

Opschaling van inductie en het pedagogisch-didactisch handelen van beginnende docenten in het voortgezet onderwijs¹

M. Helms-Lorenz, M. van der Pers, R. Maulana, R. van der Lans,
P. Moorer en P. H. van den Hoorn-Flens

Samenvatting

In de onderhavige longitudinale studie over een periode van 3 jaar, wordt onderzocht of het opschalen van een inductie-interventie de ontwikkeling van pedagogisch-didactische vaardigheden van beginnende docenten ten goede komt. Op de 147 deelnemende scholen is het pedagogisch-didactisch handelen van beginnende docenten periodiek gemeten met twee instrumenten: observaties door getrainde observatoren met het ICALT-instrument (N = 725 cross-sectioneel waarvan n = 420 longitudinaal), en met de leerlingvragenlijst "Mijn Leraar" (N = 650 cross-sectioneel waarvan n = 355 longitudinaal). De perceptie van begeleidingsaanbod werd gemeten door een begeleidingsmonitor af te nemen bij de beginnende docenten (N = 516 cross-sectioneel waarvan n = 299 longitudinaal). De resultaten wijzen op maatwerk in de begeleiding dat het meest intensief is voor docenten met een lager startniveau. Bij aanvang blijkt het geobserveerd pedagogisch-didactisch niveau voor een derde van de beginnende docenten beneden het verwachte startniveau te liggen. Over een periode van drie jaar zijn de effecten van begeleiding positief (1.01 en 0.24 standaarddeviatie voor respectievelijk het observatie- en leerlingperspectief). Docenten met een laag startniveau, vrouwen en eerstegraadsopgeleide docenten vertonen de meeste vorderingen.

Kernwoorden: opschaling, begeleiding, longitudinaal, ontwikkeling, inductie

wordt in balans gehouden met een aantal verantwoordingsmaatregelen (OECD, 2016). Binnen deze kaders worden docenten niet uniform opgeleid en bestaat er geen centraal curriculum voor de lerarenopleidingen. De uitdaging van de lerarenopleidingen en van stagescholen is om een doelgerichte bijdrage te leveren aan de ontwikkeling van aanstaande docenten. Sinds een aantal jaren dragen lerarenopleidingen én scholen gezamenlijk bij aan de ontwikkeling van aanstaande docenten. Deze samenwerking, met als predicaat "opleidingsschool" en sinds kort "samen opleiden", is geformaliseerd en gesubsidieerd door de overheid sinds 2009, mits de samenwerking is goedgekeurd door de Nederlands-Vlaamse Accreditatieorganisatie (NVAO). De effecten van de voortzetting van de begeleiding van recentelijk *bevoegde beginnende docenten* door lerarenopleidingen in samenwerking met scholen is in een experimentele setting getoetst in de periode 2009-2013 (Helms-Lorenz, van de Grift & Maulana, 2016). De gevonden positieve effecten hebben bijgedragen aan de opschaling van de begeleiding van beginnende docenten in Nederland. In het project Begeleiding van Startende Leraren (BSL)² worden lerarenopleidingen en scholen tijdelijk gesubsidieerd door de overheid met het doel de driejarige begeleiding van beginnende docenten binnen de scholen goed te organiseren en inhoudelijk te verbeteren. De centrale vraag in het onderzoeksproject is in hoeverre de opschaling van de samenwerking met betrekking tot de begeleiding van bevoegde docenten in de vorm van regionaal geïmplementeerde inductieprogramma's, eveneens tot de gewenste effecten in de praktijk leidt. De gewenste effecten zijn: 1) Minder uitval en 2) Acceleratie van de ontwikkeling van de pedagogisch-didactische vaardigheden van beginnende docenten. De onderhavige studie focust op het tweede doel: de ontwikkeling

1 Inleiding

Het onderwijssysteem in Nederland kent een hoge mate van decentralisatie en verleent een hoge mate van autonomie aan scholen en onderwijsinstellingen. De geboden autonomie

van de pedagogisch-didactische handelingen tijdens de eerste 3 beroepsjaren van de dataverzamelingsronde 1 bij beginnende docenten (de eerste van drie dataverzamelingsrondes van het BSL-project).

1.2 Bronnen van zorg

Beginnende docenten hebben naast aanstaande docenten een prominente plaats in de onderzoeksliteratuur en op de beleidsagenda van het Ministerie van Onderwijs, Cultuur en Wetenschap (OC&W, 2013). Het beleid en de daaruit voortkomende interventies zijn gedreven door zorgen (“concerns”). De zorgen kunnen onderverdeeld worden in 3 niveaus: op *macroniveau* stelt “The Organisation for Economic Co-operation and Development” (OECD) dat de meeste OECD-landen de volgende zorgen delen: 1. De (on)aanlokkelijkheid van het beroep van de leraar; 2. De (trage) ontwikkeling van de kennis en vaardigheden van docenten; 3. De knelpunten bij de werving, selectie en de werkgelegenheid van docenten; en 4. Het behoud (retentie) van (goede) docenten. Binnen de Nederlandse context duiden arbeidsmarktanalyses op een stijgend aantal (verwachte) onvervulde vacatures van specifieke schoolvakken (Adriaens, Fontein, Den Uijl, & De Vos, 2016; Ministerie van Onderwijs, Cultuur en Wetenschap, 2012; Van der Ploeg, Van de Pol, Vankan & Van Kerkhof, 2017). In Nederland worden visitaties gehouden die de kwaliteit van de lerarenopleidingen in kaart trachten te brengen met het doel om knelpunten te signaleren waar de opleidingen acties op kunnen plannen ter verbetering. De uitkomsten van de visitaties (die vaak een zorgwekkende werking hebben in plaats van een stimulerende werking) worden echter ook gebruikt in landelijke discussies over de kwaliteit van beginnende docenten en dragen bij aan de beleidsagenda die opleidingszorgen ook trachten te verminderen.

De concerns op *mesoniveau* betreffen die van de scholen. In opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) wordt tweejaarlijks het Personeels- en Mobiliteitsonderzoek uitgevoerd, waaruit o.a. blijkt dat 44% van het schoolmanagement tevreden is over de kwaliteit van

eerstegraads beginnende docenten en 56% tevreden is over de kwaliteit van de tweedegraads beginnende docenten³. De onderhavige studie tracht na te gaan of deze zorgen bevestigd kunnen worden door de startvaardigheden van beginnende docenten met verschillende achtergronden in kaart te brengen.

De zorgen op *microniveau* betreffen die van de beginnende docent zelf. De beginperiode van docenten gaat gepaard met behoorlijk oplopende emoties veroorzaakt door spanningen tussen idealen en realiteit (bijvoorbeeld Fuller & Bown, 1975). Uit recent onderzoek van onder andere Kelchtermans en Ballet (2002) en Pillen, Beijard en Den Brok (2012, 2013) wordt duidelijk dat soortgelijke spanningen nog steeds gelden. Shoffner (2011) toont aan dat beginnende docenten zich zorgen maken over de eigen negatieve emoties zoals boosheid en gevoelens van schuld ten aanzien van leerlingen als de lessen niet goed verlopen en hoe hiermee om te gaan. Daarnaast is gebleken dat beginnende docenten met hoge psychologische taakeisen, negatieve ervaringen met leerlingen en negatieve collegiale relaties sterke voorspellers zijn van ervaren onvredeheid en werkstress (Harmsen, Helms-Lorenz, Maulana, & van Veen, 2018a). Ander onderzoek laat zien dat beginnende docenten problemen ervaren met ordehandhaving, het motiveren van leerlingen, het omgaan met verschillen tussen leerlingen, het omgaan met individuele probleemleerlingen, beoordelen van prestaties, hoge taakbelasting, relaties met collega’s, en de gebrekkige leermiddelen (Créton & Wubbels, 1984; Houtveen, Versloot, & Groenen, 2006; De Jonge & De Muijnck, 2002; Van Kregten & Moerkamp, 2004; Veenman, 1984; Vonk, 1982). De onderhavige studie tracht de relatie tussen pedagogisch-didactisch handelen en gepercipieerd begeleidingsaanbod in kaart te brengen.

Bovenstaande zorgen op macro-, meso- en microniveau hebben tot diverse samenwerkingsverbanden geleid tussen lerarenopleidingen en scholen, in een poging om deze problemen samen op te lossen. De vraag is of deze maatregelen ook bijdragen aan de vermindering van de genoemde zorgen. Om een antwoord te vinden op de vraag of de zorgen

op macro-, meso- en microniveau gereduceerd kunnen worden met inductie-arrangementen, wordt in deze studie de invloed van inwerktrajecten op het pedagogisch-didactisch handelen (door de ogen van getrainde observatoren en door de ogen van leerlingen) tijdens de eerste 3 beroepsjaren in kaart gebracht.

2 Theoretisch kader

Voor een goede start in het beroep zijn een goede voorbereiding tijdens de opleiding (opleidingsroute en stageschool) en gepaste aanvangsbegeleiding op de school van de eerste aanstelling (inwerktraject) belangrijke randvoorwaarden.

2.1 Het belang van de opleidingsroute

In het Nederlandse voortgezet onderwijs (vo) wordt met verschillen tussen leerlingen omgegaan door leerlingen direct bij de overgang van het basisonderwijs naar het vo, of kort daarna, toe te wijzen aan verschillende onderwijssoorten, variërend van voorbereidend middelbaar beroepsonderwijs (vmbo), hoger algemeen voortgezet onderwijs (havo) tot voorbereidend wetenschappelijk onderwijs (vwo). Dit wordt ook externe differentiatie genoemd (Bosker, 2005). De docenten die deze diverse onderwijssoorten verzorgen, volgen verschillende opleidingsroutes. Het voornaamste verschil is dat docenten voor havo en vwo, universitair geschoold worden. Een van de overwegingen hierbij is dat deze leerlingen zodoende beter voorbereid kunnen worden voor een universitaire opleiding (b.v. van Rooij, 2018). Tweedegraadse lerarenopleidingen (hoger beroepsonderwijs) richten zich op aanstaande docenten voor het vmbo, de onderbouw van havo en vwo, en het praktijkonderwijs. Door de startvaardigheden van beginnende docenten in kaart te brengen draagt de studie ook bij aan inzicht in de verschillen in niveau van pedagogisch-didactisch handelen tussen deze opleidingsroutes en hoe de vaardigheid zich verder ontwikkelt.

2.2 Het belang van samen opleiden (de opleidingsschool)

Studies uitgevoerd in de Verenigde Staten heb-

ben aangetoond dat partnerschappen tussen scholen en opleidingsinstituten (zogenaamde “professional development schools”) een positieve bijdrage leveren aan onder andere de instructievaardigheden van aanstaande docenten (bijv. Castle, Fox & O’Hanlan Souder, 2006; Houston, Hollis, Clay, Ligons, & Roff, 1999; Sharpe, Lounsbury, Golden, & Deibler, 1999). In een longitudinaal onderzoek binnen de Nederlandse context hebben Helms-Lorenz, Van de Grift, Canrinus, Maulana en Van Veen (2018) aangetoond dat de pedagogisch-didactische vaardigheden van aanstaande docenten die op opleidingsscholen zijn opgeleid en werken, van een hoger niveau zijn dan de vaardigheden van docenten die op niet-opleidingsscholen zijn opgeleid en werken. Dit geldt voor zowel de observatiescores als voor de leerlingpercepties (tussen meetmomenten 1 en 2, m.a.w. het eerste jaar na de opleiding). Het verschil in het 2^e jaar bleek significant⁴ te zijn. Aanstaande docenten op opleidingsscholen rapporteerden meer leermogelijkheden en meer tevredenheid over de kwaliteit van de genoten opleiding en de voorbereiding op het beroep. Deze groep docenten profiteren van een leeromgeving waarin de lerarenopleiding en de school samenwerken bij het opleiden van aanstaande docenten en het verder professionaliseren van (beginnende) docenten. Gezien de positieve effecten van opleidingsscholen lijkt het relevant om te analyseren of opleidingsscholen ook positievere bijdragen leveren aan de startvaardigheid en de ontwikkeling van beginnende docenten in vergelijking met reguliere scholen.

2.3 Het belang van aanvangsbegeleiding

Studies naar retentie en ondersteuning van beginnende docenten zijn vanaf het jaar 2000 in de internationale onderzoeksliteratuur toegenomen (bijv. Den Brok, Wubbels & van Tartwijk, 2017; Fantilli & McDougall, 2009; Heikkinen, Jokinen, & Tynjälä, 2012; Skilbeck & Connell, 2003; Stokking, Leenders, De Jong, & Van Tartwijk, 2003;). In een evaluatiestudie van Smith en Ingersoll (2004) zijn positieve relaties van inductiearrangementen aangetoond met arbeidssatisfactie, beleving van eigen-bekwaamheid (self-efficacy), betrokkenheid bij het beroep en retentie. In een

grootschalige retentiestudie (N = 3.235 beginnende docenten) toonden zij aan dat begeleiding gepaard gaat met minder beroepsverlating en wisseling van school. Ook toonden zij aan dat beginnende docenten die participeren in een breed inductiearrangement (met meerdere activiteiten) meer geneigd zijn om in het beroep te blijven dan beginners die een enkele maatregel is aangeboden. Smith en Ingersoll (2004) vonden een verband tussen de volgende specifieke maatregelen en retentie:

1. een looptijd van minstens twee jaar, idealiter drie jaar
2. zorgen voor een opgeleide coach. Idealiter zou de coach hetzelfde vakgebied moeten hebben en dagelijks beschikbaar moeten zijn om de juiste ondersteuning te kunnen bieden
3. gezamenlijke lesvoorbereidingstijd bieden met collega's van hetzelfde vakgebied
4. de mogelijkheid voor regelmatige samenwerking met andere docenten bieden
5. de mogelijkheid aanbieden om deel uit te maken van een extern netwerk van docenten
6. de mogelijkheid bieden voor gestructureerde peer-coaching
7. de mogelijkheden bieden van lesobservaties met feedback
8. zorgdragen voor relevante workshops
9. de mogelijkheid bieden om met professionele ontwikkelingsplannen te werken
10. de mogelijkheid bieden om te werken met portfolio's/bekwaamheidsdossiers
11. veel activiteiten/maatregelen aanbieden, want het aantal componenten verhoogt de kans op retentie

Minder krachtige, doch effectieve, componenten van inductieregelingen zijn: een gereduceerd lesrooster, een vermindering van het aantal lesvoorbereidingen en extra klassikale assistentie.

Helms-Lorenz et al. (2016) toonden met een experimentele studie in Nederland aan dat de groei in vaardigheden significant versneld wordt door (vak)coaching (begeleiding in de klas) en het werken met een professioneel ontwikkelingsplan. De groei werd echter vertraagd door werkdrukreductie. Harmsen, Helms-Lorenz, Maulana, en van Veen (2018b) toonden ook aan dat werkdrukreductie van

groot belang is voor het welbevinden van beginnende docenten. Naar aanleiding van deze inzichten is de inductie-interventie van de onderhavige opschalingsstudie gewijzigd ten aanzien van de experimentele interventie (ijkingsstudie). De vernieuwde interventie focust op meer maatwerk met betrekking tot werkdrukreductie en de systematische begeleiding in de klas.

2.4 Onderzoeksvragen

De onderzoeksvragen zijn erop gericht om de relatie tussen startvaardigheden, geboden begeleiding en groei in vaardigheden te verhelderen bij de opschaling van de interventie met de inzet van genoemde wijzigingen. De eerste twee vragen zijn cross-sectioneel, de laatste is longitudinaal van aard:

1. Welk pedagogisch-didactisch startniveau hebben beginnende docenten die werkzaam zijn op scholen met een inductiearrangement en hoe verhouden zich de vaardigheden tot (geselecteerde) persoons- en schoolkenmerken?
2. Welke inductie-activiteiten worden beginnende docenten aangeboden die werkzaam zijn op scholen met een inductiearrangement en is het aanbod gerelateerd aan het startvaardigheidsniveau?
3. Wat is de gemiddelde vooruitgang van het pedagogisch-didactisch handelen van beginnende docenten en hoe verhoudt zich dit tot geselecteerde persoons- en schoolkenmerken en het begeleidingsaanbod?

3 Methode

3.1 Steekproef en respons

De werving van scholen in het voer verliep via de 9 regionale universitaire lerarenopleidingen in samenwerking met 9 tweedegraadse lerarenopleidingen. Deelname van de school betekende dat de school begeleid werd door de regionale lerarenopleidingen bij de ontwikkeling van een driejarig inwerktraject voor beginnende docenten. Alle coaches op de scholen is een observatiecursus aangeboden waarna zij gecertificeerd waren om lesobservaties uit te voeren met het ICALT-observatie-instrument. Bij aanmelding werd

de deelnemer om geïnformeerde toestemming gevraagd om de verzamelde gegevens te mogen gebruiken voor genoemde onderzoeksdoeleinden.

De dataverzameling van dataverzamelingenronde 1 van het landelijk onderzoek naar de effecten van inductie is in 2014 gestart en is afgerond in de zomer van 2017. De dataverzameling van dataverzamelingenrondes 2 en 3 eindigen respectievelijk in 2018 en 2019 en vallen buiten beschouwing van deze studie. De nulmeting van dataverzamelingenronde 1 werd uitgevoerd tussen oktober en december 2014. De beginnende docenten werden geobserveerd tijdens een les door de getrainde observatoren. De leerlingen van de beginnende docenten vulden een vragenlijst in, die de mening m.b.t. het pedagogisch-didactisch handelen van de docent in kaart brengt vanuit het perspectief van de leerling. De school bepaalde zelf of de data op digitale wijze of op papier werd aangeleverd. De beginnende docenten kregen jaarlijks een digitale vragenlijst om in te vullen om de gepercipieerde begeleiding te peilen. Jaarlijks ontving de beginnende docent (cumulatieve) feedback met betrekking tot de persoonlijke metingen.

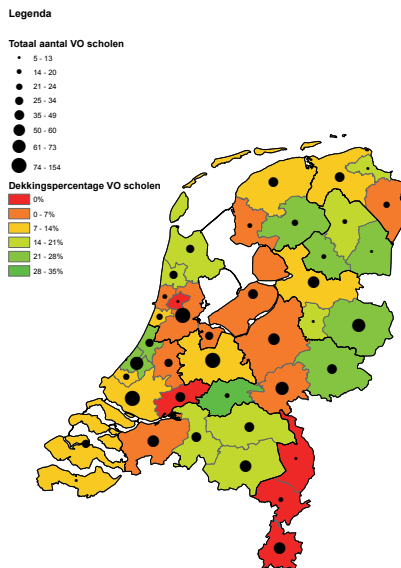
Elke school ontving voor iedere aangemelde bevoegde docent (met 0 tot 2 jaar ervaring) een aanvangsubsidiebedrag van 1000 euro, en na 3 jaar wederom 1000 euro per docent van wie alle metingen werden geleverd. Beginnende docenten kwamen in aanmerking voor loting om een prijs ter waarde van 500 euro (laptop) te winnen na drie jaar bij volledige deelname aan het onderzoek. Een deel (34%) van de beginnende docenten werkte op een van de opleidingsscholen.

Representativiteit van de steekproef

In de eerste dataverzamelingenronde werden van 147 scholen 1003 docenten aangemeld, waarvan 810 docenten (81%) actief toestemming gaven om deel te nemen aan dit onderzoek. In figuur 1 wordt de dekkinggraad van de Nederlandse vo-scholen van dataverzamelingenronde 1 uitgebeeld. De regio's Limburg en Zeeland zijn ondervetegenwoordigd, mede omdat andere overheidssubsidies lokaal worden gebruikt voor regionale impulsen van het onderwijs die ook beginnende docenten betreffen.

Van 725 docenten (89,5%) is voor de nulmeting een lesobservatie verzameld. Van deze groep hebben 420 docenten (58%, longitudinaal respons) een observatie 3 jaar later geleverd, 59,5% van deze groep is vrouw. Bij 650 docenten (80% van de docenten die bij aanvang toestemming hadden verleend) is bij een klas de leerlingvragenlijst afgenomen tijdens de nulmeting. Voor 350 van deze docenten (54 % longitudinaal respons) is drie jaar later nogmaals een leerlingvragenlijst aangeleverd. Van deze groep is 62% vrouw.

Voor de interpretatie van de resultaten van de laatste twee onderzoeksvragen is het van belang om de representativiteit van de nulmeting van de longitudinale data te vergelijken met de nulmeting van de cross-sectionele data m.b.v. t-toetsen. De nulmeting van de observatiescores van beginnende docenten met een eindmeting wijken niet significant af van de nulmeting van beginnende docenten zonder eindmeting ($t(723) = .897, p = 0.370$, Tabel 1). De groep docenten die geen eindmeting heeft, scoort lager op het domein "creëren van een veilig en stimulerend leerklimaat" vergeleken met de groep die wel een eindmeting heeft ($t(723) = -2.362, p = 0.019$, Tabel 1). Met betrekking tot de leerlingvragenlijst zijn er geen verschillen tussen de twee groepen docenten



Figuur 1
Dekkingsgraad deelnemende scholen, $N = 147$

die wel en geen eindmeting geleverd hebben ($t(648) = -1.433, p = 0.152$, Tabel 1).

De groep beginnende docenten met een nul- en een eindmeting krijgt een vergelijkbare hoeveelheid begeleidingsactiviteiten aangeboden vergeleken met beginnende docenten die niet alle metingen hebben aan-geleverd (zie Tabel 1).

We kunnen daarom vaststellen dat de

longitudinale data representatief is, met een kanttekening m.b.t. “veilig leerklimaat” vanuit het perspectief van de observatoren.

3.2 De interventie: ‘Begeleiding startende-leraren’

Het kader voor het inductiearrangement van het BSL-project, dat in alle regio’s van Nederland gebruikt wordt, is in overeenstemming

Tabel 1

Representativiteit steekproef met en zonder eindmeting vergeleken, lesobservaties ($n = 42$ vs. $n = 305$), leerlingvragenlijst ($n = 355$ vs. $n = 295$), monitor ($n = 299$ vs. $n = 203$)

	t-toets ICALT lesobservaties M0				t-toets Leerlingvragenlijst M0				t-toets Begeleidingsmonitor			
	df	t	p	effect- grootte	df	t	p	effect- grootte	df	t	p	effect- grootte
ICALT Rasch scores												
Totaal	723	-0.897	0.370	0.06	648	-1.433	0.152	0.11				
Eerste kwartiel (laagst)	155	-0.742	0.459	0.12								
Tweede kwartiel	170	-0.652	0.515	0.11								
Derde kwartiel	208	0.548	0.584	-0.06								
Vierde kwartiel (hoogst)	184	0.224	0.823	-0.03								
ICALT domein-scores												
Veilig en stimulerend leerklimaat	723	-2.362	0.019	0.19	648	-1.567	0.118	0.12				
Efficiënt klasmanagement	723	-1.540	0.124	0.13	648	-1.432	0.153	0.12				
Heldere instructie	723	-0.712	0.477	0.10	648	-1.135	0.257	0.08				
Intensieve en activerende les	723	-0.452	0.652	0.04	648	-1.763	0.078	0.12				
Afstemmen instructie en verwerking op verschillen	723	-1.719	0.086	0.12	648	-1.567	0.118	0.12				
Aanleren van meta-cognitieve vaardigheden	723	-0.672	0.502	0.06	648	-1.432	0.153	0.12				
Monitorscores												
Werkdruk reductie									500	1.389	0.166	0.13
School enculturatie									500	0.379	0.705	0.03
Begeleiding in de klas									500	1.346	0.179	0.12
Professionele ontwikkeling									500	0.379	0.705	0.03

met de zelfdeterminatietheorie van Ryan en Deci (2002). Inzichten uit deze theorie, en uit de beschreven literatuur, worden gebruikt om een werksituatie voor de beginnende docent te creëren die duurzame motivatie (om zich te verbeteren) mogelijk maakt. Daarbij staan drie psychologische basisbehoeften centraal: 1) het vergroten van de *autonomie* van de beginnende docent, 2) het bevorderen van zijn/haar *competentie* en 3) het bevorderen van de *verbondenheid* met collega's en leerlingen. Om aan de basisbehoeften recht te doen, worden effectief gebleken maatregelen en/of activiteiten (zie 2.1 en 2.2) gebundeld⁵ in "4 pijlers van driejarige inductie":

Het BSL-kader biedt:

1. Leerzaam werk met een daarbij *passende* werkdruk
2. Aandacht voor de schoolcultuur en het schoolbeleid
3. Begeleiding d.m.v. professionele ontwikkelingsplannen
4. Ondersteuning in de klas bieden in de vorm van feedback en vakdidactische coaching naar aanleiding van observaties. Samen lessen voorbereiden en nabespreken, lessen van ervaren collega's observeren en nabespreken

Het inductiearrangement heeft enerzijds betrekking op de *werksituatie* waarbinnen de beginnende docent functioneert en voor zover deze een rol speelt in het professionaliseringsproces en op de *systematiek van de begeleidingssituatie* welke gericht is op formeel, intentioneel leren. Deze systematiek sluit echter niet uit dat de daaruit voortvloeiende begeleiding, zowel wat betreft vorm als inhoud, van school tot school en per regio kan verschillen. Het inductiearrangement heeft echter op alle participerende scholen de volgende elementen.

Elementen die de draaglast verminderen

Deze elementen zijn gericht op vermindering van de (door de beginnende docent ervaren) taakeisen en zodoende op de draaglast van de beginnende docent. De verwachting is dat de realisatie van deze elementen van het arrangement een positieve uitwerking zal hebben op de situatiebeleving en de ervaren stress van de beginnende docent. De elementen in het

arrangement die hiervoor moeten zorgen zijn onder andere: geen extra taken, vrijstelling van taakuren, minder lessen, geen moeilijke klassen, gunstige roosters, rooster waarbij rekening gehouden wordt met het rooster van de mentor/vak-coach, een looptijd van het inductiearrangement van 3 jaar en een geleidelijke opvoering van de werkbelasting naar het normale eindniveau. Deze maatregelen zijn geen doel op zich maar worden ingezet waar nodig.

Elementen die de draagkracht vergroten

Tijdens de drie inductiejaren verloopt de ontwikkeling van de beginnende docent fasegevijs. Daarom wordt aan het begin van het arrangement meer aandacht geschonken aan problemen die zich de eerste maanden van de beroepsuitoefening voordoen. Later in het arrangement verschuift de focus naar de problemen die zich daarna voordoen. Met andere woorden, tijdens het arrangement verandert de inhoud van de begeleiding. Deze verandering wordt vooral bepaald door de ontwikkeling van de beginnende docent. Uiteindelijk is de inhoud van de begeleiding gericht op het vergroten van de autonomie van de beginnende docent. De verwachting is dat de begeleiding binnen het arrangement een positieve uitwerking heeft op het gevoel van bekwaamheid van de beginnende docent en dat daarmee de draagkracht zal toenemen.

Alle universitaire lerarenopleidingen en sommige partner hbo-lerarenopleidingen waren betrokken bij de ontwikkeling van inductiearrangementen in de eigen regio. Deze aanpak past naadloos bij het autonome karakter van het Nederlandse onderwijssysteem (Helms-Lorenz, et al., 2017). Alle scholen die deelnamen aan dit onderzoek hebben een inductiearrangement ontwikkeld en toegepast volgens een gemeenschappelijk kader, dat gebaseerd is op onderzoeksbevindingen. Er is ingezet op minder werkdrukreductie en meer op begeleiding in de klas, voortkomend uit de studie van (Helms-Lorenz, et al., 2016) beschreven in de theoretische inleiding.

3.3 Instrumenten

De begeleidingsmonitor

Om te achterhalen welke begeleidingsmaatre-

gelen op school aangeboden worden aan beginnende docenten wordt een vragenlijst jaarlijks uitgezet bij alle beginnende docenten die deelnemen aan dit onderzoek. De beginnende docent wordt met deze vragenlijst verzocht aan te geven welke specifieke maatregelen ze, m.b.t. 4 inductiepijlers werkdrukreductie, school-enculturatie, begeleiding in de klas en planmatige professionele ontwikkeling, wordt aangeboden. De antwoordmogelijkheden zijn: ja, nee. Deze vragenlijst meet of de beginnende docent zich bewust is van het door de school voorgenumen begeleidingsaanbod. Voorbeelden van maatregelen per pijler zijn opgenomen in Appendix A.

De vragenlijst is in eerder experimenteel onderzoek gebruikt om de implementatie van inductie-interventie te monitoren (o.a. Helms-Lorenz, Slof, Vermue, & Canrinus, 2012). Dit instrument is een inventarisatie van maatregelen en niet een instrument dat psychologische constructen meet. Per inductiepijler worden de maatregelen bij elkaar opgeteld. Dat is in overeenstemming met Smith & Ingersoll (2004) die geconstateerd hebben dat er een positieve relatie is tussen het aantal aangeboden maatregelen en retentie.

Het ICALT-observatie instrument

Om het pedagogisch-didactisch handelen in kaart te brengen is gekozen voor een instrument waarvan eerder de validiteit en de betrouwbaarheid is vastgesteld, het International Comparative Analysis of Learning and Teaching (ICALT) observatie-instrument. Voor de totstandkoming van dit instrument hebben Van de Grift en Lam (1998) en Van de Grift (2007) observatieschalen ontwikkeld om de pedagogisch-didactische vaardigheden van docenten te meten. Het instrument bestaat uit 32 items waarmee zes gedragsdomeinen worden geobserveerd: 1) veilig en stimulerend leerklimaat, 2) efficiënt klasmanagement, 3) heldere instructie, 4) intensieve en activerende les, 5) afstemmen instructie en verwerking op verschillen, 6) aanleren van leerstrategieën. Deze items zijn ontleend aan overzichtsstudies naar de effectiviteit van het handelen van docenten op de prestaties van leerlingen (Cotton, 1995; Creemers, 1994; Ellis & Worthington, 1994; Hattie, 2009; Le-

vine & Lezotte, 1990, 1995; Purkey & Smith, 1983; Sammons, Hillam, & Mortimore, 1995; Scheerens, 1992; Scheerens & Bosker, 1997; Walberg & Haertel, 1992). Alle items worden gescoord op vier antwoordmodaliteiten ('zwak'; 'meer zwak dan sterk'; 'meer sterk dan zwak'; 'sterk'). Zie Appendix B voor de items per domein.

Observatietraining

Ervaren docenten, van wie een groot deel eerder een coachcursus heeft gevolgd, werden geselecteerd om deel te nemen aan een observatietraining. De vier uur durende training beoogt het observatie-instrument en het scoren uit te leggen en de docenten te laten oefenen met behulp van twee opgenomen lessen. Tijdens het bekijken van de eerste les wordt het observatie-instrument gescoord. Na afloop wordt het percentage overeenstemming tussen de observatoren geschat en worden scores vergeleken met een landelijk norm. Deze informatie dient als input voor het nagesprek waarin de verschillen tussen observatoren worden uitgelicht en bediscussieerd, evenals de verschillen met een eerder vastgestelde normgroep. Na het nagesprek (30-45 min) wordt een tweede film bekeken en gescoord. Het doel is om consensus van 70% of hoger op de tweede les te krijgen. Bij alle observatoren is aan dit criterium voldaan.

De leerlingvragenlijst "Mijn leraar..."

Als aanvulling op de ICALT-observaties is een leerlingvragenlijst afgenomen om de ontwikkeling van de pedagogisch-didactische vaardigheden van de beginnende docenten vanuit het leerlingperspectief in kaart te brengen. Het instrument berust op dezelfde theoretische basis als het ICALT-observatie-instrument en is betrouwbaar en valide gebleken (Maulana, Helms-Lorenz, & Van de Grift 2015a; Van der Lans, Van de Grift, & Van Veen, 2018b; Van der Lans & Maulana, 2018). De vragenlijst bevat de leerlingenpercepties van pedagogisch-didactisch handelen gemeten met 41 items. Enkele voorbeelden zijn 'Mijn leraar helpt mij als ik iets niet weet', 'Mijn leraar maakt duidelijk wat de bedoeling is van de les', 'Mijn leraar geeft mij

zelfvertrouwen bij moeilijke taken', 'Mijn leraar laat mij uitleggen hoe ik de opdracht heb aangepakt' (zie Appendix C).

Raschscores

De gemiddelde gedragsdomeinscores (klassieke test theorie) zijn berekend en aangevuld met IRT Raschscores omdat deze aanvullende, nuttige informatie verschaffen over het vaardigheidsniveau over verschillende domeinen heen. Indien een instrument aan de vereisten (assumpties) van een item respons model voldoet, dan kunnen items (onafhankelijk van de steekproef) beschouwd worden als een indicator voor de mate waarin een eigenschap aanwezig is. De latente eigenschap is in dit geval "effectief gedrag in de klas". De hypothese die getoetst wordt met een IRT-model is of, voor welk vaardigheidsniveau dan ook, de kans om item j (moeilijker item) juist te beantwoorden toeneemt wanneer de vaardigheid van docenten toeneemt, doch voor welk vaardigheidsniveau dan ook kleiner is dan de kans op een juist antwoord op i (makkelijker item). Hieruit volgt de statistisch te toetsen voorspelling dat de verwachte proportie juiste antwoorden op item j kleiner is dan op item i in een willekeurige steekproef van personen (Van Til & Van Boxtel, 2015). Onderzoek met het ICALT-observatie-instrument en de "Mijn leraar" leerlingvragenlijst heeft aangetoond dat beide instrumenten aan de eisen van de Rasch model voldoen. Dit itempatroon is als eerste getoetst in het primair onderwijs (Kyriakides, Creemers, & Antoniou, 2009; Van de Grift, 2007; Van de Grift, Van der Wal, & Torenbeek, 2011) en later in het voortgezet onderwijs bij aanstaande docenten door Van de Grift, Helms-Lorenz en Maulana (2014) en voor ervaren docenten door Van der Lans, Van de Grift, en van Veen (2018b). In het huidige onderzoek wordt gebruik gemaakt van Raschscores (Warm, 1989) om het startniveau en de longitudinale ontwikkeling weer te geven gebaseerd op Van de Grift et al. (2014) en Maulana, Helms-Lorenz & Van de Grift (2015b). De betrouwbaarheid van de Raschschaal is goed voor beide instrumenten in deze steekproef. Voor de observatie Raschschaal is Cronbach's α gelijk aan 0.89 en voor de leerlingvragenlijst 0.91.

Graad en opleidingsschool

Het opleidingstype van docenten werd vastgelegd met een achtergrondvariabele in de begeleidingsmonitor, die op verschillende momenten werd afgenomen. De opleidingsschoolstatus (2009) van de deelnemende scholen werd nagevraagd bij de VO-Raad. In 2015 hebben meer scholen deze status gekregen, maar omdat dit na de start van de dataverzameling is toegekend is de status van 2009 aangehouden.

3.4 Data-analyse

Voor onderzoeksvraag 1 zijn de observatie- en leerlingvragenlijstdomeinscores en de Raschscores berekend. De Raschscores zijn vergeleken met een opgelegde norm om vast te stellen of het beginniveau in lijn is met het verwacht niveau. Daarnaast zijn de vaardigheidsscores van docenten met diverse persoons- en schoolkenmerken met elkaar vergeleken m.b.v. t-toetsen en effectmaten (verschil tussen gemiddelden gedeeld door de gepoolde standaarddeviatie).

Voor onderzoeksvraag 2 zijn de beschrijvende statistieken van de ervaren begeleiding vergeleken met de gemiddelden van de vergelijkbare, kleinschaliger experimentele studie van 2009-2013 van Helms-Lorenz et al. (2016) (hierna ook wel "de ijkingsstudie" genoemd) om na te gaan in hoeverre de opschaling tot de gewenste implementatieperceptie van begeleiding bij de beginnende docenten heeft geleid. Om de relatie tussen ervaren begeleiding en vaardigheidsniveau in kaart te brengen zijn de observatie Raschscores van de nulmeting in kwartielen gesplitst, waarmee de docenten in 4 'vaardigheidsgroepen' zijn verdeeld. Middels t-toetsen en effectmaten werd het gepercipieerde begeleidingsaanbod per vaardigheidsgroep vergeleken.

Om de verandering over tijd vast te stellen is van de longitudinale data de eindmeting vergeleken met de nulmeting (onderzoeksvraag 3) met behulp van t-toetsen en effectgroottes. Dit werd gedaan voor de hele groep maar ook voor groepen docenten met verschillende kenmerken. Om de relatie tussen het ervaren begeleidingsaanbod en de groei in pedagogisch-didactisch handelen (onderzoeksvraag 4) te onderzoeken is het begelei-

dingsaanbod van de 4 vaardigheidsgroepen met elkaar vergeleken met behulp van t-toetsen en effectmaten.

4 Resultaten

4.1 Onderzoeksvraag 1

Welk pedagogisch-didactisch startniveau hebben beginnende docenten die werkzaam zijn op scholen met een inductiearrangement en hoe verhouden zich de vaardigheden tot (geselecteerde) persoons- en schoolkenmerken?

Door de ogen van de observatoren creëert de gemiddelde beginnende docent een goed en veilig leerklimaat ($M = 3.27$, $SD = 0.56$), heeft deze een goed klassenmanagement ($M = 3.08$, $SD = 0.60$), zijn de instructies voldoende helder en duidelijk ($M = 2.93$, $SD = 0.56$) en zijn de lessen voldoende activerend voor de leerlingen ($M = 2.49$, $SD = 0.59$) (Tabel 5). Het “afstemmen op verschillen” ($M = 1.87$, $SD = 0.62$) en het aanleren van leerstrategieën ($M = 1.95$, $SD = 0.65$) kan gemiddeld genomen nog niet als voldoende bestempeld worden.

Door de ogen van de leerlingen creëert de gemiddelde beginnende docent een goed en veilig leerklimaat ($M = 3.32$, $SD = 0.32$), klassenmanagement ($M = 3.12$, $SD = 0.36$), heldere instructie ($M = 3.36$, $SD = 0.34$) en zijn de lessen activerend ($M = 3.06$, $SD =$

0.36). Het afstemmen op verschillen ($M = 2.88$, $SD = 0.37$) en het aanleren van leerstrategieën ($M = 2.80$, $SD = 0.36$) percipiëren de leerlingen als zijnde ‘voldoende’.

Het gemiddelde startniveau (nulmeting) van docenten met verschillende persoons- en schoolkenmerken worden in Tabel 2 gepresenteerd. De lesobservatie Raschscores zijn iets hoger in niet-opleidingsscholen maar het verschil is niet significant ($t(723) = 1.781$, $p = 0.075$; effectgrootte = -0.15). Tussen mannen en vrouwen en tussen eerste en tweedegraadsopgeleide docenten zijn er geen verschillen gevonden. De leerlingvragenlijst Raschscores zijn ook iets hoger voor docenten op niet-opleidingsscholen en het verschil is significant ($t(648) = 3.176$, $p = 0.002$; effectgrootte = -0.28). Mannen krijgen iets hogere scores van leerlingen hoewel het verschil niet significant is ($p = 0.073$). De leerlingsscores van tweedegraadsopgeleide docenten zijn hoger dan die van eerstegraadsopgeleide docenten ($t(630) = 2.545$, $p = 0.011$; effectgrootte = -0.20).

4.2 Onderzoeksvraag 2

Welke inductie-activiteiten worden beginnende docenten aangeboden die werkzaam zijn op scholen met een inductiearrangement en is het aanbod gerelateerd aan het startvaardigheidsniveau?

De docenten ontvangen in het eerste jaar van het inductiearrangement 37% van de voorgestelde werkdruk verlagende activitei-

Tabel 2

Niveau pedagogisch-didactisch handelen bij aanvang (M0)

Rasch scores ICALT-observaties (N = 725) en leerlingvragenlijst (N = 650)

	ICALT lesobservaties M0			Leerlingvragenlijst M0		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Rasch score totaal	725	0.42	1.33	650	2.07	0.98
Type school						
Opleidingsschool	192	0.27	1.26	160	1.86	1.05
Geen opleidingsschool	533	0.47	1.36	490	2.14	0.94
Geslacht docent						
Man	299	0.40	1.31	262	2.15	0.95
Vrouw	426	0.44	1.35	388	2.01	0.99
Graad docent						
Eerste graad	296	0.41	1.33	267	1.95	1.01
Tweede graad	404	0.44	1.34	365	2.15	0.95

ten, 60% activiteiten rond schoolenculturatie en 46% en 57% van de voorgestelde activiteiten betreft begeleiding in de klas en systematische professionele ontwikkeling.

In Tabel 3 worden de monitoruitkomsten van de ijkingsstudie en de huidige studie gepresenteerd. Hieruit blijkt dat er in de huidige studie, significant minder werkdrukreductie ($t(595) = 9.601, p = 0.000$; effectgrootte = -1.15) en minder schoolenculturatie maatregelen ($t(595) = 5.512, p = 0.000$; effectgrootte = -0.66) aangeboden worden dan in de experimentele conditie van de ijkingsstudie. In vergelijking met de ijkstudie is er gemiddeld even veel begeleiding in de klas ($t(595) = -1.071, p = 0.285$; effectgrootte = 0.13) en zijn evenveel activiteiten ten behoeven van systematische professionele ontwikkeling ($t(595) = 0.827, p = 0.409$; effectgrootte = -0.10) aangeboden.

Om te bepalen of er een relatie is tussen begeleiding en vaardigheidsniveau is de

groep in 4 vaardigheidsniveaus verdeeld door gebruik te maken van de observatie Rasch-score.

4.3 Onderzoeksvraag 3

Wat is de gemiddelde vooruitgang van het pedagogisch-didactisch handelen van beginnende docenten en hoe verhoudt zich dit tot geselecteerde persoons- en schoolkenmerken en het begeleidingsaanbod?

Om te bepalen of het algemeen startniveau “voldoende” is, is het van belang om de complexiteit van gedragsdomeinen in acht te nemen. Welke vaardigheden kunnen redelijkerwijs verwacht worden van een starter? Voor deze studie geldt de veronderstelling dat een docent pedagogisch-didactisch bekwaam is als hij/zij in staat is een veilig en stimulerend klimaat te creëren, een les adequaat kan organiseren, voldoende duidelijk instructie kan geven en enigszins activerende lessen kan

Tabel 3

Monitorscores- huidige studie (N = 516) en ijkingsstudie 2009-2013 (N = 81)

	Ijkingsstudie		Huidige studie		Ijkingsstudie vs. huidige studie t-toets en effectgroottes			
	M	SD	M	SD	t	df	p	effectgrootte
Werkdruk reductie	11.70	3.12	7.10	4.13	9.601	595	0.000	-1.15
School enculturatie	4.94	1.41	3.65	2.03	5.512	595	0.000	-0.66
Begeleiding in de klas	4.22	2.42	4.57	2.78	1.070	595	0.285	0.13
Professionele ontwikkeling	5.38	2.35	5.10	2.90	0.827	595	0.409	-0.10

Tabel 4

Gepercipieerde ontvangen begeleiding in aanvangsjaar, per vaardigheidsniveau pedagogisch-didactisch handelen

	Werkdruk reductie		School enculturatie		Begeleiding in de klas		Professionele ontwikkeling	
	M	SD	M	SD	M	SD	M	SD
Totaal (N=516)	7.10	4.13	3.65	2.03	4.57	2.78	5.10	2.90
ICALT Rasch - Eerste kwartiel (laagst) (n=107)	7.39	4.28	3.61	1.98	4.83	2.67	5.28	2.79
ICALT Rasch – Tweede kwartiel (n=72)	6.83	3.93	3.62	2.08	4.61	2.88	4.78	2.98
ICALT Rasch - Derde kwartiel (n=115)	7.17	4.34	3.68	2.05	4.50	2.76	5.24	2.93
ICALT Rasch - Vierde kwartiel (hoogst) (n=121)	7.01	4.14	3.51	2.05	4.42	2.93	5.17	2.98

Tabel 5

Pedagogisch-didactisch handelen van recentelijk bevoegde leraren in het vo; cumulatieve percentages totale groep M0 (N = 725) en longitudinale groep M0 en M3 (n = 420)

Domein	De leraar/lerares...	Warm's θ	% M0 N=725	% M0 n=420	% M3 n=420
Klimaat	...toont in gedrag en taalgebruik respect voor leerlingen	-6.73	0.1		
Instructie	...geeft duidelijke uitleg van de leerstof	-4.43	0.4		
Klimaat	...zorgt voor een ontspannen sfeer	-3.31	0.7	0.5	
Klimaat	...ondersteunt het zelfvertrouwen van leerlingen	-2.73	1.2	1.0	
Instructie	...geeft feedback aan de leerlingen	-2.33	2.5	2.1	
Lesorganisatie	...zorgt voor een ordelijk verloop van de les	-2.01	4.1	3.8	0.2
Klimaat	...zorgt voor wederzijds respect	-1.73	5.5	5.0	
Lesorganisatie	...zorgt voor doelmatig klassenmanagement	-1.49	7.6	6.2	0.7
Instructie	...geeft goed gestructureerd les	-1.27	10.2	9.0	
Instructie	...bevordert dat leerlingen hun best doen	-1.07	13.0	11.9	1.4
Lesorganisatie	...gebruikt leertijd efficiënt	-.87	15.3	14.3	2.1
Lesorganisatie	...gaat tijdens de verwerking na of leerlingen de opdrachten op een juiste manier uitvoeren	-.69	18.5	17.6	3.1
Activeren	...stelt vragen die leerlingen tot denken aanzetten	-.51	21.7	19.8	3.6
Instructie	...betreft alle leerlingen bij de les	-.34	26.9	25.5	5.2
Activeren	...verduidelijkt bij de aanvang van de les de lesdoelen	-.17	33.4	31.2	7.4
Activeren	...hanteert werkvormen die leerlingen activeren	.00	38.2	36.7	8.8
Instructie	...geeft duidelijke uitleg van het gebruik van didactische hulpmiddelen en opdrachten	.17	45.4	44.8	10.5
Activeren	...stimuleert leerlingen om over oplossingen na te denken	.34	51.2	50.7	13.3
Instructie	...gaat tijdens instructie na of leerlingen de leerstof hebben begrepen	.51	57.5	57.6	16.4
Activeren	...stimuleert het zelfvertrouwen van zwakke leerlingen	.68	61.8	62.4	21
Activeren	...zorgt voor interactieve instructie	.86	67.2	66.7	26.2
Leerstrategieën	...moedigt kritisch denken van leerlingen aan	1.04	74.3	74.3	34.3
Leerstrategieën	...bevordert het toepassen van het geleerde	1.24	80.4	79.5	41.2
Activeren	...laat leerlingen hardop denken	1.44	84.0	83.6	48.8
Leerstrategieën	...leert leerlingen hoe zij complexe problemen kunnen vereenvoudigen	1.65	87.9	87.4	54.3
Differentiatie	...gaat na of de lesdoelen werden bereikt	1.88	90.2	90.2	61.4
Leerstrategieën	...stimuleert het gebruik van controle activiteiten	2.13	92.6	92.4	66.4
Leerstrategieën	...leert leerlingen oplossingen te checken	2.42	94.8	94.5	72.6
Differentiatie	...stemt de instructie af op relevante verschillen tussen leerlingen	2.75	97.0	97.1	79.5
Leerstrategieën	...vraagt leerlingen na te denken over strategieën bij de aanpak	3.17	97.9	98.3	86.2
Differentiatie	...biedt zwakke leerlingen extra leer- en instructietijd	3.76	98.9	99.3	93.1
Differentiatie	...stemt verwerking van de leerstof af op relevante verschillen tussen leerlingen	4.93	100.0	100	100

verzorgen. De lat kan uiteraard lager of hoger gelegd worden. In Tabel 5 is te zien dat een cumulatief percentage van 33,4% onder het gewenst niveau is bij de nulmeting. Bij de eindmeting is dit cumulatieve percentage gedaald naar 7,4%. Van de docenten die onder het verwachte niveau scoren bij de nulmeting, is de verdeling tussen mannen en vrouwen gelijk en is de verdeling over docenten met een eerste- of tweedegraadse bevoegdheid eveneens gelijk. Docenten die op opleidingscholen werken scoren iets vaker laag, vergeleken met docenten die op niet-opleidingscholen werken (38% tegenover 34%).

In Tabel 6 worden de observatie Rasch-scores aan het begin en eind van het project gepresenteerd en met elkaar vergeleken. Dit betreft een tussenliggende periode van drie jaar. In het algemeen wordt een positieve ontwikkeling over de tijd gevonden. Voor de totale groep zijn de effecten groot (effectgrootte = 1.02). De vooruitgang is het grootst bij de laagste vaardigheidsgroep en het laagst bij de hoogste vaardigheidsgroep (plafondefect), waarbij alleen bij de hoogste groep de vorderingen niet significant zijn.

Als de groep verdeeld wordt naar school- en persoonskenmerken dan zijn de gemeten vor-

deringen op de opleidingsscholen even groot als op de niet-opleidingsscholen, en dat geldt ook voor mannen en vrouwen. De vorderingen van eerstegraadsopgeleide docenten zijn in zeer geringe mate hoger vergeleken met tweedegraadsopgeleide docenten (eerstegraadsopgeleide docenten scoren bij aanvang significant lager; $p = 0.0224$ en bij de eindmeting gelijk aan tweedegraadsopgeleide docenten).

In Tabel 7 worden de leerling Rasch-scores aan het begin en eind van het project gepresenteerd en met elkaar vergeleken. Vanuit het perspectief van de leerlingen groeien de vaardigheden van docenten (effectgrootte = 0.24; Tabel 7). Het kleinere effect (in vergelijking met de observaties) is wellicht te wijten aan het feit dat de leerlingen bij aanvang het handelen al hoger scoren. Additioneel is de groep docenten in 4 vaardigheidsgroepen (op basis van de observatiescores) verdeeld. Hieruit blijkt dat de vorderingen van zowel de laagste en de hoogste vaardigheidsgroepen door de ogen van de leerlingen het hoogst en significant zijn. De scores van de tussenliggende groepen aan het begin en het eind verschillen niet significant van elkaar.

Als de groep verdeeld wordt naar school- en persoonskenmerken, dan zijn de gemeten ver-

Tabel 6

Rasch observatiescores, nulmeting (M0) en eindmeting (M3), $n = 420$

	<i>n</i>	Nulmeting M0		Eindmeting M3		M0 vs. M3 t-test en effectgroottes			
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>	effectgrootte
ICALT Rasch-scores									
Totaal	420	0.46	1.26	1.82	1.41	838	14.38	0.000	1.02
Eerste kwartiel (laagst)	83	-1.26	0.66	1.44	1.52	164	14.90	0.000	2.31
Tweede kwartiel	105	-0.06	0.20	1.68	1.36	208	13.02	0.000	1.80
Derde kwartiel	124	0.69	0.26	1.85	1.28	246	9.94	0.000	1.26
Vierde kwartiel (hoogst)	108	2.01	0.83	2.21	1.45	214	1.20	0.231	0.16
Type school									
Opleidingsschool	145	0.45	1.08	1.68	1.38	288	8.44	0.000	0.99
Geen opleidingsschool	275	0.46	1.35	1.90	1.43	548	12.14	0.000	1.04
Geslacht docent									
Man	170	0.42	1.21	1.63	1.38	338	8.61	0.000	0.93
Vrouw	250	0.48	1.29	1.95	1.42	498	12.07	0.000	1.08
Graad docent									
Eerste graad	186	0.41	1.20	1.81	1.44	370	10.22	0.000	1.06
Tweede graad	219	0.52	1.31	1.87	1.41	436	10.39	0.000	0.99

schillen tussen de nul- en de eindmeting op de opleidingsscholen groter dan op de niet-opleidingsscholen, hoewel de verschillen in beide contexten significant zijn (de startwaarden verschillen niet significant van elkaar ($p = 0.087$)). De gemeten verschillen tussen de nulmeting en de eindmeting is gering bij de mannen (niet significant en effectgrootte = 0.07) en significant bij de vrouwen ($t(436) = 3.613$, $p = 0.000$; effectgrootte = 0.35). Dit zou verklaard kunnen worden door een significant verschil bij de nulmeting ($t(353) = -2.655$, $p = 0.008$; effectgrootte = 0.30) waarbij de vrouwen lager scoren. De voorde-ningen van eerstegraadsopgeleide docenten zijn in zeer geringe mate hoger vergeleken met tweedegraadsopgeleide docenten (eerstegraadsopgeleide docenten score bij aanvang lager maar niet significant lager; $p = 0.086$ en bij de eindmeting gelijk aan tweedegraadsopgeleide docenten). Om na te gaan of de 4 vaardigheidsgroepen

even veel begeleidingsaanbod op de werk-vloer percipiëren worden de beschrijvende statistieken in Tabel 8 gepresenteerd. Het gepercipieerd begeleidingsaanbod van de laagste vaardigheidsgroep is met een t-toets vergeleken met dat van de hoogste vaardigheidsgroep. Hieruit blijkt dat de laagste vaardigheidsgroep meer werkdrukreductie rapporteert (effectgrootte = 0.20) hoewel het verschil niet significant is. Dat geldt ook voor schoolenculturatie (effectgrootte = 0.25) en professionaliseringsmaatregelen (effectgrootte = 0.21). Het verschil in de maatregel “begeleiding in de klas” is marginaal signifi-cant hoger voor de laagste vaardigheidsgroep vergeleken met de hoogste vaardigheidsgroep ($t(141) = -1.764$, $p = 0.080$; effectgrootte = 0.30). Bij de interpretatie van de effectgrootte met het significantie niveau dient rekening ge-houden te worden met de steekproefomvang. Bij een grote steekproef kunnen (kleine) ef-fecten eerder significant zijn.

Tabel 7
Scores leerlingvragenlijst, nulmeting (M0) en eindmeting (M3), $n = 355$

	n	Nulmeting M0		Eindmeting M3		M0 vs. M3 t-test en effectgroottes			
		M	SD	M	SD	t	df	p	effect-grootte
Rasch scores									
Totaal Rasch Leerling-vragenlijst	355	2.12	0.97	2.34	0.86	3.198	708	0.001	0.24
ICALT Rasch - Eerste kwartiel (laagst)	77	1.75	1.11	2.10	0.96	2.093	152	0.038	0.34
ICALT Rasch - Tweede kwartiel	88	1.99	0.95	2.17	0.93	1.27	174	0.206	0.19
ICALT Rasch - Derde kwartiel	96	2.37	0.89	2.49	0.76	1.005	190	0.316	0.13
ICALT Rasch - Vierde kwartiel (hoogst)	89	2.28	0.85	2.58	0.71	2.555	176	0.012	0.38
Type school									
Opleidingsschool	112	1.99	1.01	2.31	0.87	2.540	222	0.012	0.34
Geen opleidingsschool	243	2.18	0.95	2.36	0.86	2.172	484	0.029	0.20
Geslacht docent									
Man	136	2.29	0.89	2.35	0.86	0.568	270	0.572	0.07
Vrouw	219	2.01	1.01	2.33	0.87	3.613	436	0.000	0.35
Graad docent									
Eerste graad	159	2.02	1.01	2.28	0.85	2.482	316	0.014	0.28
Tweede graad	187	2.20	0.93	2.39	0.87	1.998	372	0.042	0.21

5 Conclusies

De implementatie van inwerktrajecten op grote schaal lijkt gelukt te zijn. Werkdrukreductie wordt minder, en begeleiding in de klas wordt meer gerapporteerd door de huidige beginnende docenten vergeleken met de ijkingsstudie. Dit is in lijn met de voornemens van deze interventie. De samenwerking tussen de regionale partners heeft tot de gewenste begeleiding geleid.

Deze studie wijst uit dat het geobserveerd startniveau van de beginnende docenten lager ligt dan verwacht zou kunnen worden van bevoegde docenten. Ongeveer een derde van de docenten ontstijgt bij de nulmeting niet een basaal niveau van pedagogisch-didactisch handelen. Na 3 jaar is dit percentage gedaald naar 7,4%. De uitkomsten laten zien dat bij een eventuele wens om het startniveau te verhogen, landelijke afspraken met betrekking tot de bevoegdheidscesuur⁶ mogelijk aangescherpt dienen te worden. Daarnaast zou het wenselijk zijn om meer doelgericht te coachen en te oefenen/trainen tijdens de eerste- en tweedejarige opleidingen, maar ook tijdens de inductieperiode (Ericson, 2008).

Het negatieve beeld van schoolleiders over het startniveau van alle, en in het bijzonder van eerstegraadsopgeleide docenten in vergelijking met tweedegraadsopgeleide docenten, wordt bij aanvang bevestigd, maar na 3 jaar is een groot deel van de achterstand ingelopen. Het verschil bij aanvang is ook gering (effectgrootte is 0.09) en door de leerlingen wordt

geen significant verschil gepercipieerd. Het percentage docenten met lage vaardigheidsniveaus na diplomering is gelijk voor beide opleidingstypen. De vorderingen tijdens de eerste 3 jaar na diplomering die eerstegraadse docenten maken is groter vergeleken met tweedegraadse beginnende docenten gemeten door beide instrumenten.

Beginnende docenten die op scholen werken waar begeleidingstrajecten zijn ontwikkeld in samenwerking met lerarenopleidingen, groeien gemiddeld 1 standaarddeviatie met betrekking tot geobserveerd pedagogisch-didactisch handelen gedurende 3 jaar. Door de ogen van de leerlingen wordt een kwart standaarddeviatie aan vorderingen waargenomen. De relatief lage vorderingen door de ogen van de leerlingen zou te verklaren kunnen zijn door het hoge startniveau. In de toekomst verdient het plafondeffect van de leerlingvragenlijst nadere psychometrische aandacht.

Minder vaardige docenten rapporteren meer werkdrukreductie en meer begeleiding in de klas vergeleken met de meer startvaardige docenten. Dit is in lijn met de voornemens om met deze interventie maatwerk te leveren. De minder startvaardige docenten vertonen ook de meeste geobserveerde vorderingen over de periode van 3 jaar. Dit geldt ook voor de leerlingpercepties. Daarnaast percipiëren de leerlingen ook bij de hoogste startvaardigheidsgroep significante verbeteringen. In de toekomst is meer aandacht voor de tussenliggende groep m.b.t. vaardigheidsniveau gewenst.

Tabel 8

Gepercipieerde ontvangen begeleiding in aanvangsjaar, per vaardigheidsniveau pedagogisch-didactisch handelen bij aanvang, groep dis nul-en eindmeting heeft geleverd, n = 293

	Werkdruk reductie		School enculturatie		Begeleiding in de klas		Professionele ontwikkeling	
	M	SD	M	SD	M	SD	M	SD
ICALT Rasch - Eerste kwartiel (laagst) (n = 64)	7.81	4.33	3.86	1.88	5.23	2.66	5.61	2.89
ICALT Rasch - Tweede kwartiel (n = 72)	7.50	3.78	3.67	2.06	4.83	2.96	4.86	2.92
ICALT Rasch - Derde kwartiel (n = 78)	7.23	4.07	3.90	1.88	4.52	2.52	5.20	2.60
ICALT Rasch - Vierde kwartiel (hoogst) (n = 79)	6.89	4.35	3.35	2.14	4.38	3.02	4.97	3.12

De observatiebevindingen met betrekking tot beginnende docenten op opleidingsscholen vs. niet-opleidingsscholen duiden op een gelijk startniveau. De geobserveerde vorderingen, over een periode van 3 jaar, zijn gelijk in de opleidingsscholen en de niet-opleidingsscholen. Leerlingen percipiëren meer groei op BSL opleidingsscholen vergeleken met leerlingen op niet-opleidingsscholen die aan BSL mee doen. Dit betekent dat opleidingsscholen niet meer bijdragen aan de ontwikkeling van beginnende docenten vergeleken met reguliere scholen. Dat is gunstig voor de beginnende docenten.

Vrouwelijke docenten profiteren meer van de geboden begeleiding door de ogen van leerlingen en door de ogen van observatoren. In de toekomst is het van belang dit fenomeen nader te onderzoeken ten einde aanpassingen in het arrangement te maken om ook de ontwikkeling van de mannelijke docenten te versnellen. Observatoren geven hogere scores aan vrouwelijke beginnende docenten, terwijl leerlingen aan mannelijke beginnende docenten de hoogste scores geven. De positieve waardering van de leerlingen van de mannelijke beginnende docenten zou wellicht de motivatie om te verbeteren kunnen afremmen en bij de vrouwelijke docenten kunnen bevorderen. In de toekomst zou de reden voor deze perceptieverschillen nader onderzocht kunnen worden. Een andere mogelijkheid die onderzocht zou kunnen worden is of er geslachtsverschillen zijn in aanstellingsomvang die de ruimte voor leren kunnen bevorderen of juist belemmeren.

5.1 Beperkingen van deze studie

Om de effecten van een interventie goed te kunnen meten is een controlegroep vereist. Begeleiding van Startende Leraren is een speerpunt in de beleidsagenda van het Ministerie van Cultuur, Onderwijs en Wetenschap, zoals verwoord in de Lerarenagenda 2012-2020. De meeste scholen ontwikkelen inwerktrajecten en diverse subsidies zijn daarvoor in het leven geroepen. Dit gegeven maakte de werving van een controlegroep voor deze studie onmogelijk. Dat laat niet onverlet dat door de opschaling van de interventie in deze studie, variatie geconstateerd

is in de implementatie over de scholen heen.

De betrouwbaarheid van observaties neemt toe met het aantal observatoren per meting (van der Lans et al., 2018b). In deze grootschalige studie was het niet mogelijk om meerdere observaties per meetmoment te organiseren. De mixed-methode van zowel observaties als leerlingpercepties kan deels voor dit gebrek compenseren.

Uitval van deelname van een vrij grote groep deelnemende docenten/respondenten is over het algemeen een bron van zorg met betrekking tot de generaliseerbaarheid van de bevindingen. Het reponspercentage per geslacht wijkt niet noemenswaardig af van de verdeling in de totale steekproef. Het percentage beginnende docenten op opleidingsscholen is gelijk voor de cross-sectionele als voor de longitudinale data. Er lijken iets meer eerstegraadse docenten te zijn in de longitudinale data vergeleken met de cross-sectionele data. In hoeverre dat de uitkomsten heeft beïnvloed is niet duidelijk. In de toekomst, na afloop van de dataverzameling van dataverzamelingsronde 2 en 3, kunnen deze steekproef-fluctuaties nader onderzocht worden.

6 Discussie en implicaties

De vraag die centraal staat in deze studie is of de grootschalige interventie tot een reductie van zorgen op macro-, meso- en microniveau kan leiden. De term “zorgen” bevat een subjectieve en een objectieve component. Deze studie probeert een objectief antwoord te geven dat wellicht een invloed kan uitoefenen op het subjectieve deel.

Op macroniveau zijn er zorgen met betrekking tot de aantrekkelijkheid van het beroep, de trage ontwikkeling van vaardigheden en onvervulde vacatures. De effecten die gemeten zijn in deze studie zouden bij valorisatie de aantrekkelijkheid van het beroep kunnen verbeteren en het imago dat ontwikkeling traag verloopt kunnen nuanceren. De bevindingen zouden een onderbouwing kunnen vormen voor de structurele facilitering van inwerktrajecten. Hiervoor is het nodig dat erkenning komt voor het gegeven dat vaardighedsontwikkeling tijd (en geld) nodig heeft,

dat docenten verschillen in de mate waarin ze groeien (waarbij persoons- en werkomsomstandigheden een grote rol spelen), en dat de beginnende docent zich goed kan ontwikkelen in een ondersteunende leeromgeving. De opbrengsten verhogen ten slotte de kwaliteit van het onderwijs.

Een veel gehoorde klacht op de werkvloer is dat beginnende docenten een tijdelijke eerste aanstelling van een jaar krijgen. Als er na een jaar een positieve beoordeling én formatieruimte is, volgt een vaste aanstelling. Hierdoor ervaren beginnende docenten een enorme druk om zich te profileren. Tegen dit soort krachten is geen enkel inductie-arrangement opgewassen. Daar is meer voor nodig. Als de tijdelijke aanstelling voor 3 jaar zou zijn en/of als er duidelijke en *realistische* criteria waren voor een vaste aanstelling, waarbij de persoonlijke ontwikkeling de ruimte gegeven wordt, zal de profileringsdruk afnemen.

Op mesoniveau blijkt dat schoolleiders niet tevreden zijn over het startniveau van beginnende docenten en in hogere mate met betrekking tot eerstegraadsopgeleide docenten. Objectief gezien bevestigen onze uitkomsten dit beeld. Het startniveau van een deel van de bevoegde beginnende docenten blijkt minder hoog te zijn dan men zou verwachten. Dit kan meerdere oorzaken hebben. Beginnende docenten hebben niet veel ervaring met verschillende soorten klassen en werken gemiddeld meer lesuren dan tijdens de opleiding. Dat sluit echter niet uit dat deze docenten in opleiding onder bepaalde omstandigheden (bepaalde klassen, jaargroepen en in bepaalde schoolcontexten) betere pedagogisch-didactische vaardigheden kunnen hebben vertoond aan het eind van de opleiding. De transfer van de vaardigheden naar nieuwe omstandigheden kan problematisch zijn en kan een verklaring zijn voor de bevindingen (Perkins & Salomon, 1987). Daarnaast kan het ook zijn dat de opleidingseisen niet scherp genoeg geformuleerd zijn. De vraag blijft staan of dit een probleem is dat opgelost moet/kan worden tijdens de opleiding, of tijdens de inductieperiode en of dat beide ondersteuningsinstanties een bijdrage kunnen leveren. Daarnaast is nader onderzoek nodig naar de percepties van schoolleiders. Welke verwachtingen worden

gekoesterd en welke zijn (on)realistisch.

Op microniveau geven beginnende docenten aan dat ze onder andere problemen ervaren met ordehouden, leerlingmotivatie en omgaan met verschillen tussen leerlingen. Met deze studie is dit bevestigd. Daarnaast is ook aangetoond dat deze vaardigheden verbeteren met gepaste begeleiding.

De vraag of de zorgen op macro-, en mesoniveau ook daadwerkelijk gereduceerd zullen worden naar aanleiding van deze uitkomsten moet nog blijken. Het wegnemen van de zorgen zal afhangen van de bekendheid, de duurzaamheid van de uitkomsten en de aannemelijkheid van de uitkomsten bij de betrokken zorgdragers in diverse gremia. De boodschap is: dat er binnen samenwerkingsverbanden goede inwerktrajecten ontwikkeld kunnen worden voor beginnende docenten en dat het goed lukt om ze op grote schaal te implementeren in het VO. Wat niet bekend is of deze inwerktrajecten ook geborgd zullen blijven. Hiervoor zou de begeleiding van beginnende docenten onderdeel kunnen worden van een toezichtkader van de Inspectie van het Onderwijs van het Ministerie van Onderwijs, Cultuur en Wetenschap. De lerarenopleidingen zouden een structurele rol kunnen vervullen in de ondersteuning van scholen bij de begeleiding van beginnende docenten.

Noten

- ¹ Wij willen de beginnende docenten, de schoolcontactpersonen, de regioleiders en de schoolbegeleiders hartelijk danken voor de grote bijdrage die zij hebben geleverd aan deze studie. Ook zijn wij erkentelijk voor de ondersteunende onderzoekswerkzaamheden van Suus Siekman, Bernie Helms, Els Lugthart, Anna Verkade, Ruth Harmsen en Anna Meijer. Deze studie wordt gefinancierd door het Ministerie van Onderwijs, Cultuur en Wetenschap, projectnummer: OND/ODB-2013/45916.
- ² <http://www.begeleidingstartenderaren.nl/>
- ³ <https://www.rijksoverheid.nl/regering/bewindspersonen/sander-dekker/documenten/publicaties/2016/11/29/bijlage-met-voortgangsrapportage-dashboard-lerarenagenda>
- ⁴ Gezien de relatief kleine steekproef ligt het in

de verwachting dat de verschillen van de andere metingen ook significant hadden kunnen zijn bij een grotere steekproef.

⁵ Zie handreikingen voor inductie op <http://www.begeleidingstartendeleraars.nl/handreiking/>

⁶ De cesuur is de zak/slaaggrens bij de beoordeling van een toets. Daarbij de vraag gesteld welke score een student tenminste moet behalen om een voldoende te halen of om te slagen.

Literatuur

- Adriaens, H., Fontein, P., Den Uijl, M., & De Vos, K. (2016). *De toekomstige arbeidsmarkt voor onderwijspersoneel po, vo en mbo 2015-2025*. Tilburg: CentERdata. Retrieved from <http://www.voion.nl/downloads/26350d9c-a5c5-497c-9698-988dbfa31ff0>
- Bosker, R. J. (2005). De grenzen van gedifferentieerd onderwijs. Groningen: Rijksuniversiteit Groningen.
- Castle, S., Fox, R.K. & O'Hanlan Souder (2006). Do Professional Development Schools (PDSs) Make a Difference? A Comparative Study of PDS and Non-PDS Teacher Candidates. *Journal of Teacher Education*, 57(1), 65-80.
- Cotton, K. (1995). *Effective schooling practices: A research synthesis*. Portland, OR: Northwest Regional Educational Laboratory.
- Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
- Créton, H. A. & Wubbels, T. (1984). *Ordeproblemen bij beginnende leraren*. Utrecht: Uitgeverij W.C.C.
- De Jonge, J. F. M. & De Muijnck, J. A. (2002). *Waarom leraren de sector verlaten. Onderzoek naar de uitstroom in het primair en secundair onderwijs*. Zoetermeer: EIM.
- Den Brok, P. D., Wubbels, T., & van Tartwijk, J. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teachers and Teaching, Theory and Practice*, 23(8), 881-895. doi:10.1080/13540602.2017.1360859
- Ellis, E. S. & Worthington, L. A. (1994). Research synthesis on effective teaching principles and the design of quality tools for educators. (Technical Report No. 5). Eugene: University of Oregon.
- Ericson, K.A. (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview, presented at the Academic Emergency Medicine Consensus Conference, Washington, DC.
- Fantilli, R. D., & McDougall, D. E. (2009). A study of novice teachers: Challenges and supports in the first years. *Teaching and teacher education*, 25(6), 814-825. doi:10.1016/j.tate.2009.02.021
- Fuller, F. F., & Brown, O. H. (1975). In K. Ryan (Ed.), *Becoming a teacher*. Teacher education, 74th yearbook of the national society for the study of education, Part II, University of Chicago Press, Chicago.
- Harmsen, R., Helms-Lorenz, M., Maulana, R., & Van Veen, K. (2018a). The relationship between beginning teachers' stress causes, stress responses, teaching behaviour and attrition. *Teachers and Teaching: Theory and Practice*, 24(6), 626-643. <https://doi.org/10.1080/13540602.2018.1465404>
- Harmsen, R., Helms-Lorenz, M., Maulana, R., & Van Veen, K. (2018b). The longitudinal effects of induction on beginning teachers' stress. *British journal of educational psychology*. <https://doi.org/10.1111/bjep.12238>
- Hattie, J. A. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York, Routledge.
- Heikkinen, H. L., Jokinen, H., & Tynjälä, P. (2012). Teacher education and development as lifelong and lifewide learning. In H. Heikkinen, H. Jokinen, & P. Tynjälä (Eds.), *Peer-group mentoring for teacher development* (pp. 3-30). London: Routledge.
- Helms-Lorenz, M., Koffijberg, J., Schellings, G., Van der Wolk, W., Aarts, R., Runhaar, P., ... van den Hoorn-Flens, P.H. (2017). De driejarige begeleiding van startende leraren in Nederland. In S. de Vos, J. de Wilde, & S. Beusaert (Eds.), *Start to Teach: Inspiratiegids over aanvangsbegeleiding in het onderwijs* (pp. 55-77). Antwerpen-Apeldoorn: Garant Publishers.
- Helms-Lorenz, M., Slof, B., Vermue, C. E., & Canninus, E. T. (2012). Beginning teachers' self-efficacy and stress and the supposed effects of induction arrangements. *Educational Studies*, 38(2), 189-207. DOI: 10.1080/03055698.2011.598679

- Helms-Lorenz, M., Van de Grift, W., Canrinus, E., Maulana, R., & van Veen, K. (2018). Evaluation of the behavioral and affective outcomes of novice teachers working in professional development schools versus non-professional development schools. *Studies in Educational Evaluation*, 56, 8-20. DOI: 10.1016/j.stue-duc.2017.10.006
- Helms-Lorenz, M., Van de Grift, W., & Maulana, R. (2016). Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *School Effectiveness and School Improvement*, 27(2), 178-204. <https://doi.org/10.1080/09243453.2015.1035731>
- Houston, W. R., Hollis, L. Y., Clay, D. Ligons, C., & Roff, L. (1999). Effects of collaboration on urban teacher education programs and professional development schools. In D. Byrd & D. J. McIntyre (Eds.), *Research on professional development schools: Vol. 7. Teacher education* (pp. 6-28). Thousand Oaks, CA: Corwin Press.
- Houtveen, A. A. M., Versloot, B. & Groenen, A. (2006). *De begeleiding van startende leraren; in het voortgezet onderwijs en in het basisonderwijs*. Den Haag: Sectorbestuur Onderwijs-arbeidsmarkt (SBO).
- Kelchtermans, G., & Ballet, K. (2002). The micro-politics of teacher induction. A narrative-biographical study on teacher socialization. *Teaching and Teacher Education*, 18(1), 105-120.
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25(1), 12-23. doi:10.1016/j.tate.2008.06.001
- Levine, D. U. & Lezotte, L. W. (1990). *Unusually effective schools: a review and analysis of research and practice*. Madison, WI: Center for Effective Schools Research and Development.
- Levine, D. U. & Lezotte, L. W. (1995). Effective schools research. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of research on multicultural education*. (525-547) New York: Macmillan.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015a). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement*, 26(2), 169-194.
- Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015b). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teaching and Teacher Education*, 51, 225-245.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2012). *Werken in het Onderwijs*, DeltaHage, Den Haag.
- OC&W (2013). *Lerarenagenda 2013-2020: de leraar maakt het verschil*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- OECD (2016), *Netherlands 2016: Foundations for the Future*, Reviews of National Policies for Education, OECD Publishing, Paris, doi:10.1787/9789264257658-en
- Perkins, D. N., & Salomon, G. (1987). Transfer and teaching thinking. In D. N. Perkins, J. Lochhead, and Bishop (Eds.), *Thinking, the second international conference*, 285-303, Hillsdale, N.J. Lawrence Erlbaum Associates.
- Pillen, M., Beijaard, D., & Den Brok, P. D. (2012). Tensions in beginning teachers' professional identity development, accompanying feelings and coping strategies. *European Journal of Teacher Education*, 36(3), 240-260. doi:10.1080/02619768.2012.696192
- Pillen, M., Beijaard, D., & Den Brok, P. D. (2013). Professional identity tensions of beginning teachers. *Teachers and Teaching: Theory and Practice*, 19(6), 660-678. doi:10.1080/13540602.2013.827455
- Purkey, S. L., & Smith, M. S. (1983). Effective schools: A review. *The elementary School Journal* 83(4), 427-452.
- Ryan, R. M. & Deci, E. L. (Eds.), (2002). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools: a review of school effectiveness research*. London: Ofsted.
- Scheerens, J. (1992). *Effective schooling: research, theory and practice*. London: Cassell.
- Scheerens, J. & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Skilbeck, M., & Connell, H. (2003). *Attracting, developing and retaining effective teachers: Au-*

- stralian country background report. Canberra: Commonwealth Government of Australia.
- Smith, T. M., & Ingersoll, R. M. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal*, 41(3), 681–714.
- Sharpe, T., Lounsbury, M.F., Golden, C., & Deibler, C. (1999). Analysis of an Ongoing, District-Wide Collaborative Approach to Teacher Education. *Journal of Teaching in Physical Education*, 19(1), 79-96. DOI: 10.1123/jtpe.19.1.79.
- Shoffner, M. (2011). Considering the first year: Reflection as a means to address beginning teachers' concerns. *Teachers and Teaching: Theory and Practice*, 17(4), 417-433. doi:10.1080/13540602.2011.580518
- Stokking, K., Leenders, F., De Jong, J., & Van Tartwijk, J. (2003). From student to teacher: Reducing practice shock and early dropout in the teaching profession. *European journal of teacher Education*, 26(3), 329-350. doi:10.1080/0261976032000128175
- Van de Grift, W. J. C. M. (2007). Quality of Teaching in Four European Countries: A review of the literature and an application of an assessment instrument. *Educational Research*, 49(2), 127-152.
- Van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150-159. DOI: 10.1016/j.stueduc.2014.09.003
- Van de Grift, W. J. C. M. & Lam, J. F. (1998). Het didactisch handelen in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 23(3), 224-241.
- Van de Grift, W. J. C. M., Van der Wal, M., & Torrenbeek, M. (2011). Ontwikkeling in de pedagogisch didactische vaardigheid van leraren in het basisonderwijs. *Pedagogische Studien*, 88(6), 416-423.
- Van der Lans, R. M., Van de Grift, W. J., & Van Veen, K. (2015). Developing a teacher evaluation instrument to provide formative feedback using student ratings of teaching acts. *Educational Measurement: Issues and Practice*, 34(3), 18-27. <https://doi.org/10.1111/emip.12078>
- Van der Lans, R. M., Van de Grift, W. J., & Van Veen, K. (2018a). Developing an instrument for teacher feedback: using the rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education*, 86(2), 247-264. <https://doi.org/10.1080/00220973.2016.1268086>
- Van der Lans, R. M., & Maulana, R. (2018b). The use of secondary school student ratings of their teacher's skillfulness for low-stake assessment and high-stake evaluation. *Studies in Educational Evaluation*, 58, 112–121. DOI: 10.1016/j.stueduc.2018.06.003.416-432.
- Van der Ploeg, S., Van de Pol, G., Vankan, A., & Van Kerkhof, D. (2017). Arbeidsmarktbarometer po, vo, mbo 2016/2017. Dialogic and Ecorys. Retrieved from <https://www.voion.nl/downloads/322a7914-ffad-46b0-806a-6eda20a12901>
- Van Kregten, A. & Moerkamp, T. (2004). *Startende leraren in het buitenland*. Den Haag: SBO.
- Van Rooij, E. C. M. (2018). *Secondary school students' university readiness and their transition to university*. Dissertation, University of Groningen.
- Van Til, A. & Van Boxtel, H.W. (2015). *Wetenschappelijke verantwoording Toets 0 t/m 3, tweede generatie*. Arnhem: Cito.
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, 54(2), 143-178.
- Vonk, J. H. C. (1982). *Opleiding en praktijk*. Amsterdam: VU-boekhandel/uitgeverij.
- Walberg, H. J. & Haertel, G. D. (1992). Educational psychology's first century. *Journal of Educational Psychology*, 84(1), 6-19.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response models. *Psychometrika*, 54(3), 427–450.

Auteurs

Michelle Helms-Lorenz is als universitair hoofd-docent verbonden aan de Lerarenopleiding van de Rijksuniversiteit Groningen. **Marieke van der Pers** is als postdoctoraal onderzoeker werkzaam bij de Lerarenopleiding van de Rijksuniversiteit Groningen. **Ridwan Maulana** is als universitair docent verbonden aan de Lerarenopleiding van

de Rijksuniversiteit Groningen. **Rikkert van der Lans** is als postdoctoraal onderzoeker werkzaam bij de Lerarenopleiding van de Rijksuniversiteit Groningen. **Peter Moorer** is werkzaam als onderzoeker bij de Lerarenopleiding van de Rijksuniversiteit Groningen. **Petra van den Hoorn-Flens** was ten tijde van het onderzoek verbonden aan de Lerarenopleiding van de Rijksuniversiteit Groningen als onderzoeker en is thans projectleider bij QANU.

Correspondentieadres: Michelle Helms-Lorenz, Grote Kruisstraat 2/1, 9712 TS Groningen, Nederland; E-mail: m.helms-lorenz@rug.nl

Abstract

Scaling up induction arrangements and measuring longitudinal beginning teacher behaviour in secondary education

This longitudinal study over a period of 3 years, investigates whether the scaling up of an induction intervention benefits the development of teaching skills of beginning teachers. At the 147 participating schools, the teaching skills of novice teachers were periodically measured with two instruments: observations by trained observers using the ICALT instrument (N = 725 cross-sectional, of which n = 420 longitudinal), and the student "My Teacher" questionnaire (N = 650 cross-sectional of which n = 355 longitudinal). The perception of beginning teacher support was surveyed from the beginning teachers (N = 516 cross-sectional, n = 299 longitudinal). The results point to tailored support being the most intensive for teachers with lower starting skill levels. At the start, the observed level of teaching skills appeared to be below the expected starting level. Over a period of three years, the effects of induction support are positive (1.01 and 0.24 standard deviation for the observation and pupil perspective, respectively). Teachers with a low starting level, women and university-educated teachers reveal the highest progress rate.

Keywords: scaling up, support, longitudinal, development, induction.

Appendix A: items begeleidingsmonitor

Voorbeelden van werkdrukreductie maatregelen

- 1 In deeltijd aangesteld worden
- 2 Minder lessen te geven dan meer ervaren collega's
- 3 In het eerste jaar geen mentor te zijn
- 4 Geen extra taken toebedeeld te krijgen binnen de sectie of team
- 5 Vrijgesteld te worden van buitenschoolse activiteiten
- 6 Taken op basis van draagkracht toebedeeld te krijgen, met andere woorden er wordt maatwerk geleverd
- 7 Geen lastige klassen toegewezen te krijgen
- 8 Van klas te mogen wisselen
- 9 Samen met een ervaren collega in parallelklassen les te geven
- 10 Lessen evenwichtig over de week te verspreiden
- 11 Ruimte in te plannen voor intervisie
- 12 Op vaste tijden met begeleiders af te spreken
- 13 Met de vakcoach op dezelfde locatie te werken (de vakcoach is een medewerker belast met de ondersteuning en begeleiding van de beginnende docent. Hij/ zij geeft hetzelfde schoolvak als de beginnende docent)
- 14 Met de algemeen coach op dezelfde locatie te werken (de algemeen coach is een medewerker belast met de ondersteuning en begeleiding van één of meer beginnende docenten. Hij/ zij geeft niet per definitie hetzelfde schoolvak als de beginnende docent)
- 15 Niet overmatig te pendelen tussen locaties (wordt voor mij zoveel mogelijk beperkt)
- 16 Eigen lokaal toegewezen te krijgen
- 17 Een rustige, goede werkplek toegewezen te krijgen
- 18 Een vertrouwenspersoon toegewezen te krijgen

Voorbeelden van schoolenculturatie maatregelen

- 1 Een introductiebijeenkomst
- 2 Een speciaal informatiepakket met informatie over de structuur, organisatie en cultuur van de school
- 3 Informatie verstrekking over inductieactiviteiten
- 4 Een introductie over de schoolregels
- 5 Voorlichting over de opzet, inhoud en structuur van de zorg, door de zorgcoördinator
- 6 Introductie aan andere teams, locaties, niveaus, afdelingen binnen en buiten de school

Voorbeelden van professionaliseringsmaatregelen

- 1 Werken met een (persoonlijk) startprofiel (een startprofiel is een schriftelijk document waarin weergegeven wordt wat de sterke en minder sterke punten van de beginnend docent bij aanvang zijn, alsmede ontwikkelingsbehoeften en -wensen)
- 2 Werken met een (persoonlijk) ontwikkelingsplan
- 3 Periodieke individuele voortgangsgesprekken
- 4 In de loop van het jaar een beoordelingsgesprek
- 5 Reflecteren met collega's op professionele ontwikkeling
- 6 Adequate instrumenten voor het begeleidingsprogramma
- 7 Een goed lesobservatieformulier
- 8 Een methode voor het geven van feedback na een lesobservatie
- 9 Een systeem voor bevraging van leerlingen over hun mening over mijn onderwijs

Voorbeelden van begeleiding in de klas maatregelen

- 1 Geobserveerde lessen (van collega's) nabespreken
- 2 Minimaal 1 videoregistratie laten maken en nabespreken
- 3 Feedback ontvangen op de geobserveerde lessen
- 4 Onderwijs samen plannen met collega's of (vak)coach
- 5 Onderwijs samen voorbereiden met collega's of (vak)coach
- 6 Onderwijs samen uitvoeren met collega's of (vak)coach
- 7 Samen met collega's of (vak)coach mijn eigen onderwijspraktijk evalueren
- 8 Leerling-prestaties met een collega of (vak)coach bespreken met als doel mijn eigen onderwijspraktijk te evalueren

Appendix B: *iCalt* items per domein

Domein	Item	Omschrijving
Vellig en stimulerend leerklimaat	1	De leraar toont in gedrag en taalgebruik respect voor leerlingen
	2	De leraar zorgt voor een ontspannen sfeer
	3	De leraar ondersteunt het zelfvertrouwen van leerlingen
	4	De leraar zorgt voor wederzijds respect
Efficiënte lesorganisatie	5	De leraar zorgt voor een ordelijk verloop van de les
	6	De leraar gaat tijdens de verwerking na of leerlingen de opdrachten op een juiste manier uitvoeren
	7	De leraar zorgt voor een doelmatig klassenmanagement
	8	De leraar gebruikt de leertijd efficiënt
Duidelijke en gestructureerde instructie	9	De leraar geeft duidelijke uitleg van de leerstof
	10	De leraar geeft feedback aan de leerlingen
	11	De leraar betreft alle leerlingen bij de les
	12	De leraar gaat tijdens de instructie na of leerlingen de leerstof hebben begrepen
	13	De leraar bevordert dat leerlingen hun best doen
	14	De leraar geeft goed gestructureerd les
	15	De leraar geeft duidelijke uitleg van het gebruik van didactische hulpmiddelen en opdrachten
Intensieve en activerende les	16	De leraar hanteert werkvormen die leerlingen activeren
	17	De leraar stimuleert het zelfvertrouwen van zwakke leerlingen
	18	De leraar stimuleert leerlingen om over oplossingen na te denken
	19	De leraar stelt vragen die leerlingen tot denken aanzetten
	20	De leraar laat leerlingen hardop denken
	21	De leraar zorgt voor interactieve instructie
	22	De leraar verduidelijkt bij de aanvang van de les de lesdoelen
Afstemmen instructie en verwerking op verschillen	23	De leraar gaat na of de lesdoelen werden bereikt
	24	De leraar biedt zwakke leerlingen extra leer- en instructietijd
	25	De leraar stemt de instructie af op relevante verschillen tussen leerlingen
	26	De leraar stemt de verwerking van de leerstof af op relevante verschillen tussen leerlingen
Leerstrategieën aanleren	27	De leraar leert leerlingen hoe zij complexe problemen kunnen vereenvoudigen
	28	...stimuleert het gebruik van controle activiteiten
	29	De leraar leert leerlingen oplossingen te checken
	30	De leraar bevordert het toepassen van het geleerde
	31	De leraar moedigt kritisch denken van leerlingen aan
	32	De leraar vraagt leerlingen na te denken over strategieën bij de aanpak
Betrokkenheid van de leerlingen	33	De leerlingen zijn betrokken bij de les
	34	De leerlingen tonen zich geïnteresseerd
	35	De leerlingen zijn actief op leren gericht

Appendix C: items “Mijn leraar...” leerlingvragenlijst

Mijn leraar/lerares...

- 1 ... helpt mij als ik iets niet weet
 - 2 ... zorgt dat ik mijn tijd goed gebruik
 - 3 ... maakt duidelijk wat ik voor een proefwerk moet leren
 - 4 ... geeft mij antwoord op mijn vragen
 - 5 ... let erop dat ik anderen met respect behandel
 - 6 ... legt mij uit hoe ik iets moet doen (vertelt mij op welke manier ik iets het beste kan doen)
 - 7 ... legt mij alles duidelijk uit
 - 8 ... maakt duidelijk wat de bedoeling is van de les
 - 9 ... bereidt zijn/haar lessen goed voor
 - 10 ... benadert mij met respect
 - 11 ... zorgt dat ik oplet
 - 12 ... gebruikt duidelijke voorbeelden
 - 13 ... sluit aan bij wat ik kan (vraagt geen dingen die te moeilijk of te makkelijk zijn. Als we iets nieuws leren past dit bij iets dat ik al weet)
 - 14 ... hanteert duidelijke regels (Mijn leraar/lerares heeft duidelijke regels en zorgt dat die nagekomen worden)
 - 15 ... vertelt mij hoe ik iets moet leren
 - 16 ... geeft mij zelfvertrouwen bij moeilijke taken (geeft mij het gevoel dat ik moeilijke opdrachten aan kan)
 - 17 ... zet mij aan het denken
 - 18 ... heeft aandacht voor mij
 - 19 ... controleert of ik de lesstof heb begrepen
 - 20 ... motiveert mij (zorgt ervoor dat ik wil werken en/of leren)
 - 21 ... weet wat ik moeilijk vind
 - 22 ... zorgt dat ik mijn best doe
 - 23 ... betreft mij bij de les
 - 24 ... helpt mij als ik iets niet snap
-