

Allochtone leraren en prestaties van allochtone leerlingen

Geert Driessen, ITS, Radboud Universiteit Nijmegen

Inhoud

-
- 1 Achtergronden
 - 1.1 Achtergronden van onderwijsachterstanden
 - 1.2 Ondervertegenwoordiging van allochtone leraren
 - 1.3 Het belang van allochtone leraren

 - 2 Probleemstelling en onderzoeksopzet
 - 2.1 Vraagstelling
 - 2.2 Onderzoeksopzet
 - 2.3 Methodologische kanttekeningen

 - 3 Onderzoeksbevindingen
 - 3.1 Selectie studies
 - 3.2 Resultaten studies
 - 3.2.1 Objectieve effectmaten
 - 3.2.2 Subjectieve effectmaten
 - 3.3 Een samenvattend overzicht

 - 4 Samenvatting en conclusies
 - 4.1 Samenvatting
 - 4.2 Conclusies

Literatuur

1 Achtergronden

1.1 Achtergronden van onderwijsachterstanden

Wanneer het gaat om de verklaring van verschillen in onderwijsprestaties, worden op het niveau van de leerlingen traditioneel vaak drie relevante factoren genoemd: de sekse, het sociaal-economische milieu en de etnische herkomst. In de loop der jaren is er uitvoerig onderzoek naar deze relaties verricht, is nagegaan hoe eventuele negatieve effecten voorkomen dan wel bestreden kunnen worden en is naar aanleiding daarvan gericht beleid gevoerd om onderwijsongelijkheid te verkleinen (Driessen, 2013).

Verreweg het meeste onderzoek is uitgevoerd vanuit het perspectief van de leerling. Hierbij kunnen twee niveaus worden onderscheiden, te weten het niveau van de individuele leerling en het niveau van de leerlingen in de klas of school. Er is bijvoorbeeld nagegaan wat het effect is van de etnische/allochtone herkomst van een leerling op diens prestaties en daarnaast wat het effect is van de etnische samenstelling (c.q. compositie, diversiteit) van een klas op de prestaties van de leerlingen (Driessen, 2007; Dronkers & Van der Velden, 2013; Herweijer, 2011). Vergelijkbaar onderzoek is verricht naar de relatie sekse-prestaties en de relatie milieu-prestaties.

Er is echter ook ander relevant onderzoek denkbaar, onderzoek vanuit het perspectief van de leraar, naar de relatie tussen de sekse, het milieu en de etniciteit van de leraar en de prestaties van de leerling. Ook hier zijn weer twee varianten te onderscheiden: onderzoek waarbij de individuele relatie tussen leraar en leerling wordt onderzocht, en onderzoek waarbij de relatie tussen de compositie en diversiteit van het lerarenteam en de leerling voorwerp van onderzoek vormt. Dit type onderzoek is echter veel minder verricht, en dat geldt zeker voor Nederland.

1.2 Ondervertegenwoordiging van allochtone leraren

In het onderhavige onderzoek staat de relatie tussen de allochtone herkomst van de leraar en de prestaties van allochtone leerlingen centraal. Allereerst moet worden geconstateerd dat het aantal allochtone leraren in Nederland bijzonder laag is. In 2009 (recentere gegevens ontbreken) was slechts 3,7% van het personeelsbestand in het primair onderwijs niet-westers allochtoon; in het voortgezet onderwijs ging het om 4,7%. Daarbij dient te worden aangetekend dat een groot deel van hen niet als leraar voor de klas staat, en al helemaal niet in het management zit, maar een ondersteunende functie heeft. Bovendien is het aandeel allochtone personeelsleden in het primair onderwijs – ondanks beleidsvoornemens en -maatregelen – de afgelopen jaren nog afgenomen. Dit wordt mede veroorzaakt door de lage instroom van allochtone studenten in de lerarenopleidingen en de hoge uitval tijdens de opleiding. Dit laatste leidt ertoe dat het studierendement van allochtone leraren in opleiding bijzonder laag is: voor het primair onderwijs minder dan de helft van dat van hun autochtone jaargenoten (22,6 vs. 50,5%) (Van den Berg, Van Dijk & Grootscholte, 2011). De percentages allochtone leraren steken schril af tegen die van de allochtone leerlingen: in het schooljaar 2009/2010 was 11% van de leerlingen in het basisonderwijs een zogenoemde cumi-leerling (culturele minderheid) en 15% van de leerlingen in het voortgezet onderwijs niet-westers allochtoon (CBS, 2012, 2013). Al met al moet daarom worden geconcludeerd dat allochtone leraren fors zijn ondervertegenwoordigd in het primair en voortgezet onderwijs, en dat bovendien de kans dat dit in de toekomst – zonder stevige en succesvolle maatregelen – gaat verbeteren bijzonder gering is.

1.3 Het belang van allochtone leraren

Waarom is een goede vertegenwoordiging van allochtone leraren van belang? In algemene zin wordt doorgaans gewezen op een mogelijk positief effect van de

allochtone herkomst van leraren op de prestaties van hun allochtone leerlingen. Dat is waarschijnlijk het belangrijkste argument dat genoemd wordt, maar zeker niet het enige. In de literatuur worden uiteenlopende daaraan ten grondslag liggende of daarmee samenhangende veronderstellingen opgevoerd; hierna volgen de meest genoemde (bv. Van den Berg, Van Dijk & Grootsholte, 2011; Bone & Slate, 2011; Clewell & Villegas, 1998; FORUM, 2012; Irvine & Fenwick, 2009; Keijzer, Winthagen & De Vries, 2008; Kregten & Moerkamp, 2008; Marreveld, 2007; Maylor, 2009; Naman, 2009; National Collaborative on Diversity in the Teaching Force, 2004; Villegas & Irvine, 2010):

- De allochtone leraar fungeert als rolmodel voor niet alleen allochtone leerlingen, maar ook voor autochtone leerlingen.
- Allochtone leraren zorgen voor verbetering van de contacten tussen de school en allochtone ouders en leerlingen.
- Ouders en leerlingen herkennen zich beter in de school en krijgen als gevolg daarvan meer affiniteit met het onderwijs; ouderbetrokkenheid en -participatie nemen daardoor toe.
- Door hun kennis van de allochtone cultuur hebben de leraren een belangrijke signalerende functie; zij herkennen eerder wanneer er problemen zijn (of dreigen te ontstaan) bij betreffende leerlingen en ingegrepen moet worden.
- Allochtone leraren vervullen vanwege hun kennis van en ervaring met de andere culturen een vraagbaak- en scholingsfunctie voor hun autochtone collega's.
- Ze leveren door hun aanwezigheid en specifieke kennis een bijdrage aan het voorbereiden van alle leerlingen op het leven in de Nederlandse multiculturele samenleving.
- In een breder perspectief zijn meer allochtone leraren van belang voor het realiseren van een goede etnisch-culturele afspiegeling van de samenleving.
- Een laatste argument is dat zij (toekomstige) lerarentekorten kunnen opvullen.

In de literatuur wordt vaak aangegeven dat het noodzakelijk is meer allochtone leraren te werven. In veel landen zijn daartoe de afgelopen decennia ook verschillende initiatieven genomen (Bone & Slate, 2011). Waaróm er echter meer allochtone leraren noodzakelijk zijn, daarvoor worden weliswaar uiteenlopende veronderstellingen en argumenten aangedragen, maar deze worden nauwelijks met empirisch onderzoek onderbouwd (Villegas & Irvine, 2010). De vraag luidt daarom: in hoeverre kloppen deze veronderstellingen?

2 Probleemstelling en onderzoeksopzet

2.1 Vraagstelling

De verwachting is dat wanneer er meer allochtone leraren zijn, er daardoor qua etniciteit meer 'congruentie' (Pigott & Cowen, 2000) en 'match' (Bates & Glick, 2013) is met de allochtone leerlingen en de prestaties van die leerlingen zullen verbeteren. In hoeverre deze verwachting door empirisch onderzoek wordt ondersteund, is vooralsnog onduidelijk. Voor zover er wetenschappelijk onderzoek naar is verricht, is dat nauwelijks systematisch bij elkaar gezet. Sommige studies laten positieve effecten zien, andere géén effecten, en van weer andere zijn de resultaten ambigu. In een enkel geval wordt ook op negatieve effecten gewezen. FORUM, Instituut voor Multiculturele Vraagstukken heeft het ITS van de Radboud Universiteit Nijmegen daarom gevraagd een verkennende internationale literatuurstudie te verrichten, waarin de uitkomsten van empirisch wetenschappelijk onderzoek worden geïnventariseerd en samengevat. De hoofdvraagstelling van het onderzoek luidt:

Wat is het effect van allochtone leraren op de prestaties van allochtone leerlingen?

Bij deze vraagstelling horen enkele specificaties. Qua sector richt dit onderzoek zich op het primair en voortgezet onderwijs. Hoewel in principe geen landen uitgesloten worden,

ligt het accent op onderzoek dat betrekking heeft op Nederland, België, Duitsland, Engeland en de Verenigde Staten.

Centraal in het onderzoek staan de prestaties van de allochtone leerlingen. 'Prestaties' worden echter breed opgevat. Het kan bijvoorbeeld gaan om toetsprestaties, maar ook om rapportcijfers, de beoordeling van capaciteiten door leraren, adviezen voor vervolgonderwijs, het niveau van vervolgonderwijs en examenresultaten en slaagkansen. Naast dergelijke cognitieve opbrengstmaten kunnen er ook niet-cognitieve resultaten worden onderscheiden, zoals motivatie, werkhouding en gedrag. Dit type kenmerken kan indirect bijdragen aan betere cognitieve resultaten. Het onderzoek richt zich op zowel cognitieve als niet-cognitieve resultaten. Het hoofdaccent ligt op de allochtone leerlingen. De aanwezigheid van allochtone leraren kan echter ook de prestaties van de autochtone klasgenoten – positief dan wel negatief – beïnvloeden. Ook deze effecten worden meegenomen.

Hiervoor is er al op gewezen dat in onderzoek vaak twee niveaus worden onderscheiden, het individuele en het klas- of schoolniveau. Concreet betekent dit hier dat er onderzoek mogelijk is naar het effect van de individuele allochtone leraar op de prestaties van de leerlingen, en naar het effect van het aandeel allochtone leraren en de etnische diversiteit in het lerarenteam op de prestaties van de leerlingen. In het onderzoek zijn beide varianten meegenomen.

In de volgende twee paragrafen wordt nader ingegaan op de methode van onderzoek. In het hoofdstuk daarna (3) worden de resultaten van het onderzoek beschreven en in het laatste hoofdstuk (4) wordt een samenvatting gegeven en worden conclusies getrokken. Het rapport wordt afgesloten met een overzicht van de gebruikte literatuur.

2.2 Onderzoeksopzet

De methode van onderzoek is een internationale literatuurstudie. Daarbij zijn vier bronnen gebruikt:

- de bij de auteur beschikbare kennis en publicaties;
- deze zijn aangevuld met gerichte literatuur-searches via zoeksystemen als Scopus, Web of Science, Picarta, Psychinfo, SciVerse, ERIC, Online Contents en Scholar;
- via de zogenoemde 'snowball'-methode is vervolgens via de referentielijst van relevant gebleken publicaties steeds verder gezocht;
- daarnaast zijn (internationale) collega-onderzoekers benaderd met de vraag of ze publicaties op dit terrein kennen.

Bij het zoeken naar relevante literatuur zijn verschillende selectiecriteria aangehouden:

- het dient te gaan om verslagen van (in principe kwantitatief) empirisch onderzoek of om samenvattingen van en verwijzingen naar dergelijk onderzoek;
- de geanalyseerde steekproef dient van voldoende omvang te zijn en er dient sprake te zijn van een (quasi-)experimentele opzet met controle voor versturende kenmerken;
- het moet gaan om onderzoek dat betrekking heeft op het primair of voortgezet onderwijs;
- de publicatie dient niet ouder te zijn dan twintig jaar;
- hoewel in principe geen enkel relevant onderzoek wordt uitgesloten, ligt het accent op Nederland, België, Duitsland, Engeland en de VS;
- de etniciteit van de leraar of leraren dient een centrale verklarende variabele te zijn, evenals de etniciteit van de leerling, maar het liefst de combinatie c.q. interactie van beide;
- er dienen cognitieve en/of niet-cognitieve effecten te worden bepaald bij leerlingen.

Bij de literatuur-search zijn talloze (combinaties van) zoektermen gebruikt. Enkele voorbeelden zijn: racial pairings; ethnic/racial (mis)matching; own-race teacher; racial

(in)congruence; racial (a)symmetry; race consonance; racial disconnect; ethnic staff composition; Lehrer; Lehrkräfte; Migrationshintergrund; Ethnizität.

Van elke gevonden en relevant lijkende publicatie is een digitale fiche aangemaakt met de volgende rubrieken:

- onderzoeksjaar (jaar/jaren van dataverzameling);
- land;
- onderwijssector (PO of VO);
- steekproef (aantal scholen, leraren, leerlingen);
- te verklaren kenmerk (cognitief, niet-cognitief);
- verklarend kenmerk (ethniciteit leraar x ethniciteit leerlingen);
- methode, analysetechniek;
- resultaten;
- effectgrootte, sterkte van de samenhang.

Gaandeweg bleek dat lang niet altijd alle rubrieken konden worden ingevuld, vaak omdat de betreffende gegevens geheel of gedeeltelijk ontbraken of onduidelijk waren. In de praktijk is daarom vooral een beknopte beschrijving van de resultaten gegeven, waarbij gefocust is op 'ethniciteit leraar' in combinatie met 'ethniciteit leerling' in relatie tot prestaties.

Omdat het een internationale literatuurstudie betreft, hebben we te maken met verschillend woordgebruik, afhankelijk van de context. In Nederland zijn 'allochtoon' en 'autochtoon' en 'ethniciteit' of 'etnische herkomst' ingeburgerde aanduidingen. In de Amerikaanse situatie spreekt men doorgaans in termen van 'race', 'ethnicity', 'color' of 'minority' en wordt een onderscheid gemaakt tussen 'White', 'Black' (of 'African American'), 'Hispanic' of 'Latino'.

2.3 Methodologische kanttekeningen

Onderzoek naar effecten van de compositie of diversiteit van een school, bijvoorbeeld in termen van ethniciteit van de leerlingen, is complex en er passen allerlei methodologische voorbehouden bij; deze zijn uitvoerig beschreven in Driessen (2007). Wanneer echter effecten van kenmerken van leraren en teams centraal staan, wordt de situatie nog ingewikkelder. De vraag is namelijk hoe lang dergelijke effecten doorwerken. Doet zo'n effect zich alleen voor wanneer de leerling in de klas van de betreffende leraar of leraren zit, of ook daarna? De verwachting is dat effecten op het cognitieve vlak (zoals prestaties) langer zullen doorwerken, terwijl effecten op het niet-cognitieve vlak (zoals motivatie, werkhouding en sociaal gedrag) eerder gekoppeld zijn aan een bepaalde leraar in een bepaalde jaargroep of klas. In een vergelijkbaar onderzoek, maar dan naar het effect van mannelijke leraren in het basisonderwijs, hebben Driessen en Doesborgh (2004) daarom verschillende aan het tijdstip of de periode gekoppelde effecten onderscheiden:

- Het cumulatieve effect, dat wil zeggen het aantal mannelijke leraren van wie de leerlingen in de groepen 1–8 les hebben gehad.
- Het fase-effect, waarbij is nagaan of er effecten zijn van de fase waarin de leerlingen les hebben gehad van mannelijke leraren. Het kan immers zijn dat er verschillen zijn tussen de lagere en de hogere jaargroepen.
- Het moment-effect, dat wil zeggen het effect van het geslacht van de leraar die de leerlingen op een bepaald moment, in een bepaalde jaargroep of klas, hebben gehad.

Wanneer een dergelijk onderscheid in de analyses niet wordt gemaakt, is onduidelijk wie het effect heeft veroorzaakt. In het basisonderwijs, met meestal één leraar per jaar maar verschillende leraren over de hele loopbaan, is dat al ingewikkeld, maar in het voortgezet onderwijs – met voor elk vak andere leraren – is het nog complexer. Bij leraarbeoordelingen ligt dit waarschijnlijk anders dan bij toetscores; bij (subjectieve) oordelen gaat het in belangrijke mate om hier-en-nu-ervaringen, terwijl (objectieve)

toetscores (c.q. vaardigheden) meer het resultaat zijn van een heel traject, i.c. van leerervaringen opgedaan bij meerdere leraren in eerdere leerjaren.

3 Onderzoeksbevindingen

3.1 Selectie studies

De literatuur-search leverde, na een grondige schifting aan de hand van de eerder genoemde selectiecriteria, uiteindelijk 24 artikelen op met een beschrijving van empirische studies. Veel titels zijn uit de lijst geschrapt, met name omdat ze kleinschalig en kwalitatief van aard waren, omdat het overzichtsstudies betrof, omdat ze handelden over lerarentekorten of theoretisch-beschrijvend van aard waren (vgl. Bone & Slate, 2011; Burt, Ortlieb & Cheek, 2009; Cizek, 1995; Clewell & Villegas, 1998; Cooper, 2003; Ferguson, 2003; Irvine & Fenwick, 2009; Naman, 2009; Villegas & Irvine, 2010; Weinberg, 2013). Er is nauwelijks empirisch onderzoek naar het (al dan niet directe) effect van leraren uit minderheden op leerlingprestaties. Veel onderzoek focust op de houdingen, verwachtingen, percepties en (positieve en negatieve) feedback van leraren en de plaatsing/allocatie van leerlingen (bv. Van den Berg et al., 2010; Tenenbaum & Ruck, 2007; Uhlenberg & Brown, 2002). Meer concreet gaat het vaak om onderzoek dat voortbouwt op onderzoek rond het Pygmalion-effect (Van der Hoeven-Van Doornum, 1990; Jungbluth, 2003; Rosenthal & Jacobson, 1968). De achterliggende gedachte is dat stereotiepe verwachtingen tot 'self-fulfilling prophecies' (zichzelf waarmakende voorspellingen) leiden (Naman, 2009; Takei & Shouse, 2008). De etniciteit van de leraren is in dergelijk onderzoek echter bijna nooit meegenomen.

Kenmerkend voor de uiteindelijk geselecteerde studies is dat ze allemaal betrekking hebben op de VS. Er is dus geen enkel onderzoek gevonden dat verricht is in Nederland, België, Duitsland of Engeland. Een ruime meerderheid van de studies is uitgevoerd in het primair onderwijs, al dient daarbij meteen te worden aangetekend dat het onderscheid naar sector niet steeds eenduidig is. Dat heeft te maken met het feit dat het Amerikaanse onderwijsstelsel een andere fasering kent dan het Nederlandse en dat daar ook nog flinke verschillen bestaan naar schooldistrict.

In praktisch al deze studies worden relatief grote steekproeven geanalyseerd (vaak cohortdata, zoals NELS88, HSES en ELS-K-5) en wordt in de analyses gecontroleerd voor (c.q. rekening gehouden met) kenmerken op het niveau van de leerling/het gezin (zoals etniciteit, sociaal milieu en sekse), de leraar (bv. kwalificaties) en de klas/school (bv. etnische samenstelling leerlingenpopulatie). Tegelijk met de etniciteit van de leraar wordt soms ook de sekse van de leraar meegenomen (bv. Ehrenberg, Goldhaber & Brewer, 1995; Howsen & Trawick, 2007; Ouazad, 2007). De etnische samenstelling of diversiteit van het lerarenteam wordt praktisch nooit meegenomen. In de enkele gevallen dat dit wel gebeurt, zijn er verschillende varianten: het percentage allochtone leraren op een school op een bepaald moment (cross-sectioneel), het percentage allochtone leraren in een schooldistrict, het percentage allochtone schoolleiders in een district, en het les hebben gehad van ten minste één allochtone leraar over een bepaalde periode (longitudinaal). Soms worden in onderzoek alle leerlingen meegenomen, dat wil zeggen alle etnische minderheidsgroepen en autochtone leerlingen. Soms wordt echter ook gefocust op één groep, bijvoorbeeld zwarten of Latino's.

Dat al het onderzoek betrekking heeft op de VS, heeft consequenties voor de mate waarin de bevindingen ook gelden voor de Nederlandse situatie. Er spelen op z'n minst twee (samenhangende) problemen bij een vergelijking met de VS. De verschillen in achtergronden (incl. woon/leefsituatie van leerlingen uit minderheidsgroepen en de organisatie van het onderwijs) en prestaties tussen enerzijds White en anderzijds African American en Hispanic leerlingen, zijn daar aanmerkelijk geprononceerder (Fryer & Levin, 2004). Bovendien zijn de kwaliteitsverschillen tussen scholen met veel en weinig achterstandskinderen daar ook extremer, wat voor een belangrijk deel wordt

veroorzaakt door de andere financiering van scholen en – daarmee samenhangend – de kwaliteit van de leraren. Relatief vaak gaat het op deze scholen (meestal in stedelijke achterstandswijken) om on- of ondergekwalficeerde leraren met geringe onderwijservaring en is er bovendien sprake van veel verloop (Bireda & Chait, 2011; Clotfelter, Ladd & Vigdor, 2004; Eddy & Easton-Brooks, 2011; Irvine & Fenwick, 2009; Pitts, 2007).

3.2 Resultaten studies

Hierna volgt allereerst een (alfabetisch op auteur geordende) beschrijving van de resultaten van de studies, waarbij ook summier wordt ingegaan op de aard van het onderzoek. Daarbij is geprobeerd een onderscheid te maken tussen effecten in het objectieve en in het subjectieve domein (bv. scores op door leerlingen gemaakte toetsen, resp. beoordelingen door leraren). Daarbinnen wordt vervolgens, voor zover mogelijk, nog een onderscheid gemaakt tussen effecten op het individuele niveau, i.c. 'matching' leraar x leerling, en op het klas-, school- of districtsniveau, i.c. com-positie-effecten. Het hoofdstuk wordt afgesloten met een samenvattend overzicht van de bevindingen.

3.2.1 Objectieve effectmaten

3.2.1.1 Matching

In het onderzoek van Clotfelter, Ladd en Vidgor (2006) stond de matching leraar-leerling centraal, met name qua niveau. De vraag was of hooggekwalficeerde leraren ook vooral lesgeven aan hoogpresterende leerlingen. Er werd een serie leraar-kenmerken meegenomen, zoals ervaring, competenties, sekse en etniciteit. Qua etniciteit werd een onderscheid gemaakt naar blank, zwart en Hispanic. Het prestatieniveau van de leerlingen werd geïndiceerd middels toetsscores wiskunde en lezen. Gebruik werd gemaakt van gegevens van 3.223 leraren en 60.791 10-jarige leerlingen. De onderzoeksresultaten lieten zien dat – inderdaad – goedpresterende leerlingen in sterkere mate les kregen van hooggekwalficeerde leraren. Zwarte en andere niet-blanke leraren gaven les aan leerlingen met lagere toetsprestaties, niet alleen over scholen heen maar ook binnen scholen. Dit betekent overigens niet per se dat er sprake is van een causale relatie. Dat negatieve verband verdween echter grotendeels of helemaal wanneer rekening werd gehouden met controlevariabelen op leerlingniveau, zoals sekse, etniciteit en ouderlijke opleiding. Wanneer vervolgens ook nog de eerdere prestaties werden verdisconteerd, verdween het negatieve effect van zwarte of Hispanic leraren helemaal.

Dee (2001, 2004a,b) maakte voor zijn analyses gebruik van data uit het bekende STAR-experiment, waarin het effect van klassegrootte werd bepaald. Bij Dee ging het om de vraag of etnisch matchen van leerling met leraar van invloed is op de reken- en leesprestaties. In totaal waren er circa 11.600 5- tot 8-jarige leerlingen van 79 scholen bij zijn onderzoek betrokken. De analyses lieten zien dat er positieve effecten optraden van het les krijgen van een leraar met dezelfde etnische herkomst, voor zowel rekenen/wiskunde als lezen, en dat deze effecten nagenoeg gelijk waren voor blanke en zwarte leerlingen en ook nauwelijks veranderden wanneer gecontroleerd werd voor leerling-, klas- en schoolkenmerken. Bovendien was er sprake van cumulatieve effecten over de vier onderzochte jaren heen.

Easton-Brooks, Lewis en Zhang (2010) analyseerden longitudinale gegevens van 1.207 5- tot 10-jarige zwarte leerlingen uit het cohort ECLS-K-5. De vraag was of zwarte kinderen beter presteerden op een leestoets als ze les kregen van een zwarte leraar. Deze gecalibreerde (en daardoor over de tijd vergelijkbare) toets werd aan het eind van elk van de zes schooljaren afgenomen. Onderzocht werd of verschillen in lees-prestaties konden worden verklaard door etnisch matchen. Dit matchen werd als volgt geoperationaliseerd: zwarte leerlingen hebben in de zes leerjaren les gehad van ten

minste één zwarte leraar versus zwarte leerlingen hebben in die zes jaren alleen les gehad van blanke leraren. Uit de analyses bleek dat de leesprestaties van de zwarte leerlingen aan het eind van de Amerikaanse vroegschool, genaamd kindergarten (5 jaar), en de groei in prestaties tot grade 5 (10 jaar) beter waren wanneer ze les hadden gehad van ten minste één zwarte leraar. Maar nadat in de analyses rekening was gehouden met het aandeel zwarte leerlingen op de school, verdween dit effect.

Het onderzoek van Eddy en Easton-Brooks (2011) is in grote lijnen vergelijkbaar met dat van Easton-Brooks, Lewis en Zhang (2010). Het verschil zit voornamelijk in het feit dat bij het eerste onderzoek de rekenprestaties voorwerp van analyse vormden en bij het tweede de leesprestaties. Deze vergelijking tussen leerstofdomen is vooral interessant omdat het een gegeven is dat rekenen, veel meer dan lezen, bepaald wordt door leraar- en schoolkenmerken. Bij rekenen zou dus eerder een effect van matching verwacht kunnen worden. Deelnemers aan het onderzoek waren 1.207 zwarte 5- tot 10-jarige leerlingen uit het cohort ECLS-K-5, bij wie op vier meetmomenten een rekentoets was afgenomen. Ook hier gold dat er sprake was van etnische match wanneer een zwarte leerling in de zes bestudeerde jaren les had gehad van ten minste één zwarte leraar. De resultaten van de analyses met de rekenprestaties waren vergelijkbaar met die met de leesprestaties. De rekenprestaties van zwarte leerlingen op het eind van kindergarten en de groei in prestaties tot grade 5 waren hoger wanneer ze les hadden gehad van ten minste één zwarte leraar. Maar wanneer er gecontroleerd werd voor sekse, school, armoede en percentage allochtonen was er geen effect meer.

Ehrenberg, Goldhaber en Brewer (1995) analyseerden tegelijkertijd de etnische match en sekse-match. Ze maakten gebruik van de grootschalige NELS88-cohortdata uit 1998 en 1990 met gegevens van rond de 2.500 leerlingen. Geanalyseerd werden verschillen op toetsen die op 13- en 15-jarige leeftijd waren afgenomen op twee van de gebieden lezen/Engels, wiskunde, geschiedenis/social studies, en science. Naast deze objectieve prestatie-maten werden ook de (subjectieve) leraaroordelen over het gedrag en de capaciteiten van de leerlingen geanalyseerd. De centrale vraag was in hoeverre de etnische herkomst en sekse van de leraren de leerwinst van de leerlingen beïnvloedden. Op basis van objectieve prestatie-maten (de toetsen) kon worden geconcludeerd dat er op zijn best weinig ondersteuning is voor de verwachting dat de etniciteit en sekse van de leraren van invloed zijn op de prestaties. De analyse van de subjectieve leraarbeoordelingen liet echter zien dat er in een aantal gevallen wel sprake was van effecten van de combinatie van de etniciteit en sekse van de leraar op de beoordelingen van de leerlingen. En bovendien: dat de etnische match soms tot meer positievere beoordelingen leidde.

Fryer en Levitt (2004) maakten gebruik van data van circa 12.000 leerlingen uit het kindergarten-cohort ECLS-K. Er werden zowel objectieve maten geanalyseerd, namelijk lees- en rekentest scores op 5- en op 6- jarige leeftijd, als subjectieve maten, te weten het oordeel van de leraren over het lees- en rekenniveau van de kinderen. De analyses maakten duidelijk dat er grote verschillen waren tussen blanke en zwarte leerlingen bij hun start in de kindergarten. Deze verdwenen echter wanneer gecontroleerd werd voor enkele achtergrondkenmerken. In de twee jaren daarna raakten zwarte leerlingen echter achterop. Een belangrijke oorzaak hiervoor ligt volgens de onderzoekers waarschijnlijk bij de veel lagere kwaliteit van scholen met veel zwarte leerlingen. De kwaliteit van de leraren is waarschijnlijk de belangrijkste determinant van schoolkwaliteit. Wanneer het gaat om een mogelijk effect van de etniciteit van de leraren, luidde de these dat blanke leraren lagere verwachtingen hebben van zwarte leerlingen, dan wel dat ze discrimineren. Als dit zo zou zijn, zouden zwarte leerlingen met blanke leraren meer achterop moeten raken. De analyses lieten echter het tegenovergestelde zien: zwarte leerlingen met ten minste één zwarte leraar gingen er qua toetsprestaties juist meer op achteruit. De analyse van de subjectieve leraaroordelen liet vergelijkbare patronen zien: zwarte leerlingen met zwarte leraren kregen geen positievere beoordelingen dan hun blanke klasgenoten.

Het onderzoek van Howsen & Trawick (2007) vormt een replicatie en uitbreiding van het onderzoek van Dee (2004). Er werden analyses verricht op gegevens van 25.871 8-jarige leerlingen op 120 scholen. Centraal stonden de reken- en leestoets-prestaties. Nagegaan werd of etnisch matchen een effect had op deze prestaties. Anders dan in het onderzoek van Dee waren nu meer kenmerken beschikbaar, te weten de scores op een cognitieve-capaciteitentoets en de sekse van de leraar. Bovendien waren ook kleine en grote scholen meegenomen; in het onderzoek van Dee ontbraken die. Uit de analyses volgde dat na controle voor sekse en capaciteiten er geen significant effect meer resteerde van etnisch matchen. Daarmee werd dus de conclusie van Dee ontkracht.

3.2.1.2 Compositie

Brown-Jeffy (2008) analyseerde gegevens van 177 scholen en 3.392 leerlingen (15- tot 17-jarigen) uit het HSES-cohort. Dit onderzoek focuste op de verklaring van verschillen in de prestaties op een wiskundetoets door de etnische compositie van de leerlingenpopulatie van een school. Een onderscheid werd gemaakt tussen blanke, zwarte, Hispanic en Aziatische leerlingen. Ook werd de etnische compositie van het lerarenteam met twee categorieën meegenomen, namelijk het percentage zwarte leraren en het percentage Hispanic leraren. Er werden drie hoofdanalyses uitgevoerd per schoolcompositietype (blank vs. zwart, vs. Hispanics, vs. Aziatisch). Daaruit bleek dat na controle voor een aantal leerling- en schoolkenmerken er verschillende effecten optraden van lerarencompositie. Voor de scholen met blanke en zwarte leerlingen en de scholen met blanke en Aziatische leerlingen was er een positief effect, maar dit was niet significant. Voor de scholen met blanke en Hispanic leerlingen was er wel een significant negatief effect. Dat betekent – tegengesteld aan de verwachting van de onderzoekers – dat op scholen met meer zwarte en Hispanic leraren de Hispanic leerlingen lagere toetsscores haalden.

De studie van Meier (1993) focuste op 12 schooldistricten met relatief veel Latino-leerlingen. Er waren gegevens beschikbaar over een periode van 11 jaar; deze werden gepoold tot één data-file. De vraag was of het aandeel Latino-directeuren en -leraren in een district samenhangt met een serie cognitieve en niet-cognitieve kenmerken van Latino-leerlingen. Daarbij werden drie dimensies onderscheiden: niveaugroepen (speciaal onderwijs, onderwijs voor begaafden), discipline, en prestaties. Uit de analyse volgde dat er voor alle drie de dimensies sprake was van een significante positieve correlatie met het aandeel Latino-leraren. De samenhangen met het aandeel Latino-directeuren waren veel zwakker en soms in de niet-verwachte richting. Nadat was gecontroleerd voor sociaal milieu en andere kenmerken bleven de verbanden voor de leraren in het algemeen bestaan. Dat gold eveneens voor die met de directeuren, maar daar waren de effecten vrijwel steeds tegengesteld aan de verwachtingen. Dit werd door de onderzoekers verklaard vanuit de gedachte dat er een 'critical mass' nodig is om een (positief) effect te genereren. Omdat er maar weinig Latino-directeuren waren, leverde dat dus een (statistisch) probleem op.

De eenheid van analyse in het onderzoek van Meier, Wrinkle en Polinard (1999) was het schooldistrict. Het betrof een steekproef van 350 schooldistricten met minimaal 1.000 leerlingen van wie tussen de 10% en 90% blank was. Er waren gegevens beschikbaar van zes jaren die werden gepoold tot één databestand. Nagegaan werd of er een samenhang bestond tussen het aandeel 8- tot 15-jarige leerlingen dat geslaagd was voor een gestandaardiseerde toets en het aandeel zwarte en Latino-leraren. Wat betreft de leerlingen werd een onderscheid gemaakt tussen blanke en niet-blanke leerlingen. De analyses lieten zien dat een groter aandeel zwarte en Latino-leraren een positief effect had op de slaagkans van zowel allochtone als autochtone leerlingen, hoewel dat voor de laatste categorie niet significant was.

Pitts (2007) analyseerde de gegevens van alle schooldistricten in Texas, gepoold over de periode 1995-2002. Drie kenmerken werden geanalyseerd: het slaagpercentage op het eindexamen, het drop-outpercentage en het percentage leerlingen met een hoge score op de SAT-toets (een capaciteitentoets). Nagegaan werd wat het effect was van

de verhouding tussen het percentage leraren respectievelijk schoolleiders en het percentage leerlingen van een etnische groep (blank, zwart, Latino). Uit de analyses volgde dat het percentage schoolleiders niet uitmaakt, volgens Pitts waarschijnlijk omdat ze ver afstaan van de leerlingen. Er deed zich wel een positief effect voor van het percentage leraren met betrekking tot twee van de drie effectvariabelen: er was minder sprake van drop-out en meer leerlingen slaagden voor hun eindexamen. Het bleek echter ook dat deze effecten negatief waren voor blanken en positief voor zwarten en Latino's.

3.2.2 Subjectieve effectmaten

3.2.2.1 Matching

De centrale vraag in het onderzoek van Bates en Glick (2013) was of leerlingen anders worden beoordeeld wanneer leraren en leerlingen een verschillende etnische herkomst hebben, dan wanneer ze dezelfde herkomst hebben ('racial/ethnic disparities vs. match'). Gebruik werd gemaakt van gegevens van 16.701 5- tot 10-jarige kinderen uit het cohort ECLS-K. Qua etniciteit van leraren en leerlingen was een onderscheid gemaakt in blank, Latino, zwart en Aziatisch. De kinderen waren beoordeeld op hun externaliserend gedrag (bv. disruptief of problematisch gedrag). De analysesresultaten lieten zien dat het gedrag van leerlingen met een verschillende etnische herkomst op een stereotiepe wijze werd beoordeeld door hun leraren. Zwarte leerlingen werden eerder negatief beoordeeld, Aziatische leerlingen eerder positief. Maar dat lag anders, c.q. positiever, wanneer ze werden beoordeeld door leraren met dezelfde etnische herkomst. Deze verschillen bleven bestaan wanneer gecontroleerd werd voor leerling-, leraar- en schoolkenmerken, waaronder ook de opleiding en sekse van de leraar. Met andere woorden: etnische leraar-leerling-congruentie kan een bijdrage leveren aan het voorkómen van stereotiepe beoordelingen.

In het onderzoek van Cullinan en Kaufmann (2005) ging het om de beoordeling van emotionele en gedragsproblemen ('education disability') van leerlingen door hun leraren. De analyses concentreerden zich op 248 zwarte en 524 blanke 6- tot 18-jarige leerlingen die door hun blanke dan wel zwarte leraren werden beoordeeld. De resultaten lieten zien dat zwarte leerlingen volgens hun leraren weliswaar vaker problemen hadden (m.n. op de schalen 'Ongelukkig of depressie' en 'Fysieke symptomen of angsten'), maar dat er daarbij geen relatie was met de etniciteit van de leraren; er was dus geen etnische bias in de beoordelingen.

De vraag die Dee (2005) wilde beantwoorden was of het matchen (o.m. op etniciteit) van leraren en leerlingen van invloed is op de subjectieve leraarbeoordelingen van leerlingen qua gedrag en prestaties. Hij maakte gebruik van de grootschalige NELS88-bestanden met gegevens van 21.324 13-jarige leerlingen van 1.052 scholen. Er werden drie gedragingen geanalyseerd, namelijk: verstoring gedrag, niet opletten, huiswerk niet gemaakt. Uit de analyses kwam naar voren dat wanneer leraren leerlingen met een andere etniciteit beoordeelden (zowel blank als zwart en Hispanic), die negatiever werden beoordeeld. Dit sterke effect trad echter niet bij alle leerlingen in dezelfde mate op: het deed zich vooral voor bij leerlingen uit lagere sociaal-economische milieus en uit het zuiden van de VS.

In het onderzoek van Downey en Pribesh (2004) stond de etnische match tussen leraar en leerling centraal. Beoordeeld werd het gedrag van leerlingen in de klas. Enerzijds ging het om gegevens uit het cohort ECLS-K (12.989 5-jarige leerlingen) en anderzijds het NELS88-cohort (8.881 13-jarige leerlingen). Uit het onderzoek bleek dat zwarte leerlingen negatiever werden beoordeeld dan blanke leerlingen. Maar wanneer de etniciteit van de leraar werd verdisconteerd, verdween dit effect. Er waren zelfs indicaties dat wanneer leerlingen werden gekoppeld aan leraren van dezelfde etniciteit, zwarte leerlingen positiever werden beoordeeld dan blanke leerlingen. Er leek meer ondersteuning voor de hypothese dat het matching-effect het gevolg is van de bias van blanke leraren dan van de weerstand tegen school ('oppositional culture') van de zwarte

leerlingen. Van belang was ook dat dit effect al in de kindergarten (op 5-jarige leeftijd) begon.

In hun onderzoek focusten Jackson, Barth, Powell en Lochman (2006) op de keuze van klasgenoten (c.q. sociale status). Via sociometrische nominaties werd van 1.268 9- tot 11-jarige leerlingen in 57 klassen bepaald in hoeverre de etniciteit van de leraar daarbij een rol speelde. Er werden vier dimensies onderscheiden, namelijk 'Like most', 'Like least', 'Fights' en 'Leader'. De analyses lieten zien dat etnische klassencompositie van de leerlingen van belang is voor de keuzes. Er waren effecten van de etniciteit van de leraren in relatie tot de etniciteit van de leerlingen (etnisch matchen), maar deze waren veel zwakker dan die van klassencompositie. Bovendien was er slechts voor een van de domeinen (Leader) sprake van een significant effect: zwarte leerlingen met zwarte leraren werden eerder als leider genomineerd dan zwarte leerlingen met blanke leraren.

McGrady en Reynolds (2013) gingen na of een etnische mismatch van invloed was op de leraarbeoordelingen van het gedrag van leerlingen in de klas, de geschiktheid voor Engels en voor wiskunde. Ze analyseerden de gegevens van circa 9.500 15-jarige leerlingen uit het ELS-databestand. Daarbij maakten ze qua etniciteit van de leerlingen een onderscheid naar blank, zwart, Hispanic en Aziatisch, en qua etniciteit van de leraren naar blank en niet-blank. De analyses lieten een genuanceerd beeld zien. Blanke leraren hadden iets negatievere oordelen over het gedrag en de capaciteiten van zwarte en Hispanic leerlingen dan over blanke leerlingen, en positievere oordelen over Aziatische leerlingen. Niet-blanke leraren hadden vergelijkbare oordelen over blanke leerlingen als blanke leraren. Maar de vraag was of niet-blanke leerlingen beter af waren met niet-blanke leraren. Hispanic en zwarte leerlingen kregen nooit negatievere, en slechts in een beperkt aantal gevallen positievere beoordelingen van niet-blanke leraren. Aziatische leerlingen waren echter soms wat slechter af.

Oates (2003) maakte gebruik van gegevens van 8.222 15- tot 17-jarige blanke en zwarte leerlingen uit de NELS-cohortstudie. De centrale vraag was of etnische consonantie dan wel dissonantie van leraar-leerlingcombinaties via de perceptie van de leraren van invloed is op de prestaties van de leerlingen (bv. via 'self-fulfilling prophecies' en 'perceptual biases'). Er werden objectieve en subjectieve maten geanalyseerd, namelijk de gecombineerde score op toetsen voor lezen, wiskunde, geschiedenis en science, respectievelijk het oordeel van leraren qua verwachtingen en gedrag. Uit het onderzoek bleek dat er sprake was van een negatief effect van etnische leraar-leerlingdiscongruentie op de verwachtingen en vervolgens – maar veel geringer – op de prestaties van de leerlingen, maar dan voornamelijk op die van zwarte leerlingen.

Ouazad (2007) ging na wat het effect was van een etnische leraar-leerlingmatch op leraaroordeelen lezen en rekenen, na controle voor toetsscores. Daarbij werden de leerlingen ook beoordeeld qua gedrag. Gebruik werd gemaakt van gegevens van (maximaal) 67.855 5- tot 10- jarige leerlingen uit de ECLS-K-cohortstudie. Nadat rekening was gehouden met een aantal leerling- en leraarmerken (waaronder de toetsscores) beoordeelden leraren leerlingen met dezelfde etniciteit positiever qua lezen en rekenen. Dit 'same race'-effect was echter voornamelijk het gevolg van de hogere oordelen van zwarte en Hispanic leraren (wat niet per se wil zeggen dat zwarte leraren zwarte leerlingen, respectievelijk Hispanic leraren Hispanic leerlingen overwaardeerden of dat blanke leraren zwarte leerlingen onderwaardeerden). Het gedrag van de leerlingen werd niet beïnvloed door 'same race-matching'. Er waren ook geen aanwijzingen dat het effect van 'same teacher-matching' via de beoordelingen de toetsprestaties beïnvloedde.

In het onderzoek van Pigott en Cowen (2000) werd nagegaan of 'teacher-pupil racial congruence' de beoordeling van schoolproblemen en vaardigheden beïnvloedt, ervan uitgaande dat overeenstemming het begrip versterkt en leidt tot een sterkere acceptatie van gedrag. Behalve dit positieve effect was er volgens de onderzoekers ook een negatief effect mogelijk, wat ontstaat doordat allochtone leraren zichzelf meer als autochtone middenklasse-leraren zien en de allochtone leerlingen ook vanuit dat

perspectief beoordelen. Geanalyseerd werden gegevens van 445 5- tot 10-jarige kinderen op 24 scholen. Drie dimensies stonden centraal: leraaroordelen met betrekking tot de aanpassing van de leerlingen aan de school, negatieve stereotyperingen door leraren, en de verwachtingen wat betreft de onderwijskansen van de leerlingen. Uit het onderzoek bleek dat, hoewel zwarte leerlingen door zowel zwarte als blanke leraren lager werden beoordeeld, zwarte leraren niet van autochtone leraren verschilden. Wel was het zo dat zwarte leraren alle leerlingen positiever beoordeelden. Er werden geen significantie interactie-effecten leraar x leerling etniciteit gevonden.

Shepherd (2011) liet in zijn onderzoek 57 zwarte, blanke, Aziatische en Hispanic leraren de stimuli van 40 7- tot 9-jarige zwarte, blanke en Hispanic leerlingen beoordelen. Het ging om reacties op drie vragen: over Thanksgiving, de Amerikaanse vlag, en de politie. De analyses lieten zien dat niet-blanke leerlingen minder positief werden beoordeeld dan blanke leerlingen. Maar bovendien beoordeelden niet-blanke leraren zwarte en Hispanic leerlingen nog minder positief dan blanke en Aziatische leraren.

De studie van Saft en Pianta (2001) onderzocht het effect van etnisch matchen op de beoordeling van de relatie leraar-leerling door leraren. Het databestand telde 197 voorschool- en kindergarten-leraren en 840 kinderen. Vier etnische groepen werden onderscheiden: blank, zwart, Hispanic en anders. De analyseresultaten lieten zien dat er weliswaar een significant positief effect was, al dan niet na controle van andere kenmerken, maar dat dit slechts een zeer klein effect was.

Takei en Shouse (2008) maakten voor hun analyse gebruik van gegevens uit het NELS88-cohort. Het betrof een steekproef van 6.355 13-jarige zwarte en blanke leerlingen van 410 scholen. Geanalyseerd werden leraaroordelen over het onderwijsgedrag van de leerlingen (werken onder niveau, huiswerk maken, onoplettend in de klas). Geprobeerd werd verschillen in leraaroordelen te verklaren door de etnische match (blanke en zwarte leraren x leerlingen), waarbij werd gecontroleerd voor een serie achtergrondkenmerken op leerling- en schoolniveau. De analyseresultaten lieten geen consistente effecten zien. Eerdere bevindingen op basis van dezelfde data (consistente negatieve effecten van asymmetrische leraar-leerlingrelaties) werden dus niet bevestigd, wat volgens de onderzoekers waarschijnlijk kan worden toegeschreven aan het onvoldoende controleren voor demografische kenmerken als leerlingenpopulatie qua etniciteit en verschillen in de vakken die door de leraren werden gegeven.

3.2.2.2 Compositie

De vraag die Kloppenstein (2005) aan de orde stelde, luidde in hoeverre het aandeel zwarte leraren van invloed was op de keuze van leerlingen voor het vak wiskunde. Behalve de etniciteit van de leraren werd tegelijkertijd ook de sekse van de leraren onderzocht. In totaal waren ongeveer 20.000 zwarte en 81.000 blanke 14- tot 16-jarige leerlingen bij het onderzoek betrokken. De analyses lieten zien dat wanneer het percentage zwarte wiskundeleraren op een school toenam er meer zwarte leerlingen wiskunde kozen. Dit positieve effect trad echter alleen op voor tegengestelde leraar-leerling sekse matches, met andere woorden wanneer bijvoorbeeld een zwarte manlijke leerling les kreeg van een zwarte vrouwelijke leraar.

3.3 Een samenvattend overzicht

In tabel 3.1 volgt allereerst een overzicht van de opzet van de geselecteerde studies en de daar gerapporteerde resultaten. In tabel 3.2 wordt een poging ondernomen de gerapporteerde effecten samen te vatten, met een uitsplitsing naar objectieve en subjectieve effectmaten. Er wordt een onderscheid gemaakt naar een negatief effect (-), geen effect (0), een positief effect (+) en een sterk gedifferentieerd effect (\approx), dat wil zeggen een effect dat sterk varieert per groep en/of effectmaat. Onder in tabel 3.2 staan de resultaten van een zogenoemde 'vote count', een vrij grove manier om de

resultaten van een serie studies samenvattend te kwantificeren door ze simpelweg op te tellen.

Tabel 3.1 – Overzicht kenmerken geselecteerde studies en resultaten

Auteur	Data	Sector	Aantal	Te verklaren	Verklarend	Analyse	Resultaten
Bates & Glick (2013)	1998 - 2003	k/po (k-g5)	16.701 II	problematisch gedrag	etnische match leraar-leerling	ML	positievere beoordeling bij match
Brown-Jeffy (2009)	1990	vo (g10-12)	3.392 II, 177 s	wiskundetoets	aandelen zwarte + Hispanic leraren	ML	negatief effect bij meer zwarte en Hispanic leraren
Clotfelter, Ladd & Vigdor (2005)	2000	po (g5)	3.223 I, 60.791 II	reken- en leestoets	leraar-leerling matching niveau	regr.	na controle voor leerlingkenmerken geen negatief verband meer etniciteit leraar
Cullinan & Kauffman (2005)	?	po-vo (g1-12)	769 II	emotionele stoornissen	etnische match leraar-leerling	anova	geen bias in beoordelingen
Dee (2001, 2004a,b)	1985 - 1988	k/po (k-g3)	11.600 II, 79 s	reken- en leestoets	etnische match leraar-leerling	regr.	een positief effect bij match
Dee (2005)	1988	vo (g8)	21.324 II, 1052 s	problematisch gedrag	etnische match leraar-leerling	log. regr.	een negatief effect bij niet-match (m.n. laag milieu en zuiden VS)
Downey & Pribesh (2004)	1988	k/po, vo	21.870 II	problematisch gedrag	etnische match leraar-leerling	ML	een positief effect bij match
Easton-Brooks, Lewis & Zhang (2009)	1998 - 2003	k/po (k-g5)	1.207 II	leestoets	etnische match leraar-leerling	ML	een positief effect bij match, maar dat verdwijnt na controle voor aandeel allochtone leerlingen
Eddy & Easton-Brooks (2011)	1998 - 2003	k/po (k-g5)	1.207 II	rekentoets	etnische match leraar-leerling	ML	een positief effect bij match, maar dat verdwijnt na controle voor sekse, armoede en aandeel allochtone leerlingen

Auteur	Data	Sector	Aantal	Te verklaren	Verklarend	Analyse	Resultaten
Ehrenberg, Goldhaber & Brewer (1995)	1988 - 1990	vo (g8-10)	2.500 ll	prestatietoetsen en leraaroordelen	etnische match leraar-leerling	regr.	geen effect op toetsprestaties, soms wel positief effect op leraar-oordelen
Fryer & Levitt (2004)	1988	k/po (k-g1)	ca. 12.000 ll	lees- en rekentoetsen en leraaroordeelen lezen en rekenen	etnische match leraar-leerling	regr.	negatief effect op toetsprestaties, geen effect op leraar-oordelen
Howsen & Trawick (2007)	1990	po (g3)	25.871 ll, 120 s	lees- en rekentoetsen	etnische match leraar-leerling	regr.	na controle sekse leraar en capaciteiten leerlingen geen effect
Jackson e.a. (2006)	?	po (g5)	1.268 ll, 57 k	sociometrische nominaties	etnische match leraar-leerling	ML	zwak positief effect op dimensie 'leider'
Klopfenstein (2005)	1998	vo (g9-11)	101.000 ll	keuze wiskunde	aandelen zwarte leraren	log. regr.	positief effect, echter alleen bij andere sekse
McGrady & Reynolds (2013)	2002	vo (g10)	9.500 ll	leraaroordeelen gedrag en geschiktheid	etnische match leraar-leerling	log. regr.	effecten variëren naar etniciteit
Meier (1993)	1980 - 1990	?	12 d	niveaugroepering, discipline en prestaties	aandelen Latino-leraren en -directeuren	corr., regr.	positief effect aandeel leraren; geen of negatief effect aandeel directeuren
Meier, Wrinkle & Polinard (1999)	1991 - 1996	po/vo (g3-10)	350 d	slaagkans toets	aandelen zwarte en Latino-leraren	regr.	positief effect niet-blanke leerlingen
Oates (2003)	1990 - 1992	vo (g10-12)	8.222 ll	prestatietoetsen en leraaroordeelen gedrag en geschiktheid	etnische match leraar-leerling	Lisrel	negatief effect etnische discongruentie op prestaties en oordelen van zwarte leerlingen

Auteur	Data	Sector	Aantal	Te verklaren	Verklarend	Analyse	Resultaten
Ouazad (2007)	1998 - 2002	k/po (k-g5)	67.855 ll	lees- en rekentoets en leraaroorde len lezen en rekenen en gedrag	etnische match leraar-leerling	regr.	positief effect op oordelen van zwarte en Hispanic leerlingen; geen effect op gedrag; geen indirect effect oordelen op prestaties
Pitts (2007)	1995 - 2002	?	alle d Texas	slagen eindexamen, <i>drop-out</i> , hoge toetsscore	aandelen leraren en directeuren naar etniciteit	regr.	positieve effecten leraren op twee dimensies bij zwarte en Hispanic leerlingen; geen effect directeuren
Pigott & Cowen (2000)	?	k/po (k-g5)	445 ll, 24s	leraaroorde len schoolaan passing, negatieve stereotype ringen, onderwijsv erwachting en	etnische match leraar-leerling	manova	geen effect
Shepherd (2011)	?	po (g2-3)	40 ll, 57 lr	leraaroorde len	etnische match leraar-leerling	regr.	negatief effect van niet-blanke leraren
Saft & Pianta (2001)	?	p/k	840 ll, 197 lr	leraaroorde len relatie met leerling	etnische match leraar-leerling	anova, regr.	klein positief effect
Takei & Shouse (2008)	1988	vo (g8)	6.355 ll, 410 s	leraaroorde len	etnische match leraar-leerling	ML	geen consistente effecten

p = preschool; k = kindergarten; po = primair onderwijs; vo = voortgezet onderwijs;

g = grade.

ll = leerling; lr = leraar; k = klas; s = school; d = district.

regr. = regressie-analyse; log. regr. = logistische regressie-analyse; anova = variantie-analyse;

manova = multivariate variantie-analyse; ML = multilevel-analyse; corr. = correlatie-analyse;

Lisrel = structureel model analyse.

Tabel 3.2 – Overzicht gerapporteerde effecten in de geselecteerde studies

Auteur	Sector	Objectieve effectmaat	Subjectieve effectmaat
Bates & Glick (2013)	k/po		+
Brown-Jeffy (2009)	vo	-	
Clotfelter, Ladd & Vigdor (2005)	po	0	
Cullinan & Kauffman (2005)	po-vo		0
Dee (2001, 2004a,b)	k/po	+	
Dee (2005)	vo		+
Downey & Pribesh (2004)	k/po, vo		+
Easton-Brooks, Lewis & Zhang (2009)	k/po	0	
Eddy & Easton-Brooks (2011)	k/po	0	
Ehrenberg, Goldhaber & Brewer (1995)	vo	0	+
Fryer & Levitt (2004)	k/po	-	0
Howsen & Trawick (2007)	po	0	
Jackson e.a. (2006)	po		+
Klopfenstein (2005)	vo		+
McGrady & Reynolds (2013)	vo		≈
Meier (1993)	?	l + ; d -	l + ; d -
Meier, Wrinkle & Polinard (1999)	po/vo	+	
Oates (2003)	vo	+	+
Ouazad (2007)	k/po	0	+
Pitts (2007)	?	l + ; d 0	
Pigott & Cowen (2000)	k/po		0
Shepherd (2011)	po		-
Saft & Pianta (2001)	p/k		+
Takei & Shouse (2008)	vo		≈
Totaal		3 - / 7 0 / 5 +	2 - / 3 0 / 10 + / 2 ≈

p = preschool; k = kindergarten; po = primair onderwijs; vo = voortgezet onderwijs.

l = leraren; d = directies.

- = negatief effect; 0 = geen effect; + positief effect; ≈ = sterk gedifferentieerd effect.

De tabel bevat de resultaten van in totaal 24 studies. In het onderhavige review is in 15 studies over objectieve effectmaten gerapporteerd; 3 (20%) daarvan lieten een negatief effect zien, 7 (47%) geen effect en 5 (33%) een positief effect. In 17 studies is over subjectieve effectmaten gerapporteerd; 2 (12%) lieten een negatief effect zien, 3 (18%) geen effect, 10 (59%) een positief effect, en van 2 (12%) varieerden de effecten sterk. Zoals eerder al opgemerkt is de hier gehanteerde techniek van 'vote counting' een vrij grove manier om resultaten samen te vatten; enige terughoudendheid bij de interpretatie en het trekken van harde conclusies is daarom gepast. Niet alleen wordt hier voorbijgegaan aan verschillen qua sterkte van het effect, maar ook verdwijnen allerlei voorbehouden en specificaties uit het zicht. Dit opgemerkt hebbende, lijken we te kunnen concluderen dat bij de objectieve effect-maten (m.n. toetsresultaten) het aandeel studies zonder effect qua matching overheerst. Een derde van deze studies heeft een positief effect. Bij de subjectieve effectmaten (m.n. gedragsbeoordelingen door leraren) overheerst juist het aandeel positieve effecten, dat is ruim de helft. Alle 24 studies bij elkaar genomen blijkt dat er in iets minder dan de helft (47%) van de gevallen sprake is van een positief effect; in ruim de helft van de gevallen (53%) is er sprake van een onduidelijk, geen of een negatief effect. Er lijkt daarbij geen relatie te zijn met de onderwijssector c.q. onderwijsfase.

Dat er meer positieve effecten worden gerapporteerd met betrekking tot subjectieve effectmaten, is op zich niet zo verwonderlijk. Dat zou namelijk verklaard kunnen worden uit een positieve bias van de leraren jegens leerlingen uit de eigen etnische groep, en misschien een negatieve bias jegens die van een andere groep (vgl. Cullinan & Kaufmann, 2005; Downey & Pribesh, 2004; Oates, 2003). De leraren hebben immers

directe invloed op deze oordelen. In het geval van toetsprestaties zullen de feitelijke capaciteiten van de leerlingen doorslaggevend zijn. In hoeverre deze redenering empirisch klopt en in welke mate de subjectieve oordelen de objectieve prestaties beïnvloeden, kan op basis van de hier gerapporteerde onderzoeksbevindingen niet worden gezegd. Wel kan in dit verband gewezen worden op de studie van Ouazad (2007), die constateerde dat er geen indirecte effecten uitgingen van matching via de leraaroordelen op de leerlingprestaties. Om hier hardere uitspraken over te kunnen doen, is veel meer onderzoek op dit terrein nodig.

Als we het geheel overzien, lijkt de conclusie gerechtvaardigd dat er empirisch gezien weinig overtuigend bewijs is dat een sterkere mate van etnische match, zij het in de vorm van een één-op-één koppeling van leraren aan leerlingen met eenzelfde etnische herkomst, dan wel een groter aandeel leraren met een allochtone achtergrond (in de VS niet-blanke achtergrond) op een school, tot overwegend positieve resultaten leidt. Als daarvan al kan worden gesproken, geldt dat in sterkere mate voor de subjectieve beoordelingen dan voor objectieve prestatiematen.

4 Samenvatting en conclusies

4.1 Samenvatting

Allochtone leraren zijn in het Nederlandse onderwijs zeer sterk ondervertegenwoordigd. Minder dan 4% van de leraren in het primair en voortgezet onderwijs heeft een niet-westerse achtergrond, tegen ruim het drievoudige daarvan bij de leerlingen. De kans dat er in de toekomst meer balans komt in deze situatie is bijzonder gering.

Waarom is een goede vertegenwoordiging van allochtone leraren van belang? In algemene zin wordt vaak gewezen op positieve effecten van de allochtone herkomst van leraren op de prestaties van hun allochtone leerlingen. Verondersteld wordt dat allochtone leraren als rolmodel fungeren, dat ze voor betere contacten met allochtone ouders zorgen, dat ze de cultuur van de leerlingen beter begrijpen, dat ze een belangrijke signalerende functie hebben doordat zij problemen eerder herkennen, dat ze bijdragen aan het realiseren van een goede etnisch-culturele afspiegeling van de samenleving, en dat ze (toekomstige) lerarentekorten kunnen opvullen.

In verschillende landen zijn de afgelopen decennia op basis van dergelijke argumenten initiatieven genomen om meer allochtone leraren in de klas te krijgen. Empirisch onderzoek dat die argumenten zou moeten ondersteunen, ontbreekt echter veelal. FORUM, Instituut voor Multiculturele Vraagstukken heeft het ITS van de Radboud Universiteit Nijmegen daarom gevraagd een verkennende internationale literatuurstudie te verrichten, waarin de uitkomsten van empirisch wetenschappelijk onderzoek worden geïnventariseerd en samengevat. Centraal staat de vraag: Wat is het effect van allochtone leraren op de prestaties van allochtone leerlingen?

Voor dit onderzoek zijn uiteenlopende methodieken gebruikt: literatuursearches via bibliothecaire en webbased zoeksystemen, 'snowballing' (steeds verder zoeken via de literatuurlijsten), en benadering van (internationale) onderzoekers. Bij de selectie van publicaties zijn verschillende criteria aangehouden: kwantitatief empirisch onderzoek; steekproef van voldoende omvang; (quasi-)experimentele opzet met controle voor verstoringen kenmerken; primair of voortgezet onderwijs; niet ouder dan twintig jaar; accent op Nederland, België, Duitsland, Engeland en de VS; etniciteit van de leraar en van de leerling als centrale verklarende variabelen; cognitieve en/of niet-cognitieve effecten bij leerlingen.

De literatuur-search leverde na een grondige schifting aan de hand van de selectiecriteria uiteindelijk 24 artikelen op. Zonder uitzondering betrof het studies die betrekking hadden op de VS. De meeste hadden betrekking op de kleuter- en

basisschool, maakten gebruik van grootschalige databestanden en controleerden in hun analyses voor kenmerken op het niveau van het kind en de school. In verreweg de meeste studies stond etnisch matchen centraal (overeenstemming leraar-leerling qua ethniciteit); slechts in enkele gevallen ging het om de etnische compositie van het team (het aandeel allochtone leraren op een school).

Uit analyse van de 24 publicaties via 'vote counting' (het simpelweg tellen van de effecten) bleek het volgende. In 15 studies is over objectieve effectmaten (m.n. toetsprestaties) gerapporteerd; 20% daarvan liet een negatief effect zien, 47% geen effect en 33% een positief effect. In 17 studies is over subjectieve effectmaten (leraaroordelen) gerapporteerd; 12% liet een negatief effect zien, 18% geen effect, 59% een positief effect, en van 12% varieerden de effecten sterk. Bij de objectieve effectmaten overheerst dus het aandeel studies zonder effect; bij de subjectieve effectmaten heeft juist het aandeel positieve effecten de overhand. In totaal is er over de 24 studies heen in iets minder dan de helft (47%) van de gevallen sprake van een positief effect; in wat meer dan de helft van de gevallen (53%) is er een onduidelijk, geen of een negatief effect.

4.2 Conclusies

De conclusie lijkt gerechtvaardigd dat er empirisch gezien vooralsnog weinig overtuigend bewijs is dat een sterkere mate van etnische match, zij het in de vorm van een één-op-één koppeling van leraren aan leerlingen met eenzelfde etnische herkomst, dan wel een groter aandeel leraren met een allochtone achtergrond op een school, tot overwegend positieve resultaten leidt. Als daarvan al kan worden gesproken, geldt dat in sterkere mate voor de subjectieve beoordelingen dan voor objectieve prestatiematen.

Dat er meer positieve effecten worden gerapporteerd met betrekking tot subjectieve effectmaten is op zich niet zo verwonderlijk. Dat zou namelijk verklaard kunnen worden uit een positieve bias van de leraren jegens leerlingen uit de eigen etnische groep, en misschien een negatieve bias jegens die van een andere groep. De leraren hebben immers directe invloed op deze oordelen. In het geval van toetsprestaties zullen de feitelijke capaciteiten van de leerlingen doorslaggevend zijn. Van belang is dan vervolgens hoe zich leraaroordeel en toetsprestaties tot elkaar verhouden, welke van beide uiteindelijk doorslaggevend zal zijn en wat dat voor de leerlingen betekent. In hoeverre deze redenering empirisch klopt en in welke mate de subjectieve oordelen de objectieve prestaties beïnvloeden, of omgekeerd, kan op basis van de hier gerapporteerde onderzoeksbevindingen niet worden gezegd.

Bij dit alles mag niet worden vergeten dat al het onderzoek waarover hier gerapporteerd is, betrekking heeft op de VS. Er is geen enkel onderzoek gevonden dat is uitgevoerd in Nederland of ons omringende landen. De consequentie is dat de Amerikaanse bevindingen niet zonder meer getransponeerd kunnen worden naar ons land. De situatie daar verschilt immers in velerlei opzichten van die in Nederland, bijvoorbeeld qua onderwijsstelsel, verschillen tussen scholen en verhoudingen tussen etnische groepen. Van cruciaal belang voor de onderwijsresultaten is de kwaliteit van de leraren. In Nederland zijn de verschillen daarin veel minder groot dan in de VS, waar juist op scholen met veel niet-blanke leerlingen leraren een lagere opleiding hebben gevolgd, lager gekwalificeerd zijn, een lager salaris ontvangen en vaak van baan wisselen. Los van het feit dat in Amerikaans onderzoek voor dergelijke verschillen gecontroleerd zou moeten worden, wat slechts sporadisch gebeurt, is dus onduidelijk wat de precieze waarde is van de Amerikaanse bevindingen voor Nederland. Om hardere uitspraken over de effecten van allochtone leraren te kunnen doen, is in zijn algemeenheid veel meer onderzoek op dit terrein nodig, en in het bijzonder dus in Nederland. Grootschalig onderzoek zal echter moeilijk kunnen worden georganiseerd, gewoonweg omdat er hier nauwelijks allochtone leraren zijn. Mogelijk dat kleinschaligere longitudinale studies een uitkomst bieden.

Literatuur

Aa, R. van der, Bruin, G. de, & Hulst, B. van (2007). De beroepsloopbaan van allochtone leraren in het primair en voortgezet onderwijs. Paper voor de ORD 2007 Groningen. Rotterdam: ECORYS.

*Bates, L., & Glick, J. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social Science Research*, 42, 1180-1190.

Berg, D. van den, Dijk, M. van, & Grootsholte, M. (2011). *Diversiteitsmonitor 2011*. Den Haag: SBO.

Berg, L. van den, Denessen, E., Hornstra, L., Voeten, M., & Holland, R. (2010). The implicit prejudiced attitudes of teachers: Relations to teachers expectations and the ethnic achievement gap. *American Educational Research Journal*, 47 (2) 497-527.

Bireda, S., & Chait, R. (2011). *Increasing teacher diversity. Strategies to improve the teacher workforce*. Washington, DC: Center for American Progress.

Bone, J., & Slate, J. (2011). Student ethnicity, teacher ethnicity, and student achievement: On the need for a more diverse teacher workforce. *The Journal of Multiculturalism in Education*, 7 (1) 1-22.

*Brown-Jeffy, S. (2008). School effects: Examining the race gap in mathematics achievement. *Journal of African American Studies*, 13 (4) 388-405.

Burt, J., Ortlieb, E., & Cheek, E. (2009). An investigation of the impact of racially diverse teachers on the reading skills of fourth-grade students in a one race school. *Reading Improvement*, 22 maart 2009.

CBS (2012). *Jaarboek Onderwijs in cijfers 2012*. Den Haag: CBS.

CBS (2013). *Basisonderwijs; leerlingen in het basis- en speciaal basisonderwijs*. Op 12/11/13 gedownload van <http://statline.cbs.nl/statweb/>.

Cizek, G. (1995). On the limited presence of African American teachers: An assessment of research, synthesis, and policy implications. *Review of Educational Research*, 65 (1) 78-92.

Clewell, B., & Villegas, A. (1998). Diversifying the teaching force to improve urban schools: Meeting the challenge. *Education and Urban Society*, 31 (1) 3-17.

Clewell, B., Puma, M., & McKay, S. (2005). Does it matter if my teacher looks like me? The impact of teacher race and ethnicity on student academic achievement. Paper presented at the meeting of the American Educational Research Association, Montreal.

Clotfelter, C., Ladd, H. & Vigdor, J. (2004). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, 24, 377-392.

*Clotfelter, C., Ladd, H., & Vigdor, J. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41 (4) 778-820.

Cooper, P. (2003). Effective White teachers of Black children: Teaching within a community. *Journal of Teacher Education*, 54 (5) 413-427.

*Cullinan, D., & Kaufmann, J. (2005). Do race of student and race of teacher influence ratings of emotional and behavioral problem characteristics of students with emotional disturbance? *Behavioral Disorders*, 30 (4) 393-402.

*Dee, T. (2001). Teachers, race and student achievement in a randomized experiment. Working Paper 8432. Cambridge, MA: National Bureau of Economic Research.

- *Dee, T. (2004a). Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics*, 86 (1) 195–210.
- *Dee, T. (2004b). The Race Connection. Are teachers more effective with students who share their ethnicity? *Education Next*, 4 (2) 1-7.
- *Dee, T. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review*, 95 (2) 158–165.
- *Downey, D., & Pribesh, S. (2004). When race matters: Teacher evaluations of students' classroom behavior. *Sociology of Education*, 77 (4) 267-282.
- Driessen, G. (2013). De bestrijding van onderwijsachterstanden Een review van opbrengsten en effectieve aanpakken. Nijmegen: ITS.
- Driessen, G. (2007). 'Peer group' effecten op onderwijsprestaties. Een internationaal review van effecten, verklaringen en theoretische en methodologische aspecten. Nijmegen: ITS.
- Driessen, G., & Doesborgh, J. (2004). De feminisering van het basisonderwijs. Effecten van het geslacht van de leerkrachten op de prestaties, de houding en het gedrag van de leerlingen. Nijmegen: ITS.
- Dronkers, J., & Velden, R. van der (2013). Positive but also negative effects of ethnic diversity in schools on educational achievement? In M. Windzio (ed.), *Integration and inequality in educational institutions* (pp. 71-98). London: Springer.
- *Easton-Brooks, D., Lewis, C., & Zhang, Y. (2010). Ethnic-matching: The influence of African American teachers on the reading scores of African American students. *The National Journal of Urban Education & Practice*, 3 (1) 230-243.
- ECHO/FORUM (1998). Verslag van de PABO-expertmeeting 'Kwaliteit en multiculturaliteit', 23 januari 1998.
- *Eddy, C., & Easton-Brooks, D. (2011). Ethnic matching, school placement, and mathematics achievement of African American students from kindergarten through fifth grade. *Urban Education*, 46 (6) 1280-1299.
- *Ehrenberg, R., D. Goldhaber, D., & Brewer, D. (1995). Do teacher's race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review*, 48 (3) 547–561.
- Escayg, K.-A. (2010). Diverse classrooms, diverse teachers: Representing cultural diversity in the teaching profession and implications for pre-service admissions. *Canadian Journal for New Scholars in Education*, 3 (2) 1-8.
- Ferguson, R. (2003). Teachers' perceptions and expectations and the black-white test score gap. *Urban Education*, 38 (4) 460-507.
- FORUM (2012a). Diversiteit in het onderwijs – achtergrondinformatie. Utrecht: FORUM.
- FORUM (2012b). Diversiteit in het onderwijs – maatregelen. Utrecht: FORUM.
- *Fryer, R., & Levitt, S. (2004). Understanding the Black-White test score gap in the first two years of school. *Review of Economics and Statistics*, 86 (2) 447-464.
- Grootscholte, M., & Jettinghoff, K. (2010). Diversiteitsmonitor 2010. Den Haag: SBO.

- Hanushek, E., Kahn, J., & Rivkin, S. (2004). Why public schools lose teachers. *The Journal of Human Resources*, 39 (2) 326-354.
- Hauck, D. (2011). Teachers' characteristics: Understanding the decision to refer to special education placement. *Dissertation Abstracts*, 71 (11A) 3983.
- Herweijer, L. (2009). Making up the gap. Migrant education in the Netherlands. Den Haag: SCP.
- Herweijer, L. (2011). Gemengd leren. Etnische diversiteit en leerprestaties. Den Haag: SCP.
- Hoeven-Van Doornum, A. van der (1990). Effecten van leerlingbeelden en streefniveaus op schoolloopbanen. Nijmegen: ITS.
- *Howsen, R., & Trawick, M. (2007). Teachers, race and student achievement revisited. *Applied Economics Letters*, 14, 1023-1027.
- Irvine, J., & Fenwick, L. (2009). Teachers and teaching for the new millennium: The role of HBCU's. Washington, DC: Department of Education.
- *Jackson, M., Barth, J., Powell, N., & Lochman, J. (2006). Classroom contextual effects of race on children's peer nominations. *Child Development*, 77 (5) 1325-1337.
- Jungbluth, P. (2003). De ongelijke basisschool. Etniciteit, sociaal milieu, sekse, verborgen differentiatie, segregatie, onderwijskansen en schooleffectiviteit. Nijmegen: ITS.
- Keijzer, L., Winthagen, T., & Vries, S. de (2008). Multiculturele teams in het onderwijs. Den Haag: SBO.
- *Kloppenstein, K. (2005). Beyond test scores: The impact of black teacher role models on rigorous math taking. *Contemporary Economic Policy*, 23 (3) 416-428.
- Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, 48 (2) 361-386.
- Kregten, A. van, & Moerkamp, T. (2008). Meer kleur voor de klas. Wat kan Nederland leren van buitenlandse voorbeelden. Den Haag: SBO.
- Learning Point Associates (s.a.). Quality teaching in at-risk schools: Key issues. Key issue: Recruiting minority teachers. Op 08/10/13 gedownload van www.tolerance.org.
- Marinell, W. (2009). Will mid-career entrants help avert a teacher shortage, reduce racial and gender imbalances, and fill vacancies in hard-to-staff subjects and schools? Paper annual meeting AERA, San Diego, CA, 15 april 2009.
- Marreveld, M. (red.) (2007). Meer kleur voor de klas. Special. *Didactief*, december 2007.
- Maylor, U. (2009) 'They do not relate to Black people like us': Black teachers as role models for Black pupils. *Journal of Education Policy*, 24 (1) 1-21.
- *McGrady, P., & Reynolds, J. (2013). Racial mismatch in the classroom: Beyond black-white differences. *Sociology of Education*, 86 (1) 3-17.
- *Meier, K. (1993). Latinos and representative bureaucracy. Testing the Thompson and Henderson hypotheses. *Journal of Public Administration Research and Theory: J/PART*, 3 (4) 393-414.
- *Meier, K., Wrinkle, R., & Polinard, J. (1999). Representative bureaucracy and distributional equity: Addressing the hard question. *The Journal of Politics*, 61 (4) 1025-1039.

Naman, W. (2009). Who should teach black students? Research on the role of white and black teachers in American schools. *Ethnicity and Race in a Changing World: A Review Journal*, 1 (2) 26-39.

National Collaborative on Diversity in the Teaching Force (2004). *Assessment of diversity in America's teaching force. A call to action*. Washington, DC: NEA.

*Oates, G. (2003). Teacher-student racial congruence, teacher perceptions, and test performance. *Social Science Quarterly*, 84 (3) 508-525.

*Ouazad, A. (2007). *Assessed by a teacher like me: Race, gender and subjective evaluations*. London: London School of Economics.

*Pigott, R., & Cowen, E. (2000). Teacher race, child race, racial congruence, and teacher ratings of children's school adjustment. *Journal of School Psychology*, 38 (2) 177-196.

*Pitts, D. (2007). Representative bureaucracy, ethnicity, and public schools: Examining the link between representation and performance. *Administration & Society*, 39(4) 497-526.

*Saft, E., & Pianta, R. (2001). Teachers' perceptions of their relationships with students with students: Effects of child age, gender, and ethnicity of teachers and children. *School Psychology Quarterly*, 16 (2) 125-141.

*Shepherd, M. (2011). Effects of ethnicity and gender on teachers' evaluation of students' spoken responses. *Urban Education*, 46 (5) 1011-1028.

STAMOS (2013). *Onderwijsarbeidsmarkt in cijfers*. www.stamos.nl.

*Takei, Y., & Shouse, R. (2008). Ratings in Black and White: Does racial symmetry or asymmetry influence teacher assessment of a pupil's work habits? *Social Psychology of Education*, 11 (4) 367-387.

Tenenbaum, H., & Ruck, M. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99 253-73.

Uhlenberg, J., & Brown, K. (2002). Racial gap in teachers' perceptions of the achievement gap. *Education and Urban Society*, 34 (4) 493-530.

Villegas, A., & Irvine, J. (2010). Diversifying the teaching force: An examination of major arguments. *Urban Review*, 42 (3) 175-192.

Weinberg, M. (2013). Academic impacts of highly effective teachers of color: A literature review. Op 26/09/13 gedownload van http://www.growyourownteachers.org/~index.php?option=com_content&view=article&id=130&Itemid=145.