken en (lees)prestaties van jongens en meisjes, vonden Åhslund en Boström (2018) dat wanneer jongens en meisjes zich hetzelfde gedragen, de gedragingen anders worden geïnterpreteerd als gevolg van traditionele genderstereotypen die leerkrachten hanteren. Meisjes die veel praten tijdens de les worden bijvoorbeeld eerder gezien als sociaal, spraakzaam, en worden geprezen om hun hoge communicatieve vaardigheden. Jongens die veel praten tijdens de les worden echter ervaren als onoplettend, ongemotiveerd en ongeconcentreerd. Dit toont aan dat traditionele genderstereotypen die door leerkrachten worden gehanteerd, niet alleen rechtstreeks kunnen leiden tot lagere verwachtingen van leerkrachten over de leesvaardigheid van jongens. Deze genderstereotypen kunnen ook de perceptie van leerkrachten wat betreft verschillen in gedragskenmerken van jongens en meisjes vergroten. Deze verschillen in percepties van het gedrag van jongens en meisjes leiden vervolgens tot genderverschillen in verwachtingen die leerkrachten hebben. Hieruit kan de volgende hypothese worden afgeleid (zie ook Figuur 1):

H 1. Er bestaan lagere leerkrachtverwachtingen over de leesvaardigheid van jongens in vergelijking met die van meisjes.

2.1.2 *Het effect van leerkrachtverwachtingen op de leesprestaties van leerlingen*

Leerkrachtverwachtingen kunnen de prestaties van leerlingen beïnvloeden via *self-fulfilling prophecies* ('zelfvervullende voorspelling'): bepaalde verwachtingen worden werkelijkheid doordat mensen zich gedragen in overeenstemming met de verwachtingen. Toegepast op leerkrachtverwachtingen staat dit bekend als het 'Pygmalion-effect' (Rosenthal & Jacobsen, 1968). Leerkrachten communiceren hun verwachtingen over leerlingen via hun gedrag, waardoor het gedrag en de verwachtingen van leerlingen zodanig veranderen dat de oorspronkelijke verwachtingen van de leerkrachten over hen worden bevestigd (Gentrup & Rjosk, 2018; Jones & Dindia, 2004; Rosenthal & Jacobson, 1968; Trouilloud e.a., 2002; Tsiplakides & Keramida, 2010). Volgens dit mechanisme kan verwacht worden dat jongens gemiddeld lagere leesprestaties hebben dan meisjes omdat leerkrachten dit verwachten en jongens op hun beurt in overeenstemming met die lage verwachtingen zullen presteren. Bij *self-fulfilling prophecies* spelen drie processen een rol die de positieve associatie tussen leerkrachtverwachtingen en leesprestaties kunnen verklaren en die we verder zullen toelichten.

Ten eerste het verschil in leermogelijkheden die leerkrachten bieden. Leerlingen waarvan leerkrachten hoge verwachtingen hebben, lijken gemiddeld meer leermogelijkheden te krijgen en daardoor eerder hoger te presteren vergeleken met leerlingen waarvan leerkrachten lage verwachtingen hebben (De Boer e.a., 2010; Good, 1987; Jussim & Harber, 2005; McKown & Weinstein, 2002; Tsiplakides & Keramida, 2010; Wang e.a., 2018). Deze leermogelijkheden kunnen het aanleren van meer en moeilijkere leerstof zijn, leerlingen vaker de beurt te geven tijdens de les waardoor zij moeilijkere vragen kunnen beantwoorden om hun beheersing van de stof te laten zien, en leerlingen meer feedback over hun prestaties geven (Good, 1987; Jussim & Harber, 2005; Tsiplakides & Keramida, 2010). Verder laten Boerma en collega's (2016) zien dat leerlingen van wie de leerkracht lage verwachtingen heeft op het gebied van lezen, andere kansen krijgen om leesvaardigheid te verwerven dan andere leerlingen. Leerlingen worden bijvoorbeeld in aparte groepen geplaatst met andere leerlingen die laag presteren op het gebied van lezen.

Ten tweede kunnen deze verschillende manieren waarop leerkrachten meer en andersoortige kansen geven aan leerlingen ook indirect leiden tot verschillen in leesprestaties door de ontwikkeling van een negatiever zelfconcept bij leerlingen met lage leerkrachtverwachtingen (Boerma, Mol & Jolles, 2016). Het zelfconcept van leerlingen heeft betrekking op hun eigen verwachtingen, competentieovertuigingen en prestatie gerelateerd gedrag ten aanzien van hun leesvaardigheid (Fan, 2011; Martin & Mullis,

2013; Retelsdorf e.a., 2015). Leerkrachten hebben dagelijks sociale interacties met hun leerlingen en empirische bevindingen laten zien dat leerkrachten anders met leerlingen omgaan op basis van hun verwachtingen (Wang e.a., 2018). Afhankelijk van de hoogte van de verwachtingen van de leerkracht, kan het zelfconcept van leerlingen positief of negatief beïnvloed worden. Dit zelfconcept wordt waarschijnlijk positief beïnvloed wanneer leerkrachten laten zien dat zij hoge verwachtingen hebben van leerlingen door hen te motiveren en te stimuleren om veel moeite te doen voor en zo goed mogelijk te presteren bij het lezen (Fan, 2011; Good, 1987; Retelsdorf e.a., 2015). Aan de andere kant wordt het zelfconcept van leerlingen waarschijnlijk negatief beïnvloed wanneer leerkrachten laten zien dat ze lage verwachtingen hebben van leerlingen door een kille en minder uitdagende omgeving voor hen te creëren (Retelsdorf e.a., 2015).

Ten slotte suggereren attributie-theoretische modellen dat leerkrachtverwachtingen van invloed zijn op het attributieve denken van leerlingen over de redenen van hun successen en mislukkingen (Good, 1987). Leerlingen die hoge leerkrachtverwachtingen ervaren zullen waarschijnlijk geloven dat ze de mogelijkheid hebben om te slagen in lezen als ze een redelijke inspanning leveren en hun best doen. Dit wijst op een positieve ontwikkeling van het zelfconcept van de leerlingen. Vervolgens wordt van die leerlingen verwacht dat ze hun best doen om zo goed mogelijk te presteren en, als gevolg, daadwerkelijk hoog presteren bij het lezen. Maar wanneer een leerkracht weinig verwacht van de leesvaardigheid van bepaalde leerlingen, zullen deze leerlingen eerder denken dat zij niet de capaciteiten hebben om te slagen in lezen. Zij lopen vervolgens meer kans om in een 'faalsyndroom' (*failure syndrome*) of 'aangeleerd hulpeloosheidspatroon' (*learned helplessness pattern*) terecht te komen. Leerlingen zijn dan geneigd hun successen toe te schrijven aan geluk en hun mislukkingen aan een gebrek aan vaardigheden in plaats van onvoldoende inspanning of het gebruik van ineffectieve strategieën. Dit kan het zelfbeeld van leerlingen negatief beïnvloeden, wat ertoe kan leiden dat ze minder hun best doen en, vervolgens, daadwerkelijk minder goed presteren bij het lezen (Good, 1987). Toegepast op ons onderzoek, zouden we kunnen stellen dat jongens gemiddeld lager presteren in lezen dan meisjes omdat jongens door lage leerkrachtverwachtingen eerder denken dat ze lagere leesvaardigheden hebben en zich daarom minder inspannen voor lezen. Leerlingen lijken aan te voelen of leerkrachten geloven dat ze het vermogen hebben om te slagen of niet door middel van verbale en non-verbale interacties en gedragingen van leerkracht naar leerling, bijvoorbeeld door het minimaliseren van de eisen die aan leerlingen gesteld worden, het overreageren op kleine

successen, en door op mislukkingen te reageren met medelijden in plaats van te proberen de oorzaak van de mislukking te achterhalen (Boerma e.a., 2016; Good, 1987; Wang e.a., 2018).

Eerder onderzoek wijst uit dat waarschijnlijk alleen rigide en inflexibele verwachtingen bepaald gedrag zullen veroorzaken en als 'self-fulfilling prophecies' op leerling prestaties zullen werken (Jones & Dindia, 2004). Dit is bijvoorbeeld het geval wanneer leerkrachten hun verwachtingen baseren op demografische kenmerken van leerlingen zoals hun gender. Door genderstereotypen die leerkrachten kunnen hebben, hebben leerkrachten gemiddeld lagere verwachtingen van de leesvaardigheid van jongens dan van meisjes. Daardoor is de kans groter dat jongens negatieve zelfconcepten ontwikkelen over hun leesvaardigheden en minder gemotiveerd zijn om hoog te presteren in lezen dan meisjes (Retelsdorf e.a., 2015; Wolter e.a., 2015).

Uit de bovenstaande theorieën kunnen de volgende hypothesen afgeleid worden over het effect van leerkrachtverwachtingen op leerlingen hun leesprestaties (H₂) en over het totale mediatie-effect (H₃):

H₂. Hoe lager de leerkrachtverwachtingen over de leesvaardigheid van leerlingen, des te lager de daadwerkelijke leerprestaties van leerlingen zullen zijn.

H₃. De lagere leesprestaties van jongens versus meisjes wordt deels gemedieerd door leerkrachtverwachtingen.

2.2 Het moderatie-effect van leerkrachtverwachtingen

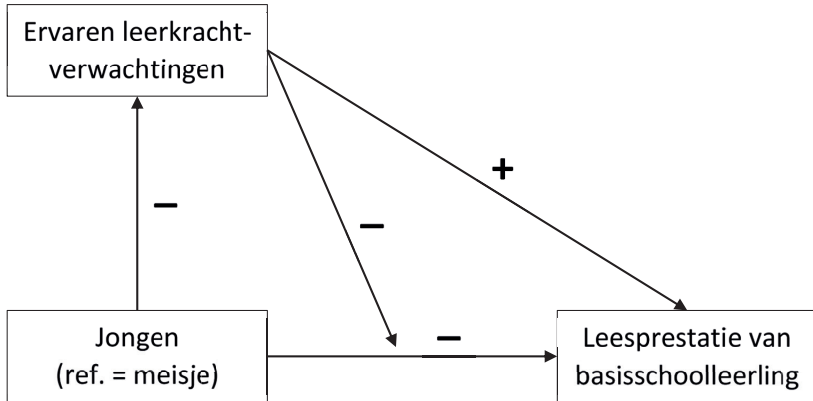
Leerkrachtverwachtingen kunnen het effect van het gender van leerlingen op hun leesprestaties modereren, waarbij het genderverschil in leesprestaties afhangt van hoe hoog de verwachtingen van de leerkracht voor jongens en meisjes zijn.

Onderzoek over stereotypedreiging (*stereotype threat*) geeft hiervoor een mogelijke verklaring (Gentrup & Rjosk, 2018; McKown & Weinstein, 2002). Cognitieve capaciteiten en motivatie zullen waarschijnlijk afnemen wanneer men vreest een negatief stereotype te bevestigen. Eerdere literatuur stelt dat de impact van lage verwachtingen groter is voor gestigmatiseerde dan voor niet-gestigmatiseerde leerlingen. Gestigmatiseerde leerlingen zullen bij het ervaren van lage leerkrachtverwachtingen namelijk eerder gaan denken dat ze niet de capaciteiten hebben en/of zullen eerder een afkeer krijgen tegen school omdat ze ervaren dat de lage leerkrachtverwachtingen gebaseerd zijn op een negatief stereotype over de gestigmatiseerde groep waartoe ze behoren. Vanwege de angst om het negatieve stereotype te bevestigen en/

of vanwege een ontwikkelende afkeer tegen school, zouden leerlingen die behoren tot een bepaalde gestigmatiseerde groep slechter presteren (McKown & Weinstein, 2002; Pansu e.a., 2016). Vanwege het genderstereotype dat jongens gemiddeld minder succesvol zijn in lezen dan meisjes, horen jongens tot de gestigmatiseerde groep als het gaat om leesprestaties. Lage leerkrachtverwachtingen zullen meer negatieve gevolgen hebben voor jongens dan meisjes, omdat stereotypedreiging bij meisjes niet speelt; zij hoeven immers niet te vrezen dat ze een negatief stereotype met betrekking tot lezen bevestigen. Wanneer leerkrachtverwachtingen verhogen, zullen de negatieve gevolgen van stereotypedreiging bij jongens minder spelen en kan het genderverschil in lezen kleiner worden. Hoge leerkrachtverwachtingen kunnen bijdragen aan een verminderde stereotypedreiging, omdat hogere leerkrachtverwachtingen met betrekking tot leesprestaties van jongens zouden kunnen duiden op dat het negatieve stereotype over leesprestaties van jongens minder prevalent is bij leerkrachten (Gentrup & Rjosk, 2018; Hinnant e.a., 2009; Jussim & Harber, 2005; Pansu e.a., 2016; Robinson & Lubienski, 2011).

Pansu en collega's (2016) onderzochten of en hoe verminderde stereotypedreiging de leesprestaties van jongens beïnvloedt. Bij het maken van een leesopdracht werden jongens en meisjes willekeurig toegewezen aan de 'bedreigingsconditie,' waarbij werd verteld dat de opdracht diende voor het evalueren van de leesvaardigheid, of de 'gereduceerde-dreigingsconditie,' waarbij werd verteld dat de opdracht diende als het testen van een spelletje. Hun bevindingen tonen aan dat jongens significant lagere leesscores hebben dan meisjes in de bedreigingsconditie, maar even goed presteren als meisjes in de gereduceerde-dreigingsconditie. De leesscores van meisjes verschillen niet significant tussen de dreigings- en de gereduceerde-dreigingsconditie. Verder evalueren jongens hun eigen leesprestaties gemiddeld positiever dan meisjes in de conditie met verminderde dreiging, terwijl ze dit minder positief evalueren wanneer er sprake is van stereotypedreiging. Deze bevindingen wijzen erop dat verminderde stereotypedreiging in sterkere mate leidt tot meer zelfvertrouwen en de ontwikkeling van een positief zelfconcept over leesvaardigheid bij jongens dan bij meisjes, wat op zijn beurt leidt tot hogere leesscores bij jongens.

Ook uit ander onderzoek blijkt dat hoge leerkrachtverwachtingen vooral de leesvaardigheid van gestigmatiseerde groepen verhogen (Hinnant e.a., 2009; Jussim & Harber, 2005). Hinnant en collega's (2009) vonden bijvoorbeeld dat jongens uit minderheidsgroepen de grootste winst in leesprestaties behalen wanneer hun capaciteiten door leerkrachten worden overschat. Verder wees onderzoek uit dat leerlingen die sociaal afhankelijk zijn van anderen vooral worden beïnvloed door hoge verwachtingen van leerkrachten



Figuur 1. Conceptueel model van de hypothesen

(Good & Nichols, 2001). Ervan uitgaande dat jongens gemiddeld afhankelijker zijn dan meisjes als het om lezen gaat – aangezien meisjes beschikken over hoge organisatievaardigheden, meer leergierigheid en meer discipline (Åhslund & Boström, 2018; Downey & Yuan, 2005) – kan deze bevinding betekenen dat hoge leerkrachtverwachtingen een sterker positief effect hebben op de leesprestaties van jongens in vergelijking met die van meisjes, op een zodanige manier dat het gendersverschil in lezen kleiner wordt.

Wanneer we de genoemde effecten van stereotypedreiging en stigmatisering samennemen, kan de volgende hypothese worden afgeleid:

H4. Hoge leerkrachtverwachtingen hebben een grotere impact op het verbeteren van de leesprestaties van jongens dan die van meisjes.

3 Data & Methoden

3.1 Data

Voor het beantwoorden van onze onderzoeksvragen maken we gebruik van de *Progress in International Reading Literacy Study* (PIRLS) dataset (Martin, Mullis & Hooper, 2017). PIRLS is een internationaal vergelijkend onderzoek naar de leesvaardigheid in het vierde leerjaar (groep 6 in de Nederlandse context) van het primair onderwijs. Tijdens de dataverzameling worden leestesten afgenomen en informatie verzameld over achtergrondkenmerken van leerlingen en hun ouders, leerkrachten en schoolleiders. Sinds 2001 worden de PIRLS testen en vragenlijsten elke vijf jaar afgenomen in vele landen over de wereld. In ons onderzoek gebruiken we de leestesten en

leerlingvragenlijsten van de Nederlandse PIRLS data uit 2016 (Gubbels e.a., 2017).

Voor de Nederlandse PIRLS 2016 was het responsepercentage van scholen in de hoofdstekproef 67 procent en het totale responsepercentage 88 procent. Uiteindelijk hebben 4206 leerlingen verdeeld over 226 klassen in 132 scholen deelgenomen aan de PIRLS 2016 test. De dataset voldoet daarmee aan de responsnorm van PIRLS en is daarom representatief voor de leerlingen uit groep 6 in Nederland. Voor meer informatie over de dataverzameling, kan de documentatie bij Gubbels en collega's (2017) en Martin en collega's (2017) geraadpleegd worden. Het oorspronkelijke aantal observaties in de Nederlandse PIRLS 2016 data was 4206. De observaties met missende waarden worden uitgesloten, wat leidt tot een analytische steekproef van 3995 leerlingen.

3.2 Operationalisatie

3.2.1 Leesprestatie

De PIRLS 2016 leestest bestaat uit het lezen van teksten die geselecteerd zijn op basis van de analyses van een proefmeting om de betrouwbaarheid en validiteit te waarborgen (Gubbels e.a., 2017). Voor het meten van de afhankelijke variabele, nemen we van deze leestest het gemiddelde van vijf plausibele waarden die de algehele leesvaardigheid van leerlingen meten. Deze plausibele waarden zijn berekend door het PIRLS-team door de toepassing van zogeheten *Item Response Theory* (IRT). Op deze manier kunnen betrouwbare schattingen van de leesprestatie van leerlingen op de toetsing als geheel gemaakt worden, ondanks dat elke leerling maar één leestestboekje heeft gekregen dat slechts een deel van de gehele toetsing bevat. Het werken met plausibele waarden waarborgt de nauwkeurigheid van de schattingen van de leesvaardigheid verdelingen, zodat de leesprestaties vergeleken kunnen worden tussen verschillende individuele leerlingen en subpopulaties (Martin e.a., 2017). Dit is belangrijk voor ons onderzoek, omdat de verschillen in leerprestaties tussen jongens en meisjes en tussen leerlingen met hoge versus lage waargenomen leerkrachtverwachtingen geanalyseerd worden.

De geconstrueerde schalen voor leesprestaties hebben scores tussen de 300 en 700. De internationale benchmarks geven aan of leerlingen laag (400), gemiddeld (475), hoog (550), of geavanceerd (625) presteren in lezen (Martin e.a., 2017). Als we kijken naar de beschrijvende statistieken in tabel 1, zien we dat de gemiddelde score 548,46 is met een standaarddeviatie van 55,70. De gemiddelde leesscore in Nederland is dus relatief hoog volgens de internationale benchmark van PIRLS (Martin e.a., 2017).

Tabel 1. Beschrijvende statistieken (N = 3.995)

	Min.	Max.	Gem.	SD
Leesprestaties	347,26	737,13	548,38	55,44
Leerkrachtverwachtingen	1,00	4,00	3,40	0,56
Jongen (ref. = meisje)	0,00	1,00	0,49	-
Leeftijd	8,50	12,08	10,04	0,45
Economische hulpbronnen ouders	-6,90	0,22	0,00	0,62
Culturele hulpbronnen ouders	1,00	5,00	2,92	1,08
Spreektaal thuis	0,00	1,00	0,15	-

Bron: PIRLS, 2016.

3.2.2 *Leerkrachtverwachtingen*

We meten leerkrachtverwachtingen aan de hand van leerkracht-leerling interacties. Een meer directe maat voor leerkrachtverwachtingen ervaren door leerlingen was niet aanwezig in deze data.¹ We denken echter niet dat dit problematisch is. Volgens het Pygmalion effect behandelen leerkrachten leerlingen anders afhankelijk van of ze hoge of lage verwachtingen van hen hebben (Jussim & Harber, 2005). Leerkrachten hebben bijvoorbeeld de neiging om in hun gedrag meer ondersteunend en warmer te zijn jegens leerlingen van wie ze hoge verwachtingen hebben (Meece, Glienke en Burg, 2006). Zo wordt er beargumenteerd dat leerlingen van wie de leerkracht hogere verwachtingen heeft, vaker de beurt krijgen en meer feedback krijgen (Good, 1987; Jussim & Harber, 2005). Dit komt terug in items zoals 'Mijn leerkracht laat mij zien wat ik heb geleerd' en 'Mijn leerkracht vertelt me hoe ik het beter kan doen als ik een fout maak'.

Daarnaast zijn het juist de gedragingen van leerkrachten die een belangrijke rol spelen in de percepties van leerlingen over de verwachtingen die leerkrachten van hen hebben. Leerkrachtverwachtingen hebben een invloed op prestaties van leerlingen deels doordat leerlingen zelf attributies maken over hun leesvaardigheid op basis van het aanvoelen van verwachtingen die leerkrachten over hen hebben (zie bijvoorbeeld Boerma e.a., 2016; Good, 1987; Wang e.a., 2018). Leerkracht-leerling interactie is een belangrijke manier waarop leerlingen deze verwachtingen aanvoelen (cf.

Tabel 2. Items voor de schaal 'leerkrachtverwachtingen'

	Min.	Max.	Gem.	SD	Cronbach's Alpha bij uitsluiten van item
Mijn leerkracht moedigt mij aan om te zeggen wat ik denk over wat ik heb gelezen	1,00	4,00	2,84	0,97	0,684
Mijn leerkracht laat mij zien wat ik heb geleerd	1,00	4,00	3,43	0,78	0,592
Mijn leerkracht doet verschillende dingen om ons te helpen leren	1,00	4,00	3,71	0,59	0,640
Mijn leerkracht vertelt me hoe ik het beter kan doen als ik een fout maak	1,00	4,00	3,63	0,66	0,634
Cronbach's Alpha					0,700

Bron: PIRLS, 2016.

Cooper, 1979). Al met al denken we dat de gebruikte items over leerkracht-leerling interacties een goede indicatie kunnen zijn voor de verwachtingen van de leerkracht. We meten deze leerkracht-leerling interacties aan de hand van vier items (zie tabel 2) met antwoordcategorieën van 1 (erg mee eens) tot 4 (erg mee oneens). We hebben deze items zo gecodeerd dat een hogere score overeenkomt met hogere leerkrachtverwachtingen. We hebben een schaal (Cronbach's alpha = 0,700) geconstrueerd op basis van de gemiddelde itemscores; respondenten zijn alleen meegenomen in de schaal als ze een geldig antwoord hadden op drie van de vier items.

3.2.3 Gender van de leerling

We gebruiken de dichotome variabele 'jongen' om aan te geven of de leerling een meisje (0) of jongen (1) is. Tabel 1 laat zien dat er evenveel meisjes als jongens in de steekproef zitten.

3.2.4 Controlevariabelen

In de analyses nemen we verschillende controlevariabelen mee die alternatieve verklaringen kunnen vormen voor de veronderstelde effecten.

Leeftijd van de leerling. Hoewel iedereen in groep 6 zit, is er wel wat variatie in leeftijd (zie tabel 1). Oudere kinderen in de klas zijn vaker zittenblijvers en presteren daardoor slechter in lezen dan de jongere, niet-zittenblijvende klasgenoten (Kotte, Lietz & Lopez, 2005; Kretschmann, Westphal & Vock, 2021). Verder laten resultaten zien dat hoe ouder kinderen worden, hoe meer ze hun competentiepercepties en zelfconcepten linken aan hun

daadwerkelijke onderwijsprestaties (Boerma e.a., 2016) en hoe beter ze de bestaande leerkrachtverwachtingen over hen kunnen afleiden (McKown & Weinstein, 2002; Retelsdorf e.a., 2015).

Sociaal-economische status (SES) van de ouders. Eerdere literatuur suggereert dat leerlingen uit lagere SES gezinnen vaker lagere leerkrachtverwachtingen ontvangen, lagere leesprestaties hebben en bevattelijker zijn voor leerkrachtverwachtingen (Fan, 2011; Good, 1987; Jussim & Harber, 2005; Tsiplakides & Keramida, 2010; Wang e.a., 2018; Wolter e.a., 2015). In dit onderzoek houden we rekening met de economische en culturele dimensies van ouders' SES.

Economische SES wordt gemeten middels de aanwezigheid van studiegerelateerde eigendommen in huis. Zulke hulpbronnen zijn belangrijke algemene indicatoren die de SES en het economisch kapitaal van de ouders kunnen meten (Hooper, Mullis & Martin, 2016; Sieben & Lechner, 2019). Voor zeven vragen (o.a. over het hebben van een eigen kamer, een bureau) konden leerlingen antwoorden met ja (1) of nee (2). Voor het hebben van een eigen kamer en/of internetverbinding en het hebben van een computer/tablet en/of internet connectie konden leerlingen antwoorden of ze geen (0), één (1) of beide (2) tot hun beschikking hadden. Alle items zijn gestandaardiseerd en zo gecodeerd dat een hogere score een hogere economische SES weergeeft. Naar aanleiding van de betrouwbaarheidsanalyse hebben we een schaal (Cronbach's Alpha = 0,694) geconstrueerd op basis van het gemiddelde van zes van de items, waarbij respondenten alleen meegenomen werden als ze een geldig antwoord hadden op minimaal drie items.

De culturele dimensie van ouders' SES wordt gemeten middels hoeveel boeken leerlingen thuis hebben. Het aantal boeken is een goede indicator voor cultureel kapitaal van de ouders (Sieben & Lechner, 2019). Leerlingen konden aangeven dat ze geen of een paar boeken hadden (1), genoeg om een plank mee te vullen (2), een boekenkast (3), twee boekenkasten (4), of drie of meer boekenkasten (5).

Spreektaal thuis. Leerlingen die thuis niet de Nederlandse taal spreken, hebben vaak lagere leesprestaties omdat ze meer moeite hebben met het herkennen en begrijpen van Nederlandse woorden tijdens het maken van leestoetsen (Cutting & Scarborough, 2006). Verder is het waarschijnlijk dat leerkrachten lagere verwachtingen hebben van de leesprestaties van leerlingen wanneer ze niet de Nederlandse taal thuis spreken (Gentrup & Rjosk, 2018). We meten spreektaal thuis met de vraag 'Hoe vaak spreek je thuis Nederlands'. We hebben een dummy-variabele aangemaakt die aangeeft of leerlingen (bijna) altijd (0) of soms of nooit (1) Nederlands thuis spreken.

3.3 Analysemethode

Allereerst kijken we naar de bivariate associaties tussen gender en leesprestaties en gender en leerkrachtverwachtingen. Voor het toetsen van de hypothesen maken we gebruik van multilevelanalyses. In de dataset zijn individuele leerlingen ‘genest’ binnen klassen en scholen, wat als gevolg heeft dat de observaties van leerlingen in dezelfde klas of school afhankelijk zijn van elkaar. Als hier geen rekening mee wordt gehouden in de statistische analyses, leidt dit tot kleinere standaardfouten en daarmee mogelijk valse ‘significante’ resultaten (Hox, 2010). Daarom gebruiken we een multilevel model met 3.995 leerlingen genest in 226 klassen in 132 scholen.

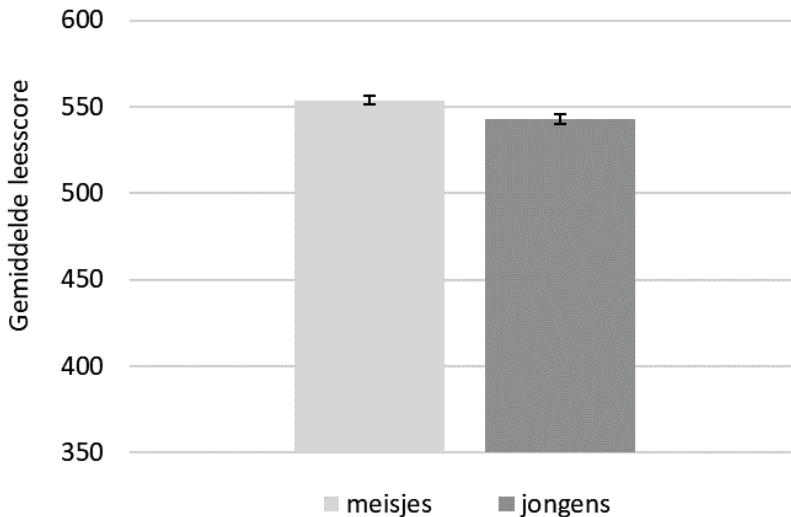
Het eerste model, ook wel ‘nulmodel’ genoemd, bevat een random intercept voor leesprestaties en geen predictoren. De variantie in leesprestaties kan verdeeld worden over drie niveaus: leerling-, klas-, en schoolniveau. Dit gebruiken we om de Intraklasse Correlatie Coëfficiënten (ICC's) voor het klas- en schoolniveau te berekenen (zie bijvoorbeeld Hox, 2010). In het tweede model voegen we gender toe als predictor en de controlevariabelen. Dit laat zien of er een genderverschil is in leesprestaties in het voordeel van meisjes, terwijl er rekening gehouden wordt met de geneste datastructuur en de controlevariabelen. In het derde model voegen we de predictor voor leerkrachtverwachtingen toe om te testen of dit een significant positief effect heeft op de leesprestaties van leerlingen en om te kijken of dit een deel van het gendereffect verklaart. Vervolgens analyseren we in het vierde model het moderatie-effect door het toevoegen van een interactieterm van leerkrachtverwachtingen en het zijn van een jongen (referentiecategorie = meisje). Om meer inzicht te krijgen in welke specifieke leerkracht-leerling interacties een rol spelen in het genderverschil in leesprestaties, testen we ook de mediatie- en moderatie-effecten voor de losse aspecten van ‘leerkrachtverwachtingen’.

4 Resultaten

4.1 Beschrijvende resultaten

Jongens hebben gemiddeld een lagere leesscore (542,818) dan meisjes (553,794), zoals te zien is in figuur 2. Dit genderverschil van 10,794 punten (0,099 standaard deviatie) op de PIRLS-leestoets is statistisch significant ($p < 0,001$), maar valt te interpreteren als een klein effect.

We verwachtten dat leerkrachtverwachtingen, middels leerkracht-leerling interacties, invloed hebben op dit genderverschil in lezen. Als we kijken naar de genderverschillen in leerkrachtverwachtingen, zien we dat jongens



Figuur 2. Leesprestaties naar gender

Noot: Gemiddelde leesprestaties van meisjes en jongens, inclusief 95 procent betrouwbaarheidsintervallen. Het verschil tussen jongens en meisjes is statistisch significant ($p < 0,001$).

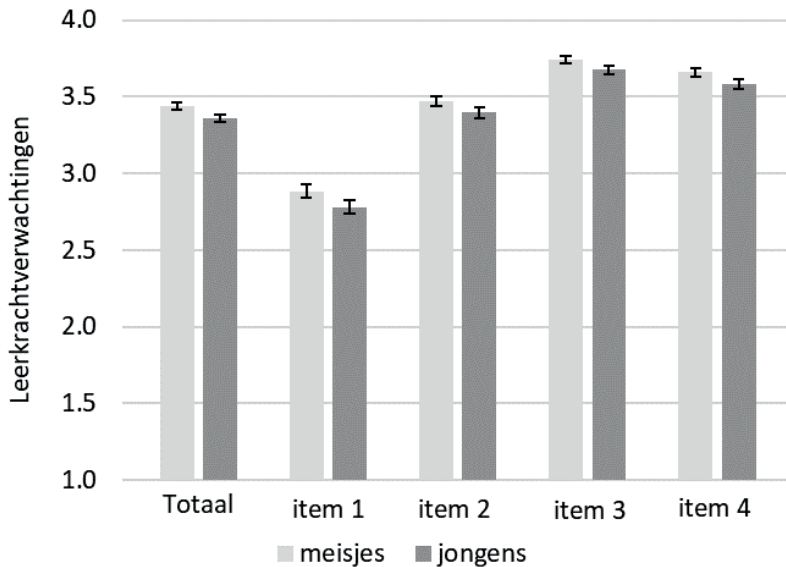
Bron: PIRLS 2016.

gemiddeld lager scoren op de schaal voor leerkrachtverwachtingen (3,358) dan meisjes (3,438) en dat dit verschil statistisch significant is ($p < 0,001$). Daarnaast blijkt dat dit geldt voor alle vier de aspecten van leerkrachtverwachtingen (zie figuur 3). Gemiddeld scoren jongens consequent lager op leerkrachtverwachtingen. De genderverschillen zijn klein te noemen en verschillen amper tussen de vier verschillende aspecten (zwakste associatie $\beta = -0,047$ voor item 2, sterkste associatie $\beta = -0,058$ voor item 4).

4.2 Toetsende resultaten

4.2.1 Nulmodel

De resultaten van de multilevelanalyses zijn te zien in tabel 3. Model 1 geeft de resultaten van het nulmodel weer, dat we gebruiken voor het berekenen van de ICCs. De ICC op klasniveau is 0,103, wat betekent dat 10,3 procent van de variantie in leesprestaties verklaard kan worden door verschillen tussen klassen. Daarnaast bestaat 5,9 procent van de totale variantie uit verschillen tussen scholen. Het overgrote deel van de variantie (83,8 procent) in leesprestatie ligt op het individuele niveau, dat wil zeggen, komt voort uit verschillen tussen individuele leerlingen.



Figuur 3. Leerkrachtverwachtingen naar gender

Noot: Gender verschillen voor de totale schaal (Totaal), leerkracht moedigt aan (1), leerkracht laat zien wat ik heb geleerd (2), leerkracht doet verschillende dingen om te helpen leren (3) en leerkracht vertelt hoe ik het beter kan doen (4). Inclusief 95 procent betrouwbaarheidsinterval. Alle verschillen tussen jongens en meisjes zijn statistisch significant op ten minste het vijf procent niveau.

Bron: PIRLS 2016.

4.2.2 Mediatie-effect

In Model 2 wordt gender toegevoegd als predictor, net als de controlevariabelen. Vergelijkbaar met de bivariate analyse, vinden we een genderverschil in leesprestaties waarbij jongens significant lager scoren dan meisjes ($b = -10,686$, $p < 0,001$). Het gestandaardiseerde effect (β) is $-0,096$ (niet gepresenteerd in tabel 3). Dit suggereert een klein genderverschil in leesprestaties waarbij de score van jongens gemiddeld 10,686 punten (0,096 standaarddeviatie) lager is dan die van meisjes op een schaal van 347,26 (minimum) tot 737,13 (maximum). Dit model beschrijft de data beter dan het nulmodel, wat blijkt uit de lagere AIC-waarde en de significante Chi^2 test ($Chi^2 = 475,563$, $df = 5$, $p < 0,001$).

In de beschrijvende analyses zagen we al dat jongens lagere leerkrachtverwachtingen rapporteren dan meisjes, in lijn met onze eerste hypothese. Deze bevinding blijft ook overeind in de multivariate analyse waarin we ook de controlevariabelen en de daadwerkelijke leesprestaties meenemen: jongens rapporteren significant lagere leerkrachtverwachtingen dan meisjes ($b = -0,077$, $\beta = -0,069$, $p < 0,001$) (zie online Appendix A²).

In Model 3 testen we of leerkrachtverwachtingen vervolgens ook invloed hebben op leesprestaties (H₂) en daarmee een deel van de associatie tussen gender en leesprestaties mediëren (H₃). Ten eerste valt er te zien dat het genderverschil waarbij jongens lager presteren dan meisjes van vergelijkbare grootte is als in Model 2 en statistisch significant (Model 3: $b = -10,583$, $p < 0,001$). Het negatieve directe effect van gender neemt slechts met 1,0 procent af wanneer leerkrachtverwachtingen worden toegevoegd. Dit lage percentage suggereert dat leerkrachtverwachtingen geen rol spelen in het verklaren van een deel van de genderverschillen in leesprestaties. Dit wordt verder onderstreept door de bevinding dat leerkrachtverwachtingen geen statistisch significant effect hebben op de leesprestaties van leerlingen. Daarnaast is het zo dat als we leerkrachtverwachtingen toevoegen als predictor aan het model, de modelfit niet verbetert aangezien de AIC toeneemt en de Chi^2 test niet significant is ($Chi^2 = 0,823$, $df = 1$, $p = 0,364$). Al met al kunnen we concluderen dat de resultaten onze hypothesen niet bevestigen. Hoewel er wel een genderverschil is in leerkrachtverwachtingen, speelt dit geen aantoonbare rol in het verklaren van het genderverschil in leesprestaties.

4.2.3 *Moderatie-effect*

Om te testen of hoger ervaren leerkrachtverwachtingen een grotere invloed hebben op de leesprestaties van jongens dan die van meisjes (H₄), voegen we een interactieterm toe tussen leerkrachtverwachtingen en 'jongen' in Model 4. Het toevoegen van de interactieterm zorgt voor een slechtere modelfit ten opzichte van model 3, zoals te zien is aan de hogere AIC-waarde en niet-significante Chi^2 test ($Chi^2 = 0,1$, $df = 1$, $p = 0,752$). De resultaten laten zien dat het effect van gender op leesprestaties van de leerlingen niet significant gemodereerd wordt door leerkrachtverwachtingen. We kunnen de vierde hypothese dus niet bevestigen.

4.2.4 *Specifieke aspecten van leerkrachtverwachtingen*

Wanneer we wat gedetailleerder naar leerkrachtverwachtingen kijken door te focussen op de specifieke aspecten hiervan in tabel 4, zien we dat er sprake is van gemengde resultaten.³ Het laten zien wat de leerling heeft geleerd ($b = 3,218$, $p = 0,009$) en het helpen leren ($b = 5,408$, $p = 0,001$) hangen beide positief samen met leesprestaties, zoals verwacht (zie Model 2). Echter vinden we ook een negatieve associatie: hoe meer een leerling waarneemt dat de leerkracht aanmoedigt om te zeggen wat de leerling denkt over wat de leerling heeft gelezen, hoe lager de leesprestaties ($b = -4,456$, $p < 0,001$). De mate waarin de leerkracht vertelt hoe het beter kan, is niet gerelateerd aan leesprestaties. Aangezien verschillende leerkracht-leerling interacties effecten in verschillende

Tabel 3. Multilevel analyses voor het testen van het mediërende en modererende effect van leerkrachtverwachtingen op leesprestaties

Parameters	Model 1		Model 2		Model 3		Model 4	
	Coëfficiënt	SE	Coëfficiënt	SE	Coëfficiënt	SE	Coëfficiënt	SE
Intercept	546,846***	(1,941)	696,560***	(18,360)	692,906***	(18,795)	694,640***	(19,581)
Jongen (ref. = meisje)			-10,686***	(1,561)	-10,583***	(1,565)	-13,631	(9,766)
Leerkrachtverwachtingen					1,320	(1,455)	0,853	(2,072)
Leerkrachtverwachtingen X jongen							0,896	(2,834)
Leeftijd leerling			-17,552***	(1,783)	-17,646***	(1,786)	-17,661***	(1,787)
Economische SES ouders			5,518***	(1,277)	5,462***	(1,278)	5,463***	(1,278)
Culturele SES ouders			11,737***	(0,761)	11,755***	(0,761)	11,766***	(0,761)
Taalvaardigheid			-11,417***	(2,272)	-11,351***	(2,273)	-11,370***	(2,273)
<i>Varianties</i>								
σ leerlingniveau	2607,369***	(60,076)	2339,315***	(53,942)	2338,844***	(53,932)	2338,910***	(53,935)
(proportie)	0,838		0,873		0,873		0,873	
σ klasniveau	319,986***	(67,124)	238,871***	(53,009)	238,629***	(52,984)	238,484***	(52,953)
(proportie)	0,103		0,089		0,089		0,089	
σ schoolniveau	183,713**	(68,689)	101,519*	(48,735)	101,693*	(48,751)	101,326*	(48,690)
(proportie)	0,059		0,038		0,039		0,039	
<i>Model fit</i>								
-2 Log Likelihood	43076,059		42600,496		42599,673		42599,573	
AIC	43084,059		42618,496		42619,673		42621,573	

* p < 0,05; ** p < 0,01; ***p < 0,001 (tweezijdige toets).

Noten: Standaardfouten (SE) staan tussen haakjes. AIC = Akaike's Information Criterion. N leerlingen = 3.995, N klassen = 226, N scholen = 132.

Bron: PIRLS, 2016.

richtingen hebben, is dit een mogelijke verklaring voor de afwezigheid van een effect wanneer we ze allemaal samennemen in een schaal.

Wanneer we de interactie-effecten toevoegen in Model 3, vinden we een significant interactie-effect voor 'leerkracht vertelt hoe het beter kan' ($b = 7,752, p = 0,006$). Het hoofdeffect is in dit model ook significant ($b = -4,454, p = 0,038$). Dit betekent dat voor meisjes er een negatief effect is: hoe meer een meisje waarneemt dat de leerkracht haar vertelt hoe ze het beter kan doen wanneer ze een fout maakt, hoe lager haar leesprestaties ($b = -4,454, p = 0,038$). Dit negatieve effect voor meisjes, is een positief effect voor jongens: hoe meer een jongen waarneemt dat de leerkracht hem vertelt hoe hij het beter kan doen, hoe hoger zijn leesprestaties ($b = -4,454 + 7,752 = 3,298$). Voor de andere items zijn geen statistisch significante interactie-effecten gevonden.

5 Conclusie en discussie

Het vaak gevonden genderverschil in leesprestaties bij basisschoolleerlingen, waarbij jongens lager presteren dan meisjes, heeft negatieve lange-termijn gevolgen voor jongens zelf en de samenleving als geheel, omdat lagere leesvaardigheden negatieve gevolgen hebben voor de algemene onderwijsprestaties van jongens. Het is daarom belangrijk om te onderzoeken welke factoren dit genderverschil zouden kunnen verklaren en welke dit zouden kunnen verkleinen. In deze studie hebben we onderzocht in hoeverre leerkrachtverwachtingen een mogelijke factor is. Onze hoofdconclusies zijn dat in het Nederlandse basisonderwijs leerkrachtverwachtingen (1) de relatie tussen gender en leesprestatie niet mediëren en (2) de invloed van gender op leesprestaties ook niet modereren.

Het huidige onderzoek suggereert dus dat het antwoord op de eerste onderzoeksvraag 'In hoeverre kunnen genderverschillen in de leesprestaties van leerlingen verklaard worden door leerkrachtverwachtingen?' is dat het genderverschil in leesprestaties van leerlingen inderdaad bestaat, maar dat leerkrachtverwachtingen het onderpresteren van jongens in lezen niet verklaren in het Nederlandse basisonderwijs. Het Pygmalion effect, dat aangeeft te verklaren hoe leerkrachtverwachtingen de leesprestaties van leerlingen beïnvloeden via *self-fulfilling prophecy*, wordt in onze studie niet ondersteund. Dit is in tegenstelling tot de meeste eerdere studies waarin wel een invloed van leerkrachtverwachtingen op leerling prestaties werd gevonden (zie bijvoorbeeld Robinson & Lubienski, 2011). De eerdere bevindingen voor een mediatie-effect waren wat meer gemengd. De bevindingen

Tabel 4. Multilevel analyses voor het testen van de effecten van de losse aspecten van leerkrachtverwachtingen op leesprestaties

Parameters	Model 1		Model 2		Model 3	
	Coëfficiënt	SE	Coëfficiënt	SE	Coëfficiënt	SE
Intercept	695,483***	(18,629)	674,612***	(19,255)	679,359***	(20,260)
Jongen	-11,225***	(1,586)	-11,081***	(1,583)	-19,760	(11,402)
<i>Leerkrachtverwachtingen</i>						
Leerkracht moedigt aan			-4,456***	(0,937)	-3,190*	(1,312)
... × jongen					-2,461	(1,850)
Leerkracht laat zien wat ik heb geleerd			3,218**	(1,229)	2,891	(1,780)
... × jongen					0,738	(2,448)
Leerkracht doet verschillende dingen om te helpen leren			5,408**	(1,624)	7,856**	(2,450)
... × jongen					-4,037	(3,262)
Leerkracht vertelt hoe het beter kan			-0,091	(1,419)	-4,454*	(2,142)
... × jongen					7,752**	(2,842)
<i>Variaties</i>						
σ leerlingniveau	2346,870***	(54,929)	2325,188***	(54,417)	2319,138***	(54,277)
(proportie)	0,874		0,989		0,989	
σ klasniveau	231,415***	(53,064)	221,212***	(51,864)	223,340***	(52,087)
(proportie)	0,086		0,009		0,009	
σ schoolniveau	105,741*	(49,138)	110,667*	(49,204)	111,151*	(49,403)
(proportie)	0,039		0,002		0,002	
<i>Model fit</i>						
-2 Log Likelihood	41451,435		41413,712		41405,194	
AIC	41479,435		41439,712		41439,194	

* p < 0,05; ** p < 0,01; *** p < 0,001 (tweezijdige toets).

Noten: Standaardfouten (SE) staan tussen haakjes. Gecontroleerd voor: leeftijd leerling, economische SES ouders, culturele SES ouders, spreektaal thuis.

AIC = Akaike's Information Criterion. N leerlingen = 3.886, N klassen = 226, N scholen = 132.

Bron: PIRLS, 2016.

van het huidige onderzoek komen overeen met die van Gentrup en Rjosk (2018), die ook geen mediërend effect van leerkrachtverwachtingen vonden terwijl er wel genderverschillen waren in leerkrachtverwachtingen.

De verscheidenheid in resultaten betreffende het mediatie-effect zou, ten eerste, kunnen liggen aan verschillen in context waarin de onderzoeken zijn gedaan. De meeste studies die een effect van leerkrachtverwachtingen op leesprestaties vonden, bestudeerden niet de (lees) prestaties van basisschoolleerlingen (zie bijvoorbeeld De Boer e.a., 2010; Robinson & Lubienski, 2011; Trouilloud e.a., 2002) terwijl de bevinding van een insignificant mediatie-effect van Gentrup en Rjosk (2018) wel gebaseerd is op basisschoolleerlingen. Vervolgonderzoek zou daarom moeten kijken in hoeverre leerkrachtverwachtingen belangrijker worden en een mediërend effect hebben naarmate leerlingen ouder worden en hun onderwijscarrière vervolgen op andere schoolniveaus (middelbare school, middelbaar beroepsonderwijs, hoger onderwijs). Ten tweede is er een grote variëteit in operationalisaties van leerkrachtverwachtingen, zoals middelbare schoolniveau adviezen van de leerkracht (De Boer e.a., 2010), leerkrachtgedragingen zoals in ons onderzoek (Robinson & Lubienski, 2011) en daadwerkelijke leerkrachtverwachtingen (Gentrup & Rjosk, 2018). Dit is een bekend probleem in dit onderzoeksveld (zie Wang e.a., 2018). Het systematisch in kaart brengen van de verschillende operationalisaties en de verkregen resultaten, evenals het komen tot een eenduidige conceptualisatie en operationalisatie van leerkrachtverwachtingen, is een belangrijke opgave voor vervolgonderzoek.

Hoewel we geen mediatie-effect vinden, vinden we wel een (klein) genderverschil waarbij jongens lagere leerkrachtverwachtingen rapporteren dan meisjes. Onze resultaten suggereren dat dit verschil verder geen gevolgen heeft voor hun lagere leesprestaties, aangezien we geen associatie vinden tussen leerkrachtverwachtingen en leesprestaties. Dit betekent niet noodzakelijk dat genderverschillen in verwachtingen helemaal geen gevolgen hebben voor de onderwijsprestaties. Jongens kunnen genderstereotypeverwachtingen waarnemen en/of ontvangen van andere belangrijke personen in hun leven zoals ouders of leeftijdsgenoten (Ma, 2008; Robinson & Lubienski, 2011). Verwachtingen van deze andere sociale groepen, maar ook van leerlingen zelf, kunnen ook dienen als alternatieve verklaringen voor het genderverschil in leesprestaties van leerlingen, onder andere via de invloed op het zelfconcept van leerlingen. De invloed van verwachtingen van ouders, leeftijdsgenoten en leerlingen zelf kan in toekomstig onderzoek op verschillende manieren worden onderzocht. Ten eerste kunnen de verwachtingen op dezelfde manier worden geanalyseerd als de verwachtingen

van de leerkrachten in ons onderzoek: door na te gaan of ze het effect van gender op de leesprestaties van de leerling mediëren en/of modereren. Naast het afzonderlijk onderzoeken van de invloeden van verwachtingen van anderen, zou ook onderzocht kunnen worden of de opeenstapeling van al deze maatschappelijke genderverwachtingen over jongens die slecht zijn in lezen het genderverschil in leesprestaties kan verklaren en/of verkleinen. Wellicht spelen verwachtingen vooral een rol wanneer deze universeel zijn.

Het antwoord op de tweede onderzoeksvraag 'In hoeverre hebben leerkrachtverwachtingen een sterker effect op de leesprestaties van jongens dan de leesprestaties van meisjes?' is dat hoge leerkrachtverwachtingen niet meer impact hebben op de leesprestaties van jongens in vergelijking met meisjes en dus geen 'buffer' vormen tegen de lagere leesprestaties van jongens. Het huidige onderzoek biedt geen ondersteuning voor stereotype dreiging en stigmatiseringseffecten waardoor de leesprestaties van jongens gevoeliger zijn voor de verwachtingen van de leerkracht. Onze resultaten zijn in lijn met die van De Boer en collega's (2010) en Gentrup en Rjosk (2018). Stereotypedreiging en stigmatisering lijken dus wel de rekenrepresentaties van meisjes gevoeliger te maken voor leerkrachtverwachtingen (zie Wang e.a., 2018; McKown & Weinstein, 2002), maar dit lijkt niet voor de leesprestaties van jongens te gelden.

Dit onderzoek heeft ook haar beperkingen. Ten eerste zou de operationalisatie van leerkrachtverwachtingen waargenomen door leerlingen kunnen verklaren waarom leerkrachtverwachtingen geen effect blijken te hebben in ons onderzoek. Wij hebben gedragingen van leerkrachten gebruikt om een schaal te creëren die de verwachtingen van de leerkracht indirect meten. We kunnen echter niet uitsluiten dat het veronderstelde positieve effect van leerkrachtverwachtingen gevonden zou zijn als we een meer directe maat zouden hebben van door leerlingen waargenomen leerkrachtverwachtingen. Daarnaast laten de analyses met de losse items zien dat verschillende leerkrachtgedragingen verschillende effecten hebben: het meer uitoefenen van sommige gedragingen zorgt voor gemiddeld betere leesprestaties terwijl het meer uitoefenen van andere gedragingen juist leidt tot slechtere leesprestaties. Sommige specifieke gedragingen wijzen wel op een mediatie- of moderatie-effect. Dit laat zien dat de rol van leerkrachtverwachtingen complexer en minder eenduidig is dan gedacht. Om beter te kunnen begrijpen hoe leerkrachtverwachtingen via leerkrachtgedragingen (genderverschillen in) leesprestaties beïnvloeden, zou dit gehele proces onderzocht moeten worden in vervolgonderzoek. Daarnaast zou toekomstig onderzoek kunnen kijken naar de verschillende impact van de perceptie van leerkrachtverwachtingen door leerkrachten versus leerlingen. Leerlingen

zouden directer gevraagd kunnen worden wat zij denken dat leerkrachten van hen verwachten, bijvoorbeeld met stellingen als: 'Mijn leraar verwacht veel van mijn leesprestaties' of 'Mijn leraar gelooft dat ik hoog kan scoren op mijn leestoetsen'. Wat betreft de perceptie van de leerkracht, zouden leerkrachten hun verwachtingen over elke leerling afzonderlijk kunnen identificeren.

Ten tweede weten we niet in hoeverre onze bevindingen robuust zijn over de tijd. Er valt te zien dat het verschil in leesprestaties tussen jongens en meisjes klein is in onze studie gebaseerd op data uit 2016. Het genderverschil was mogelijk groter geweest wanneer we recentere data hadden kunnen gebruiken, aangezien het genderverschil in leesprestaties groeit (Gubbels & Netten, 2017; Inspectie van het Onderwijs, 2017). Vervolgonderzoek zou daarom ons onderzoek kunnen uitvoeren met recentere data en kijken naar de verschillen in resultaten om zeker te weten dat leerkrachtverwachtingen geen invloed hebben op het genderverschil in leesprestaties in recentere jaren.

Ten derde hebben we in dit onderzoek geen gebruik gemaakt van longitudinale data, waardoor we de precieze processen niet in kaart kunnen brengen en geen causale uitspraken kunnen doen. Het doen van longitudinaal onderzoek met betrekking tot de associatie tussen verwachtingen en prestaties is relevant omdat ter discussie gesteld kan worden of leerkracht- en andere verwachtingen leiden tot bepaalde leerkrachtgedragingen en daarmee bepaalde leesprestaties van leerlingen, zoals wij in deze studie veronderstellen, of dat bepaalde leesprestaties van leerlingen leiden tot bepaalde verwachtingen (zie bijvoorbeeld Hughes, Kwok & Loyd, 2008; Kriegbaum, Steinmayr & Spinath, 2019; Retelsdorf, Köller & Möller, 2014). Hierbij hoeft het ene causale verband het andere causale verband niet uit te sluiten en kunnen deze processen wederkerig zijn en elkaar versterken (lage verwachtingen leiden tot lage prestaties die vervolgens leiden tot lage verwachtingen, etc.). Om deze causaliteitsproblemen te voorkomen en erachter te komen op welke manier 'self-fulfilling prophecy' processen plaatsvinden is longitudinaal onderzoek nodig.

De leesprestaties in Nederland nemen af, vooral voor jongens (Inspectie van het Onderwijs, 2018), waardoor het des te belangrijker is om meer inzicht te krijgen in de factoren die van invloed zijn op (genderongelijkheid in) leesprestaties. Het is aan toekomstig onderzoek met andere designs en data om uitsluitsel te geven over de vraag of de verwachting van de leerkracht een rol speelt bij genderverschillen in leesprestaties.

Noten

1. In tegenstelling tot leerkrachtverwachtingen ervaren door leerlingen, bevat de data wel gegevens over leerkrachtverwachtingen ervaren door de leerkracht. Echter is dit gemeten op klasniveau en daardoor niet geschikt voor het beantwoorden voor beide onderzoeksvragen. Desalniettemin hebben we deze alternatieve operationalisatie gebruikt voor het testen van het moderatie-effect. Net als bij de schaal van leerkrachtverwachtingen ervaren door leerlingen, vinden we ook voor de alternatieve maat geen significant effect (zie online Appendix B).
2. De online bijlage bij dit artikel is te vinden op <https://doi.org/10.5117/aup.19706581.v1>.
3. Het meenemen van de losse items, in plaats van gecombineerd in een schaal, leidt tot iets meer missings waardoor we nu een steekproef hebben van $N = 3.886$ in plaats van $N = 3.995$. Dit zorgt ook voor een klein verschil in de intercept en het hoofdeffect van gender in model 1 van tabel 3.

Literatuur

- Åhslund, I., & Boström, L. (2018). Teachers' Perceptions of Gender Differences: What about Boys and Girls in the Classroom?. *International Journal of Learning, Teaching and Educational Research*, 17(4), 28-44.
- Beaman, R., Wheldall, K., & Kemp, C. (2006). Differential teacher attention to boys and girls in the classroom. *Educational review*, 58(3), 339-366.
- Boerma, I. E., Mol, S. E., & Jolles, J. (2016). Teacher perceptions affect boys' and girls' reading motivation differently. *Reading Psychology*, 37(4), 547-569.
- Cooper, H. M. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational research*, 49(3), 389-410.
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific studies of reading*, 10(3), 277-299.
- De Boer, H., Bosker, R. J., & van der Werf, M. P. C. (2010). Sustainability of Teacher Expectation Bias Effects on Long-Term Student Performance. *Journal of Educational Psychology*, 102(1), 168-179.
- Epstein, D. (1998). Failing boys?: Issues in gender and achievement. *McGraw-Hill Education (UK)*.
- Fan, W. (2011). Social influences, school motivation and gender differences: an application of the expectancy-value theory. *Educational Psychology*, 31(2), 157-175.

- Gentrup, S., & Rjosk, C. (2018). Pygmalion and the gender gap: do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling? *Educational Research and Evaluation*, 24(3–5), 295–323.
- Good, T. L., & Nichols, S. L. (2001). Expectancy effects in the classroom: A special focus on improving the reading performance of minority students in first-grade classrooms. *Educational Psychologist*, 36(2), 113-126.
- Good, T. L. (1987). Two decades of research on teacher expectations: Findings and future directions. *Journal of teacher education*, 38(4), 32-47.
- Gubbels, J., Netten, A., & Verhoeven, L. (2017). *Vijftien jaar leesprestaties in Nederland: PIRLS-2016*. Nijmegen, The Netherlands: Expertisecentrum Nederlands, Radboud Universiteit Behavioural Science Institute.
- Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of educational psychology*, 101(3), 662.
- Hooper, M., Mullis, I. V., & Martin, M. O. (2016). PIRLS 2016 context questionnaire framework. *PIRLS*, 31-53.
- Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications* (2nd ed.). New York, NY: Routledge.
- Hughes, J. N., Luo, W., Kwok, O. M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of educational psychology*, 100(1), 1.
- Inspectie van het Onderwijs (2017). *De Staat van het Onderwijs. Onderwijsverslag 2015-2016*. Utrecht, The Netherlands: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2018). *De staat van het onderwijs. Onderwijsverslag 2016-2017*. Utrecht, the Netherlands: Inspectie van het Onderwijs.
- Jones, S. M., & Dindia, K. (2004). A meta-analytic perspective on sex equity in the classroom. *Review of educational research*, 74(4), 443-471.
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and social psychology review*, 9(2), 131-155.
- Kotte, D., Lietz, P., & Lopez, M. M. (2005). Factors Influencing Reading Achievement in Germany and Spain: Evidence from PISA 2000. *International Education Journal*, 6(1), 113-124.
- Kretschmann, J., Westphal, A., & Vock, M. (2021). Does it pay to be one of the oldest in class? Relative age effects on academic self-concept, peer relations, and teacher judgments in German primary schools. *Learning and Instruction*, 74, 101463.
- Kriegbaum, K., Steinmayr, R., & Spinath, B. (2019). Longitudinal reciprocal effects between teachers' judgments of students' aptitude, students' motivation, and grades in math. *Contemporary Educational Psychology*, 59, 101807.

- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.) (2017). *Methods and Procedures in PIRLS 2016*. Geraadpleegd van Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>
- Martin, M. O., & Mullis, I. V. (2013). *TIMSS and PIRLS 2011: Relationships among Reading, Mathematics, and Science Achievement at the Fourth Grade-Implications for Early Learning*. International Association for the Evaluation of Educational Achievement.
- Ma, X. (2008). Within-school gender gaps in reading, mathematics, and science literacy. *Comparative Education Review*, 52(3), 437-460.
- McKown, C., & Weinstein, R. S. (2002). Modeling the role of child ethnicity and gender in children's differential response to teacher expectations. *Journal of Applied Social Psychology*, 32(1), 159-184.
- McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of school psychology*, 46(3), 235-261.
- Meece, J. L., Glienke, B. B. & Burg, S. (2006). Gender and motivation. *Journal of School Psychology*, 44(5), 351-373.
- Nilsen, T., & Bergem, O. K. (2020). Teacher Competence and Equity in the Nordic Countries. Mediation and moderation of the relation between SES and achievement. *Acta Didactica Norden*, 14(1), 1-26.
- OECD (2014). (Revised ed.). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science, Volume I, . PISA, OECD*.
- OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*. Paris: OECD Publishing.
- Onderwijsraad (2020). Een verkenning van sekse verschillen in het onderwijs. Geraadpleegd van <https://www.onderwijsraad.nl/publicaties/adviezen/2020/10/07/verkenning-sekseverschillen-onderwijs>
- Pansu, P., Régner, I., Max, S., Colé, P., Nežlek, J. B., & Huguet, P. (2016). A burden for the boys: Evidence of stereotype threat in boys' reading performance. *Journal of Experimental Social Psychology*, 65, 26-30.
- PIRLS 2016 International Database. (2018). *TIMSS and PIRLS International Study Center*. Geraadpleegd van <https://timssandpirls.bc.edu/pirls2016/international-database/index.html>
- Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept—Testing the reciprocal effects model. *Learning and Instruction*, 29, 21-30.
- Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, 107(1), 1-33.

- Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal*, 48(2), 268-302.
- Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, 3, 16-20.
- Sieben, S., & Lechner, C. M. (2019). Measuring cultural capital through the number of books in the household. *Measurement Instruments for the Social Sciences*, 1(1), 1-6.
- Trouilloud, D. O., Sarrazin, P. G., Martinek, T. J., & Guillet, E. (2002). The influence of teacher expectations on student achievement in physical education classes: Pygmalion revisited. *European Journal of Social Psychology*, 32(5), 591-607.
- Tsiplakides, I., & Keramida, A. (2010). The Relationship between Teacher Expectations and Student Achievement in the Teaching of English as A Foreign Language. *English Language Teaching*, 3(2), 22-26.
- Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, 47(2), 497-527.
- Wang, S., Rubie-Davies, C. M., & Meissel, K. (2018). A systematic review of the teacher expectation literature over the past 30 years. *Educational Research and Evaluation*, 24(3-5), 124-179.
- Wolter, I., Braun, E., & Hannover, B. (2015). Reading is for girls!? The negative impact of preschool teachers' traditional gender role attitudes on boys' reading related motivation and skills. *Frontiers in psychology*, 6, 1-11.

Over de auteurs

Mayke Nollet behaalde haar bachelordiploma Sociologie aan de Universiteit Utrecht en is momenteel masterstudent Sociology and Social Research aan de Universiteit Utrecht. Ze is geïnteresseerd in onderwijsongelijkheden, sociale ongelijkheid en familiesociologie.

Kim Stienstra is promovendus aan de afdeling Sociologie van de Universiteit Utrecht/ICS. In haar proefschrift onderzoekt ze of de schoolomgeving de invloed van sociale afkomst, geslacht en genen op de onderwijsresultaten van kinderen vergroot of verkleint.