

ORATIE
18 MEI 2017



GERICHT
ONTWIKKELEN
VAN LEERKRACHT-
KWALITEITEN

PROF. DR. A.J. VISSCHER

UNIVERSITY OF TWENTE.



PROF. DR. A.J. VISSCHER

GERICHT ONTWIKKELEN VAN LEERKRACHT- KWALITEITEN

Rede uitgesproken bij de aanvaarding van het ambt van
hoogleraar binnen ELAN, Vakgroep Docentontwikkeling
aan de faculteit Behavioural, Management & Social Sciences
van de Universiteit Twente op donderdag 18 mei 2017

PROF. DR. A.J. VISSCHER

COLOFON

Prof. dr. A.J. Visscher (2017)

© Prof. dr. A.J. Visscher 2017

All rights reserved. No parts of this publication may be reproduced by print, photocopy, stored in a retrieval system or transmitted by any means without the written permission of the author.

May 2017

GERICHT ONTWIKKELEN VAN LEERKRACHTKWALITEITEN

Meneer de rector, beste familie, collega's, vrienden, dear international friends,

“De leerkracht doet ertoe!”, “De leraar maakt het verschil!”

Deze en soortgelijke uitspraken hoort en leest men regelmatig. Ze zijn nogal uitgesproken en uitbundig in de zin van ‘Zie je wel, leerkrachten zijn wél belangrijk!’

Opmerkelijk, alsof we daaraan zouden twijfelen.

Misschien is het een tegengeluid tegen andere, meer bezorgde uitingen over de kwaliteit van ons onderwijs, bijvoorbeeld over de prestaties van Nederlandse leerlingen in internationaal vergelijkende studies zoals PISA en TIMSS, of over het, inderdaad, onacceptabel hoge aantal laaggeletterden in ons land.

Nog maar anderhalf jaar geleden sprak ik in mijn Groningse oratie over 1,3 miljoen Nederlanders tussen de 16 en 65 jaar, oftewel meer dan 10% van alle Nederlandse 16-65 jarigen, die laaggeletterd zijn. Inmiddels heeft de Algemene Rekenkamer (2016) dit probleem ook onder de Nederlanders boven de 65 jaar onderzocht en ze komt op basis daarvan tot een totaal van 2,5 miljoen (!) Nederlanders van 16 jaar en ouder die laaggeletterd zijn en/of moeite met rekenen hebben. Men zou kunnen zeggen dat het maar goed is dat dit mijn laatste oratie is, maar het probleem is natuurlijk allesbehalve grappig.

Het slechte nieuws over laaggeletterden lijkt me geen reden om het belang van de leerkracht ter discussie te stellen. Zonder leerkrachten leren kinderen/jongeren ook wel wat, maar veel minder en veel minder doelgericht en efficiënt, dan wanneer leerkrachten hen lesgeven. We weten uit onderzoek dat leerlingen in de schoolvakanties veel minder leren, ja, ze gaan dan vaak zelfs achteruit qua kennis en vaardigheden (Luyten, Schildkamp & Verachtert, 2009; Verachtert, Van Damme, Onghena & Ghesquière, 2009). Vooral kinderen uit lagere sociaaleconomische milieus

zouden zonder school veel minder leren, omdat het thuisfront veel minder compensatie biedt voor ontbrekend onderwijs.

In verband met het voorgaande wil ik in deze inaugurele rede ingaan op de volgende vragen:

1. Hoe groot is de invloed van leerkrachten op de leerprestaties van hun leerlingen?
2. Hoe ontwikkelen leerkrachten zich tijdens hun carrière?
3. Hoe ontwikkelen experts zich in het algemeen?
4. Wat betekent dit alles voor de professionele ontwikkeling van leerkrachten?

Ad 1: Hoe groot is de invloed van leerkrachten op de leerprestaties van hun leerlingen?

De invloed van leraren op het leren van jongeren is op uiteenlopende manieren onderzocht. Het merendeel van de studies betreft observationeel onderzoek. De prestaties van leerlingen worden dan met een bepaalde test gemeten en statistisch gerelateerd aan factoren op leerkrachtniveau. Het nadeel van een dergelijke onderzoeksopzet is dat er sprake kan zijn van één of meer versturende (confounding) variabelen. Zulke variabelen zijn gerelateerd aan zowel de onafhankelijke als de afhankelijke variabele en verstoren daarmee het verband tussen die twee en dus ook ons inzicht in dat verband.

Leerlingen worden doorgaans niet aselekt toegewezen aan scholen en klassen waardoor die scholen en klassen niet volledig vergelijkbaar met elkaar zijn. In de ene klas zitten mogelijk meer leerlingen met een hoge of lage sociaaleconomische achtergrond en met andere prestatieniveaus, dan in de andere klas waardoor de leerkracht in de ene klas effectiever kan zijn, dan in de andere. Leerkrachten worden binnen scholen ook vaak selectief aan klassen toegewezen, bijvoorbeeld de beste leerkrachten aan de zwakste, of juist aan de beste klassen.

In onderwijskundig onderzoek wordt zoveel mogelijk gecorrigeerd voor die versturende variabelen waarvan we weten dat ze een rol spelen, bijvoorbeeld de correctie voor de sociaaleconomische status van leerlingen, omdat die sterk samenhangt met leerprestaties. Zo worden scholen en klassen meer vergelijkbaar en kan de samenhang tussen hen en één of meer andere factoren beter onderzocht worden. Corrigeren kan echter alleen voor die variabelen waarvan we weet hebben en waarover we data

hebben. Het kan dus zijn dat op de achtergrond een variabele een rol speelt zonder dat we daarmee rekening houden en dat die variabele ons inzicht in de samenhang tussen de factor leerkracht en de factor leerlingprestaties vertroebelt.

Als we de impact van leerkrachten op leerlingprestaties onderzoeken, dan werken we daarom het liefst met klassen met leerlingen, die zoveel mogelijk vergelijkbaar zijn op al hun observeerbare en niet observeerbare kenmerken en bovendien dat de leerkrachten die aan die klassen lesgeven ook overeenkomstig zijn qua observeerbare en niet observeerbare kenmerken. Als we daarin slagen, dan kunnen we prestatieverschillen tussen leerlingen/klassen toeschrijven aan verschillen in de effectiviteit van leerkrachten. In de praktijk is dit heel erg lastig realiseerbaar, omdat leerlingen en leerkrachten dan aselekt aan klassen moeten worden toegewezen. Scholen zullen dat normaliter, terecht, niet willen. In het STAR-project, een onderzoek naar de effecten van klassenverkleining op leerprestaties in Tennessee, gebeurde dit echter wel: leerlingen en leerkrachten werden gerandomiseerd over drie klassengroottesvarianten. De factor leerkracht bleek bij lezen voor 7% en bij rekenen voor 13% de prestatieverschillen tussen de leerlingen in één leerjaar te verklaren. (Nye, Konstantopoulos, & Hedges, 2004; Haertel, 2013).

Misschien valt het u wat tegen dat op basis van de STAR-studie leerkrachten voor, laten we zeggen, gemiddeld 10% verantwoordelijk blijken te zijn voor leerprestatieverschillen. Daarbij moet u bedenken dat andere factoren dan de leerkracht ook een rol spelen voor de prestaties van leerlingen. John Hattie (2003) onderscheidt bijvoorbeeld zes factoren die volgens hem leerlingprestatieverschillen verklaren (één van die factoren is de leerkracht):

1. de *leerling* zelf: zijn/haar geslacht, leeftijd, intelligentie, motivatie, inzet e.d., maar ook de sociaaleconomische status van de ouders van de leerling;
2. de *schoolleider*: zijn/haar invloed is vooral afhankelijk van de mate waarin deze zich richt op de (verbetering van de) kwaliteit van het onderwijs in zijn/haar school;
3. de *schoolorganisatie*: bijvoorbeeld een schoolklimaat dat minder of juist meer op het presteren in de kernvakken gericht kan zijn, maar ook de *gemiddelde* sociaaleconomische status van de leerlingpopulatie van een school;
4. de *klasgenoten*: te denken valt bijvoorbeeld aan de stimulerende invloed die van klasgenoten uit kan gaan en aan leerlingen die andere leerlingen helpen bij coöperatief leren;

5. de *ouders*: hun verwachtingen, stimulansen en support;
6. datgene wat de *leerkracht*, zoals Hattie het noemt, 'brings to the table' tijdens de plus minus 12000 lesuren die Nederlandse leerlingen tot hun 16e les krijgen.

Individuele leerlingkenmerken verklaren leerlingprestatieverschillen in studies doorgaans voor 50 á 60% (Nye, Konstantopoulos, & Hedges, 2004; Haertel, 2013). Het is belangrijk dat we ons realiseren dat deze *buitenschoolse* factoren er zo sterk toe doen voor leerlingprestaties. De leerkracht doet er m.a.w. toe, maar zijn/haar invloed kent dus ook zijn grenzen. De door Hattie genoemde factor 'ouders' zien we hier vooral als een belangrijk leerlingkenmerk, namelijk de sociaaleconomische status van de ouders. Binnenschoolse factoren als de schoolleider, schoolorganisatie en de klasgenoten verklaren samen plus minus 10 á 15% van de leerprestatievariantie (Hattie, 2003; Luyten, Scheerens, Visscher, Maslowski, Witziers, & Steen, 2005).

Hoewel de invloed van de leerkracht op leerprestaties veel kleiner is, dan die van de buitenschoolse factoren is deze factor van de *beïnvloedbare* factoren wel een hele belangrijke, zo niet de belangrijkste (zie ook Goldhaber, Brewer, & Anderson, 1999; Haertel, 2013). In het STAR experiment zagen we dat die invloed in elk van de eerste drie leerjaren van het Amerikaanse basisonderwijs gemiddeld ongeveer 10% is. Dat is echter de uitkomst van één, Amerikaans, onderzoek in drie specifieke leerjaren, met één specifieke leerlingensteekproef. We hebben meer van dat soort studies nodig om met meer zekerheid te kunnen stellen hoe groot de leerkrachtinvloed precies is. Hanushek en Rivkin (2012) komen op basis van tien observationele studies tot een gemiddeld leerkrachteffect van 13% voor lezen (range: 0.07-0.18) en 17% voor rekenen/wiskunde (range: 0.11-0.26). Het lijkt op basis van de nu beschikbare gegevens waarschijnlijk dat het gemiddelde leerkrachteffect voor deze twee vakken in één Amerikaans leerjaar in het basisonderwijs ergens in de range van 10 á 15% ligt. In een analyse van de TIMSS data van Nederlandse leerlingen uit groep 6 bleken kenmerken van het onderwijs in de klas 8,6% van de leerlingprestatieverschillen voor het vak natuuronderwijs te verklaren, voor rekenen/wiskunde was dit percentage 11,7 (Meelissen & Luyten, 2011)¹.

¹ *Het voorgaande impliceert dat een deel van de spreiding in de leerprestaties (10-30% afhankelijk van hoe de groot de invloed van de andere factoren wordt geschat) niet verklaard werd door de in de uitgevoerde studies gemeten factoren, hetzij doordat niet alle relevante factoren werden gemeten, dan wel, doordat ze imperfect gemeten werden.*

We weten dat de school en de leerkracht niet bij alle *schoolvakken* even invloedrijk zijn, dat zagen we zojuist al bij de verschillen tussen lezen en rekenen die in het STAR-experiment werden gevonden. Voor taal en lezen spelen factoren als het thuisvocabulaire, de (voor)leescultuur en de lectuur die thuis beschikbaar is een belangrijke rol en wat er op school en in de klas gebeurt dus een kleinere rol. Bij rekenen is de invloed van de leerkracht juist groter en die van het thuisfront kleiner. Verder is het allesbehalve zeker dat leerkrachteffecten even groot zijn in het *basis-, voortgezet en hoger onderwijs*. Het zou me niet verbazen als de invloed van de leerkracht in het hoger onderwijs kleiner is, dan in het basisonderwijs. We zagen al dat de lagere sociaaleconomische milieus het meest baat hebben bij 'schoolonderwijs', leerkrachten hebben daar dus meer impact dan bij de hogere *sociaaleconomische milieus*. De impact van de leerkracht verschilt bovendien tussen *leerjaren*, omdat je als leerkracht bijvoorbeeld in de onderbouw van het primair onderwijs meer toe kunt voegen aan het kennisniveau van leerlingen, dan in de hogere leerjaren van de basisschool, wanneer het leeuwendeel van de leerstof in de kernvakken al aan de orde is geweest. Het is ook aannemelijk dat de invloed van leerkrachten varieert tussen *landen*. Bijvoorbeeld door verschillen tussen landen in de opleiding van leerkrachten, verschillen in de mate waarin scholen en leerkrachten vrijheid van handelen bezitten, verschillen qua financiële middelen die beschikbaar zijn, de mate waarin er sprake is van vrije schoolkeuze, de leermaterialen (o.a. de lesboeken) die gebruikt kunnen worden enzovoort.

De invloed van meerdere factoren op de prestaties van leerlingen maakt onderwijsverbeteringsaanpakken waarin zowel de leerkracht als de ouders, klasgenoten, de schoolorganisatie en schoolleiding op een coherente manier worden meegenomen kansrijk (denk bijvoorbeeld aan het Amerikaanse Success for All programma; Quint, Zhu, Balu, Rappaport & DeLaurentis, 2015).

Het feit dat we een samenhang vinden tussen leerkrachtgedrag enerzijds en anderzijds de prestaties van hun leerlingen betekent dat leerkrachten in professionele kwaliteit *verschillen*, anders zouden we immers geen statistische samenhang vinden tussen de factor 'leerkracht' en het leren van hun leerlingen. Die verschillen tussen leerkrachten zullen o.a. samenhangen met de manier waarop ze de leerstof uitleggen en met hun leerlingen omgaan, de leermaterialen die ze gebruiken en hoe, hun inzet, de kwaliteit van hun toetsen. Dylan Wiliam stelt in dit verband dat het niet

zoveel uitmaakt naar welke school je gaat, maar vooral welke leerkracht(en) je binnen die school treft. Scholen als de som van in kwaliteit verschillende leerkrachten, verschillen volgens hem onderling aanzienlijk minder in hun prestaties, dan leerkrachten binnen scholen (William, 2016).

Onderzoek van de Amerikaanse economen Hanushek & Rivkin (2006) bevestigt de kwaliteitsverschillen tussen leerkrachten. Zij komen tot de conclusie dat de meest effectieve leerkrachten met hun leerlingen in een half jaar bereiken waar de minst effectieve leerkrachten twee jaar over doen. En waar de meest effectieve leerkrachten in een schooljaar 1,5 jaar leerwinst met hun leerlingen realiseren komen de minst effectieve leerjaren niet verder dan 0,5 jaar leerwinst. Een verschil van een heel schooljaar leerwinst dus tussen leerkrachten in de top en op de bodem van de prestatieverdeling. Hamre & Pianta (2005) lieten in hun onderzoek bovendien zien dat leerlingen uit achterstandsmilieus in klassen van effectieve leerkrachten net zo goed presteren als leerlingen uit 'advantaged' milieus.

We zouden kunnen zeggen dat je als leerling 'een beetje geluk moet hebben', of, dat een leerling juist pech kan hebben. Met een beetje veel pech kom je tijdens je schoolloopbaan meerdere minder sterke leerkrachten tegen, met alle gevolgen van dien voor je leren en voor je toekomst. Dat is ook een belangrijke implicatie van de vaststelling dat leerkrachten het verschil (kunnen) maken!

Het voorgaande maakt duidelijk dat het goed opleiden van leerkrachten van groot belang is: leerkrachten hebben een behoorlijke impact op het leren van leerlingen en die impact verschilt tussen hen. De nu vastgestelde leerkrachteffecten zijn niet per definitie *maximale* effecten. Waarom zouden we de impact van leerkrachten op het leren van leerlingen niet kunnen vergroten met nog betere opleidingstrajecten? Daarbij zou het mooi zijn als de onderlinge effectiviteitsverschillen tussen leerkrachten gereduceerd zouden kunnen worden. We moeten leerkrachten zo goed mogelijk voor hun complexe taak toerusten en de pechansen van leerlingen minimaliseren. Iedere leerling heeft recht op een goede leerkracht!

Ad 2: Hoe ontwikkelen leerkrachten zich tijdens hun carrière?

We weten opmerkelijk veel minder over het leren en de ontwikkeling van leerkrachten, dan over hoe hun leerlingen leren. Hoe ontwikkelen leerkrachten de diverse aspecten van hun vakmanschap en tegen welke

obstakels lopen ze daarbij tijdens hun carrière op?

Veel van de studies naar de ontwikkeling van leerkrachten zijn cross-sectionele studies. Daarin wordt bij wijze van spreken een 'foto' genomen: op één moment wordt de leerkrachtkwaliteit gemeten en de uitkomsten worden gerelateerd aan het aantal ervaringsjaren van leerkrachten, om zo conclusies te kunnen trekken over de relatie tussen leerkrachtkwaliteit en de werkervaring van leerkrachten. Van de Grift (2010) laat in zo'n cross-sectionele studie zien dat beginnende leraren na ongeveer vijf ervaringsjaren *het pedagogisch-didactisch niveau*, zoals gemeten met het ICALT-instrument, van de gemiddelde leerkracht bereiken. Daarna groeien leerkrachten nog een beetje, totdat ze 10 jaar ervaring hebben. Het decennium daarna is er geen sprake meer van groei en bij meer dan 20 ervaringsjaren treedt een daling op in de beheersing van de didactische vaardigheden.

Van de Grift laat ook zien dat na een aantal jaren de basisvaardigheden, zoals het bieden van een veilig klasklimaat, klassenmanagement en het uitleggen van de stof door de meeste leerkrachten goed beheerst worden. De meer *complexe* didactische vaardigheden blijven daarna echter voor veel leerkrachten een probleem vormen. Dit strookt met gegevens van de Inspectie van het Onderwijs (2016) die op basis van de beoordeling van 1400 lessen in het voortgezet onderwijs stelt dat in de schooljaren 2013/2014 en 2014/2015 ongeveer twee derde van de leerkrachten alle algemeen-didactische vaardigheden beheerst. Oftewel, een derde van de leerkrachten in dat onderzoek beheerst niet alle algemeen-didactische vaardigheden! Een minderheid van ongeveer een derde van de VO-leerkrachten beheerst ook alle differentiatievaardigheden.

De *interactie tussen leerkrachten en leerlingen* verschilt eveneens tijdens de verschillende fasen van een leerkrachtcarrière. Brekelmans (2010) liet in haar inaugurele rede zien dat startende leerkrachten het meest tolerant zijn in de klas, ze hebben dan nog weinig ervaring met leiding geven en grenzen stellen. Na 6 jaar zijn hun gezag en invloed echter aanzienlijk gegroeid. In de tweede helft van hun loopbaan nemen het leeftijdsverschil met de leerlingen en de door hen ervaren routine in het werk toe terwijl de energie geleidelijk aan afneemt. De emotionele afstand tussen leerkracht en leerlingen neemt dan toe en leerkrachten worden autoritairder.

Day (2008) onderscheidt op basis van zijn studie in 100 Engelse scholen met 300 leerkrachten in navolging van de bekende studie van Huberman (1993) drie leerkrachtcarrièrefasen.

The Early years (1 - 7 jaar)

Tijdens de eerste drie jaren hebben leerkrachten veel commitment, ze vinden dan dat ze voor hun leerlingen 'het verschil kunnen maken'. Ze ervaren echter ook klassenmanagementproblemen. Na die drie jaar groeit veelal het vertrouwen in het eigen professionele kunnen, maar bij 20% van de leerkrachten zijn in deze fase hun effectiviteit, self-efficacy en professionele identiteit *at risk*.

The Mid years (8 - 23 jaar)

Deze fase wordt gekenmerkt door spanning, toegenomen verantwoordelijkheden binnen de school, het zoeken van een balans tussen persoonlijk leven en werk en door stagnatie van de carrière (niet kunnen doorgroeien). Een behoorlijk deel van de leerkrachten verliest aan motivatie en commitment.

The Late years (24 - 31+)

Ruim de helft van de leerkrachten behoudt na dertig jaar werkervaring zijn/haar motivatie. Ongeveer een derde van deze leerkrachten voelt zich echter 'tired and trapped', hun effectiviteit daalt in vergelijking met de Mid-years en de Early-years leerkrachten.

In elk van de drie fasen variëren leerkrachten dus behoorlijk qua effectiviteit, uitdagingen en zorgen. De gedachte dat alle leerkrachten gedurende hun hele loopbaan steeds effectiever worden blijkt in stijf te zijn met de onderzoeksresultaten. Day wijst op de invloed van de (verschillen tussen de) *contexten* waarin leerkrachten werken, bijvoorbeeld qua percentages achterstandsléerlingen waarmee leerkrachten werken en ook wat betreft de steun die leerkrachten bij hun werk ontvangen binnen hun school. Die contextverschillen beïnvloeden de effectiviteit van de leerkrachten.

Kini & Podolsky (2016) hebben recentelijk een meta-analyse uitgevoerd naar de ontwikkeling van leerkrachteffectiviteit waarin 30 Amerikaanse studies in het basisonderwijs werden meegenomen, die aan door hen gehanteerde methodologische criteria voldeden. Ze kwamen tot de conclusie dat de *effectiviteit* van leerkrachten, zoals gemeten met een gestandaardiseerde

test voor leerlingen, vooral sterk groeit tijdens de eerste tien jaren van hun carrière (zie ook Leigh, 2010), om vervolgens minder te groeien in het tweede decennium van hun loopbaan en daarna. Na 25 jaar neemt de effectiviteit af.

In de studie bleken twee factoren belangrijk te zijn voor de groei van leerkrachteffectiviteit:

1. Het aantal jaren dat een leerkracht ervaring heeft met lesgeven in een *leerjaar* en met het *vak* waarin hij/zij lesgeeft.
2. Een ondersteunende professionele werkomgeving: onderlinge samenwerking en support vooral ook aan startende leerkrachten, een ondersteunende schoolleider, professionele evaluatie en feedback en, tot slot, de geboden mogelijkheden voor professionele ontwikkeling.

Op basis van het voorgaande kunnen we het volgende vaststellen:

- a. 'De' leerkracht bestaat niet. De kenmerken van de *leerlingen* waarmee zij werken variëren evenals *de mogelijkheden en de ondersteuning* die hen vanuit de schoolorganisatie geboden worden. Bovendien *variëren leerkrachten in elke carrièrefase* aanzienlijk qua competenties, effectiviteit, uitdagingen en zorgen.
- b. Hoewel de *motivatie* en het *commitment* van veel leerkrachten aanvankelijk hoog zijn verlaat ook een aanzienlijk deel van hen al snel het beroep. In Nederland gaat het hierbij om maar liefst 18%, 31% en 35% startende leerkrachten die reeds binnen vijf jaar respectievelijk het PO, VO en MBO verlaten. De groep onbevoegden blijkt in deze percentages, in ieder geval in het voortgezet onderwijs, sterk oververtegenwoordigd te zijn. Als ook gecorrigeerd wordt voor het percentage leerkrachten dat op een tweede school wel slaagt, dan blijkt het vertrekpercentage voor docenten die de lerarenopleiding afmaakten 11% te zijn (Platform Beleidsinformatie, 2014)! Halverwege de carrière neemt de motivatie af, niet alleen door factoren gerelateerd aan het werk, maar ook door factoren die te maken hebben met de thuissituatie en de levensfase van leerkrachten. Gedurende de laatste tien jaren van hun arbeidzame leven voelt een derde van hen zich 'caught in a trap'.
- c. Leerkrachten maken logischerwijs tijdens de eerste fase van hun carrière een sterke groei door naar het gemiddelde niveau van *didactische vaardigheden* en naar een niveau van *interactie met leerlingen* waarbij zij gezag en invloed hebben en waarbij hun leerlingen het gevoel hebben dat ze gehoord en begrepen worden. Leerlingen krijgen van beginnende leerkrachten dus gemiddeld minder goed

onderwijs, dan leerlingen die in de lessen van meer ervaren leerkrachten zitten.

- d. De *groei in de didactische basiscompetenties* stukt na tien jaar en na twintig jaar nemen deze competenties af. De beheersing van *complexe didactische vaardigheden* vormt voor de meerderheid van leerkrachten gedurende hun hele carrière een probleem.
- e. Het duurt lang voordat leerkrachten het *gemiddelde effectiviteitsniveau* bereiken. Ervaring met lesgeven in hetzelfde leerjaar (bijvoorbeeld Havo 4) en met hetzelfde gedoceerde vak bevorderen dit. Leerkrachten hebben ook tijdens de laatste fase van hun carrière *minder impact* op het leren van hun leerlingen.

De begeleiding van startende leerkrachten blijkt positieve effecten te hebben op de ontwikkeling van hun competenties en ook hun uitval te verminderen (Helms-Lorenz, van de Griff, Maulana, 2016). Mogelijk kunnen we ook in de laatste fase van een leerkrachtcarrière met behulp van leeftijdsafhankelijk personeelsbeleid de kwaliteit van leerkrachten in positieve zin beïnvloeden. De inzet van ervaren leerkrachten voor de begeleiding van beginnende leerkrachten ligt voor de hand. Naast een landelijk kader voor personeelsbeleid dat dit soort maatregelen mogelijk maakt is hiervoor, gegeven de geschetste variëteit in contexten, ruimte voor schoolspecifiek personeelsbeleid belangrijk.

Zouden we verder de tijd die nu nodig blijkt te zijn om door te groeien naar een gemiddeld niveau van professioneel functioneren niet kunnen bekorten? Misschien moeten we leerkrachten meer en anders opleiden, voordat ze volledig zelfstandig les gaan geven? En zouden we kunnen bevorderen dat de professionele groei van leerkrachten niet stukt na een jaar of tien werken, dat het gemiddelde prestatieniveau van alle leerkrachten hoger komt te liggen en dat we hen ook goed toerusten met belangrijke complexe didactische vaardigheden, zoals formatief evalueren en differentiëren?

Veel van deze vragen hebben te maken met de ontwikkeling van de professionele expertise van leerkrachten. Het is daarom interessant om stil te staan bij hoe men expertise ontwikkelt op andere terreinen dan het onderwijs.

Ad 3: Wat weten we over de ontwikkeling van experts in het algemeen?

Een expert typeer ik hier als een 'topper', als 'iemand die op *een bepaald terrein* consistent qua prestaties boven zijn/haar collega's uitsteekt'. We weten dat expertise domeinspecifiek is, er is niet zoiets als 'een expert in alles'. Ton Sijbrands speelde 32 dampartijen blind zonder er eentje te verliezen, maar gaf in een interview aan dat hij een briefje mee moet nemen als hij naar de supermarkt gaat om een paar producten te kopen. En de echtgenote van Nobelprijswinnaar Feringa antwoordde op de vraag hoe het is om met zo'n geweldig slimme man in één huis te wonen, "*Hij is slim in scheikunde he?*". Feringa's slimheid gaat echter verder, want onlangs zei hij in het programma *College Tour* dat voor de bevordering van kansengelijkheid in het onderwijs en voor topwetenschap het belangrijkste dat men kan doen het zorgen voor goede leerkrachten is.

Uit de literatuur komt een beeld naar voren van experts die na het jarenlang ontwikkelen en verfijnen van hun competenties snel en accuraat problemen op hun vakgebied herkennen, over verfijnde en gestructureerde mentale representaties van die problemen beschikken, en die problemen snel en voor een deel automatisch kunnen oplossen. Ze kunnen, wanneer problemen daar om vragen, bovendien snel en adequaat switchen van oplossingsaanpak (Berliner, 2001).

Een interessante vraag in het kader van de bespreking van de ontwikkeling van leerkrachtkwaliteiten is hoe je een expert wordt. Daarover zijn de experts op dit onderzoeksterrein het grappig genoeg niet helemaal eens. Hambrick, Brooke, Macnamara, Campitelli & Mosing (2016) stellen op basis van een recente meta-analyse van 88 studies naar expertgedrag dat experts 'are born and made'. Experts hebben volgens hen bepaalde talenten, genetische kenmerken en persoonskenmerken, maar ze hebben hun expertise ook ontwikkeld op basis van veel ervaring en veel oefenen.

Anders Ericsson huldigt een andere opvatting op basis van ruim dertig jaar onderzoek naar uitzonderlijk presteren en de ontwikkeling van expertise. Hij werd op jonge leeftijd door Nobelprijs-laureaat Herbert Simon vanuit Zweden naar de USA gehaald. Simon was niet alleen een vermaard econoom en organisatiekundige, maar ook zeer geïnteresseerd in het handelen van experts. Zo leerde hij naar eigen zeggen Nederlands om het beroemde werk, het *Denken van den schaker*, van A.D. de Groot uit 1946

over probleemoplossen door topschakers te kunnen lezen. De vertaling verscheen op aanraden van Simon in 1965 als *Thought and Choice in Chess* en heeft bijgedragen aan de cognitieve revolutie in de psychologie. Diezelfde de Groot ontving in 1991 aan deze universiteit overigens een eredoctoraat voor zijn wetenschappelijk oeuvre.

Ericsson heeft op tal van terreinen, zoals wetenschap, schaken, ballet, turnen, muziek, geheugentraining en medische specialismen onderzocht hoe je in die domeinen excellent wordt. Volgens hem zijn experts vooral 'made' en wel door heel veel, *gericht* trainen. Expertise-ontwikkeling is naar zijn mening vooral een kwestie van *deliberate practice* (DP), in het Nederlands zouden we waarschijnlijk spreken van 'doelgericht/weloverwogen oefenen'. Iedereen kan volgens Ericsson zijn/haar prestaties enorm verbeteren door middel van DP (Ericsson, 2006, Ericsson, Krampe & Tesch-Römer, 1993). Hij heeft op een aantal terreinen ook laten zien hoe je toppers kunt 'maken'.

Wat is DP? Nou, normaliter ontwikkelen professionals door opleiding en werkervaring geleidelijk aan een basisrepertoire aan professionele competenties tot op een *acceptabel*, een 'goed genoeg' niveau. Bij het bereiken van een natuurlijk plafond stopt de expertise-ontwikkeling. Volgens Ericsson is hetgeen we doen, wanneer we dat natuurlijke plafond bereiken, vaak niet meer dan 'hetzelfde herhalen'. Goed worden in iets is echter niet, zoals Wesley Sneijder onlangs op tv stelde, 'gewoon een kwestie van 10000 uur trainen'. Als simpelweg veel oefenen voldoende zou zijn, dan zou elke leerkracht bij een volledige werkweek na ongeveer tien jaar een expert zijn en we zagen al dat dit niet het geval is. Van herhalen wat je al kunt word je niet beter; practice does *not* make perfect. Op die manier groeien we niet door en neemt na twintig jaar de kwaliteit van het professionele handelen zelfs af. Wat ik besprak m.b.t. de ontwikkeling van leerkrachten is hier illustratief: sterke groei in de eerste professionele jaren, daarna weinig groei meer en na verloop van tijd afname van professionele competenties en effectiviteit.

Doorgroeien vereist *deliberate practice*. Experts groeien volgens Ericsson wél continu door, doordat zij door middel van *langdurig, intensief en heel gericht trainen* verfijnde mentale voorstellingen ontwikkelingen van de situaties, relaties, regels, problemen en oplossingen die karakteristiek zijn voor hun domein. Die mentale voorstellingen van een domein vormen naar zijn mening de basis voor excellente prestaties, mede doordat experts die over de mentale modellen beschikken hun eigen handelen kunnen monitoren op basis van die ideale werkwijzen.

Er moet in het geval van DP aan een aantal voorwaarden worden voldaan. Als je niet slechts 'goed genoeg', maar 'de beste' wil worden, dan:

1. Ben je zeer *gemotiveerd en gedisciplineerd* om te verbeteren: je bent bereid om veel en lang te werken aan datgene waarin je nog niet goed bent.
2. *Verlaat je je comfort zone* en zoek je bewust naar je zwakke (professionele) competenties: *waar kan ik nog verbeteren?* Daarbij vergelijk je je eigen vaardigheden met dat van ideaal handelen (zoals het handelen van een topper).
3. Bepaal je vervolgens *heel nauwkeurig wat je aan je handelen wilt verbeteren*.
4. *Oefen je frequent en kort, maar geconcentreerd en intens gedurende een lange periode*, net zo lang totdat je het gestelde doel hebt bereikt. Volgens Ericsson ontwikkelt niemand zonder reusachtig veel gericht oefenen exceptionele vaardigheden, ook bij zogenaamde 'wonderkinderen' is het aantal uren DP beslissend.
5. Heb je een *kundige coach*, die voor een bepaald domein weet hoe het handelen van een expert in een bepaald domein eruit ziet, bijvoorbeeld hoe een leerkracht tijdens de uitleg van de leerstof fantastisch kan differentiëren, of, op een heel ander terrein, hoe een perfecte karatetrainer eruit ziet. De coach weet welke problemen waarschijnlijk zijn bij het trainen van een bepaalde vaardigheid (vergelijk de vakdidactische kennis van leerkrachten). Bovendien weet de coach hoe je die potentiële problemen kunt proberen te voorkomen, welke leertaken een professional naar een hoger plan kunnen tillen en hoe je effectief feedback geeft aan de expert-in-ontwikkeling (Ericsson & Pool, 2016).

Ad 4: Wat betekent dit alles voor de professionele ontwikkeling van leerkrachten?

De hier beschreven kernkarakteristieken van DP zijn in de nascholing van leerkrachten ongebruikelijk: het uit je comfortzone treden, zoeken naar professionele deficiënties in het licht van heldere professionele standaarden, precieze verbeterdoelen formuleren, daar hard aan werken terwijl je regelmatig kwaliteitsfeedback ontvangt van een coach die weet hoe 'het ideaal' eruit ziet evenals wat effectieve wegen naar het ideaal zijn.

De initiële scholing van leerkrachten kan niet anders dan beperkt zijn, gegeven de daarvoor beperkt beschikbare tijd in verhouding tot de complexiteit van het leraarschap, die niet alleen de beheersing

van uiteenlopende algemeen pedagogisch-didactische didactische vaardigheden veronderstelt, maar ook grondige vakinhoudelijke kennis en vakdidactische kennis en vaardigheden. Het is onmogelijk om hierin allemaal te voorzien tijdens de initiële opleiding van leerkrachten. Een leerkracht is na de initiële opleiding als het goed is startbekwaam. Daarna gaat veel tijd en energie naar het met vallen en opstaan verwerven van de didactische basisvaardigheden. Startende leerkrachten worden de laatste jaren gelukkig steeds meer begeleid in de BSL- (begeleiding startende leraren) trajecten. Het zou mooi zijn als we die ontwikkeling door zouden kunnen trekken naar een doorlopende leerlijn waarbinnen ook gevorderde leerkrachten (BGL: begeleiding gevorderde leerkrachten) systematisch door kunnen werken aan hun professionele ontwikkeling. Ook als ze na een jaar of tien werken hun natuurlijke plafond bereiken. Bij een complexe activiteit als lesgeven is er zelfs voor zeer ervaren en kundige leerkrachten altijd ruimte voor professionele groei. Dat vereist een doordachte leerlijn van startende tot vakbekwame leerkrachten en lukt niet met incidentele, hap snap professionaliseringsinitiatieven.

Nu is de na- en bijscholing van leerkrachten nog te vaak een *one-size-fits-all* aangelegenheid. Niet adaptief, niet aansluitend bij de uiteenlopende individuele professionele ontwikkelingsbehoeften van leerkrachten terwijl we zagen dat leerkrachten nogal verschillen in wat ze kunnen en presteren. Dat adaptief bijscholen, dat kan nu ook niet, daarvoor ontbreekt het aan inzicht in de mate waarin individuele leerkrachten beschikken over belangrijke professionele kennis en vaardigheden. Verder is de rol van leerkrachten in na- en bijscholingsactiviteiten nog te vaak een passieve, receptieve, doordat het accent teveel op informeren en kennisoverdracht ligt. Hoewel er uitzonderingen zijn (o.a. de lesson study aanpak, docentontwikkelteams en professionele leergemeenschappen waarbij buiten de klas wordt voorbereid wat leerkrachten in de klas gaan doen) zijn nog te weinig scholingsactiviteiten gericht op het ontwikkelen van direct in de klas toepasbare leerkrachtvaardigheden. Ja, naar mijn mening ontwikkelen we leerkrachtvaardigheden nog veel te weinig gericht.

Van DP als bewezen effectieve methodiek voor het ontwikkelen voor topprestaties kunnen we het nodige leren voor de scholing van leerkrachten, maar ze is niet zomaar een, twee, drie volledig te vertalen naar de professionalisering van leerkrachten. Zo hebben experts bijvoorbeeld een uitzonderlijke drive om te verbeteren. Hoewel leerkrachten goed

onderwijs een warm hart toedragen is het niet realistisch om ervan uit te gaan dat de gemiddelde leerkracht een dergelijk fanatisme aan de dag zal leggen om zich professioneel te verbeteren. Dat hoeft ook niet, we kunnen alleen al logisch gezien niet allemaal boven onze collega's uitsteken, maar we kunnen wel allemaal verbeteren.

Jansen in de Wal (2016) liet recentelijk in zijn dissertatie-onderzoek zien dat de motivatie om te verbeteren varieert tussen leerkrachten, maar ook in positieve zin beïnvloed kan worden wanneer professionalisering gestimuleerd en gefaciliteerd wordt vanuit de schoolorganisatie en de schoolleiding. Ook bleken leerkrachten meer gemotiveerd te zijn voor professionaliseringsactiviteiten wanneer ze inspraak hadden in hun eigen professionalisering. Oudere leerkrachten die hun carrière er grotendeels op hebben zitten, hadden vaak geen trek meer in verbeteren, ze achtten zich bekwaam genoeg.

'Sociale motivatie' is ook van belang voor professionalisering, denk bijvoorbeeld aan een heel schoolteam dat besluit om aan de differentiatievaardigheden te gaan werken. 'Kunnen differentiëren' wordt daarmee dan de norm en een leerkracht hoeft dan niet in zijn eentje te zitten zwoegen met iets dat erg moeilijk is. Sociale steun en tips van collega's kunnen enorm helpen, met name wanneer het verbeteren niet lukt. En, 'success breeds success' en dat is onder anderen afhankelijk van hoe doordacht en effectief de gehanteerde leer- en verbeterstrategieën zijn. Op die leerpsychologische basis voor de trainingsactiviteiten ga ik straks in.

Naast de 'verbetermotivatie' van leerkrachten zijn er nog diverse andere grote uitdagingen bij de vertaling van DP naar leerkrachtprofessionalisering, zoals de organisatie van het frequent en langdurig trainen van leerkrachtvaardigheden binnen de realiteit van de schoolpraktijk, het werken met coaches, die voldoen aan al de hierboven onder de vijfde DP-karakteristiek genoemde kenmerken en het ontwikkelen van ondersteunende schoolorganisatieculturen gericht op zo goed mogelijk onderwijs. Zelfs als leerkrachten zeer gemotiveerd zijn om aan hun competenties te werken is het namelijk niet zeker dat dat ook daadwerkelijk gebeurt. Leerkrachten hebben het druk met de diverse taken waaraan ze werken, de kans dat professionalisering er bij inschiet is groot. De school moet professionalisering daarom niet alleen mogelijk maken, maar er ook op toezien dat productief gewerkt wordt aan de kerncompetenties, die verbetering behoeven.

In het resterende deel van deze oratie zal ik meer uitgebreid ingaan op drie andere voorwaarden voor leerkrachtprofessionalisering op basis van DP:

- De beschikbaarheid van heldere standaarden voor het professionele handelen van leerkrachten.
- Een leerpsychologisch fundament voor de training van leerkrachtvaardigheden.
- Een instrumentarium om de professionele verbeterruimte van leerkrachten te bepalen.

Heldere professionele standaarden en een leerpsychologisch basis: de match van Deliberate Practice met 4C/ID

Het feit dat bij DP in de inhoud van de trainingsactiviteiten wordt voortgebouwd op hoe de toppers in een domein handelen is aantrekkelijk. Zo kunnen ook anderen, dan de toppers die bewezen effectieve aanpakken ontwikkelen en gaat niet steeds, keer op keer weer enorm veel tijd en energie verloren met het zelf uitvinden van het bekende wiel. Dit veronderstelt voor leerkrachtprofessionalisering heldere standaarden voor vormen van 'goed lesgeven'. Volgens mij hebben we die standaarden voor bepaalde leerkrachtcompetenties al meer dan voor andere. Hoe differentieer je bijvoorbeeld als leerkracht effectief, hoe ziet formatieve evaluatie er idealiter uit en hoe bevordert een leerkracht de zelfregulering van leerlingen met succes? Daar waar dergelijke standaarden nog niet beschikbaar zijn moeten we ervoor zorgen dat we die krijgen. Naar mijn mening matcht het Four components Instructional design (4C/ID) model van Van Merriënboer & Kirschner (2013) hier heel goed met DP. In 4C/ID is een cognitieve taakanalyse van expertgedrag als input voor het trainingsontwerp van groot belang. Bovendien biedt 4C/ID een sterk leertheoretisch fundament voor het trainen van complexe vaardigheden, iets dat DP ontbeert. De cognitieve taakanalyse van bijvoorbeeld de didactische vaardigheid 'differentiëren' omvat een analyse van de te verrichten differentiatietaken en de kennis, vaardigheden en attitudes die daarvoor essentieel zijn. Uit welke routines en niet-routine activiteiten bestaan die differentiatietaken? En welke kennis (o.a. van het vak in kwestie, maar bijvoorbeeld ook algemene didactische kennis, vakdidactische kennis en evaluatiekennis) en strategische probleemaanpakken hanteren leerkrachten wanneer zij differentiatiebeslissingen nemen?

Op basis van de resultaten van de cognitieve taakanalyse worden binnen 4C/ID vervolgens, op grond van een expliciete *leerpsychologische rationale*,

leertaken ontworpen die representatief zijn voor de beoogde complexe differentiatiepraktijk. Die leertaken worden geordend van simpele naar complexe taken. De ondersteuning die bij het leren uitvoeren van deze leertaken gegeven wordt, is op een gegeven niveau van complexiteit eerst uitgebreid, maar neemt af naarmate leerkrachten de taken beter beheersen. De lerenden (in dit geval leerkrachten) worden tijdens het leren van de vaardigheden steeds voorzien van informatie die belangrijk is voor het uitvoeren van routine en non-routine deeltaken (de cognitieve strategieën en mentale modellen die belangrijk zijn voor het nemen van beslissingen in uiteenlopende situaties). Omdat een lerende uit de leertaken die taken kan kiezen, die gezien zijn/haar competenties het meest relevant voor hem/haar zijn, biedt deze aanpak ook een benadering voor het omgaan met de competentieverschillen tussen leerkrachten. De 4C/ID aanpak heeft zich in andere contexten waarin sprake is van complex leren (bijvoorbeeld bij de opleiding van medisch specialisten en luchtverkeersleiders) al veelvuldig bewezen.

Hoe we dat nu doen in ELAN-projecten

Op dit moment worden in onderzoeksprojecten binnen ELAN in nauwe samenwerking met Maastricht University en de Onderwijsinspectie cognitieve taakanalyses gemaakt van differentiëren en van formatieve evaluatie als complexe leerkrachtvaardigheden. Bij differentiëren richt de taakanalyse zich in één project op differentiëren in het primair onderwijs voor het vak rekenen. In een ander project wordt de leerkrachtvaardigheid differentiëren in het voortgezet onderwijs geanalyseerd voor het vak Engels. De analyse in het primair onderwijs wordt uitgevoerd voor twee verschillende contexten: voor de setting waarin digitale systemen worden gebruikt om de voortgang van leerlingen tijdens de les te monitoren (de leerkracht krijgt dan digitale informatie die gebruikt kan worden om te differentiëren) en voor de traditionele situatie waarin geen gebruik wordt gemaakt van digitale systemen. De op formatieve evaluatie gerichte taakanalyse in het voortgezet onderwijs wordt voor meerdere vakken uitgevoerd. De resultaten van de taakanalyses vormen de input voor het ontwerp van leerkrachttrainingen op basis van de leertheoretische aanpak, zoals die in het 4C/ID model gehanteerd wordt: van simpele taken naar steeds complexere leertaken en van veel support voor de lerende (hier de leerkracht) naar afnemende ondersteuning. De trainingen worden vervolgens in scholen uitgevoerd en de impact ervan op leerkrachten en op leerlingen wordt onderzocht. Tevens willen we de projectuitkomsten graag

benutten voor de initiële opleiding van leerkrachten, zodat leerkrachten daar al met deze complexe vaardigheden kennis kunnen maken.

Het bepalen van de individuele verbeterruimte van leerkrachten

Omdat in de DP-aanpak het gericht werken aan nauwkeurig geïdentificeerde verbeterpunten zo'n belangrijke plaats inneemt zal ik nog ingaan op hoe leerkrachten die verbeteringspunten zouden kunnen (laten) bepalen.

Zelfevaluatie door leerkrachten

Eén manier om dat te doen is op basis van zelfevaluatie. Leerkrachten worden net als hoger opgeleide professionals in andere sectoren op basis van hun jarenlange opleiding in staat geacht om te kunnen gaan met complexe en onzekere professionele situaties. Een leerkracht die ad hoc moet kunnen reageren op alles wat er in de klas gebeurt is te vergelijken met een advocaat die van tevoren nooit weet hoe een pleidooi zal verlopen en met een chirurg die tijdens een operatie onverwachte dingen tegen kan komen. Ze genieten allen veel professionele autonomie, omdat de kern van hun werk nogal onvoorspelbaar en slechts beperkt te standaardiseren is (Mintzberg, 1979).

Van hoog opgeleide professionals wordt ook verwacht dat ze op basis van zelfregulering hun professionele competenties onderhouden en verbeteren. Eva en Regehr (2005) hebben voor de medische sector laten zien dat dit 'overly optimistic' is in verband met het 'Lake Wobegon effect'. Kruger en Dunning (1999) stelden dit effect vast in vier studies naar het vermogen van mensen om de kwaliteit van hun eigen vaardigheden in te schatten. Het effect komt er in het kort op neer komt dat de meesten van ons vinden dat we professioneel bovengemiddeld presteren. Lake Wobegon verwijst naar een fictief stadje in een Amerikaanse radioserie waarin alle vrouwen sterk zijn, de mannen mooi en alle kinderen bovengemiddeld intelligent. Kruger en Dunning toonden aan dat met name de laagst presterenden binnen een bepaald domein de metacognitie voor valide zelfbeoordelingen missen: "*It takes one to know one.*" Zij stelden verder op basis van hun onderzoeksresultaten:

"In essence, we argue that the skills that engender competence in a particular domain are often the very same skills to evaluate competence in that domain- one's own or anyone else's". (p. 1130).

Diegenen waarvoor verbetering het meest urgent is zien zelf met andere woorden de noodzaak daartoe niet en zijn ook niet gemotiveerd voor verbetering. Alleen de echte toppers blijken zich te *onderschatten*, waarschijnlijk doordat zij het beste inzicht hebben in wat excellent handelen vergt en omdat ze heel kritisch op hun eigen handelen zijn. Kruger en Dunning citeren Thomas Jefferson in dit verband: "*He who knows best, best knows how little he knows*" (p. 1121).

Robert Dijkgraaf refereerde vorig jaar in een NRC-column aan een studie waaruit zou zijn gebleken dat hoogleraren ook super bedreven zijn in zelfoverschatting. In een bepaald, door mij niet te achterhalen onderzoek, zou maar liefst 97% van de professoren zichzelf als professioneel bovengemiddeld kwalificeren. Wij krijgen hier als Universiteit Twente van Neelie Kroes steeds te horen dat we onze bescheidenheid van ons moeten werpen, dus misschien kwam die 3% uit Twente.

Als we de stap naar het Nederlandse onderwijs maken dan leert onderzoek van de Inspectie van het Onderwijs (2013) ons dat de door Kruger & Dunning (1999) gesignaleerde verschijnselen ook hier gelden. Leerkrachten bleken in het onderzoek vrijwel allemaal een positief beeld van hun eigen lesgevende vaardigheden te hebben, vooral waar het de basisvaardigheden (o.a. uitleg leerstof, zorgen voor een taakgerichte sfeer) betrof. In het voortgezet onderwijs had 35% van de leerkrachten een reëel beeld van de eigen sterke en zwakke professionele vaardigheden. In het primair onderwijs was dit percentage 66%. Veel leerkrachten gaven wel aan hun complexere didactische vaardigheden, zoals differentiatie, te willen verbeteren. Verder bleken de scholen waarin leerkrachten werkten:

- slecht zicht te hebben op de sterke en zwakke kanten van het pedagogisch-didactisch handelen van hun leerkrachten;
- weinig planmatig te sturen op leerkrachtkwaliteit én weinig scholingsactiviteiten te richten op de pedagogisch-didactische vaardigheden van hun leerkrachten.

Niet echt een ideale context voor DP dus: zowel de professionals als hun werkgevers blijken slecht zicht te hebben op de mate waarin de professionele kernvaardigheden beheerst worden. Naar mijn mening onderstrepen deze uitkomsten het belang van goede, regelmatige feedback over de professionele competenties van leerkrachten.

Als zelfevaluaties door leerkrachten tekort schieten, hoe kunnen we dan wèl zo objectief mogelijk vaststellen waar de ruimte voor professionele verbetering van leerkrachten zich bevindt, als basis voor de verdere groei van leerkrachten?

De bepaling van de toegevoegde waarde van leerkrachten

We zouden de *added-value*, de toegevoegde waarde, van leerkrachten kunnen bepalen. Toegevoegde waarde hier gedefinieerd als datgene wat leerkrachten toevoegen aan het startniveau van hun leerlingen, dan wel, indien een voormeting ontbreekt, de mate waarin ze boven of onder het gemiddelde leerkrachtprestatieniveau presteren. Deze maat is bedoeld om faire vergelijkingen tussen leerkrachten mogelijk te maken, door niet slechts te kijken naar het prestatieniveau van leerlingen op een bepaald moment zonder dat duidelijk is in hoeverre de leerkracht daaraan heeft bijgedragen. Deze outputmaten hebben een grote aantrekkingskracht in de context van het afrekenen van leerkrachten. Tegelijkertijd zijn ze daarvoor op dit moment niet geschikt. Het aantal factoren waarvoor gecorrigeerd zou moeten worden om leerkrachten op een eerlijke wijze te kunnen vergelijken is groot (denk bijvoorbeeld aan de factoren van Hattie die we bespraken). De toegevoegde waarde scores van dezelfde leerkrachten blijken nogal te variëren tussen schooljaren. Dezelfde leerkracht kan het ene jaar het label 'zeer effectief' en het volgend jaar een veel negatiever label krijgen terwijl hij niet fundamenteel anders lesgeeft en de leerlingen waaraan hij lesgeeft evenmin veel anders zijn. Je zou mogen verwachten dat 'leerkrachteffectiviteit' in zekere mate een stabiel leerkrachtkenmerk is (Darling-Hammond, 2015). De grote uitdaging betreft het valide meten en isoleren van de leerkrachtfactor te midden van andere relevante factoren. De pretentie om leerkrachten tot op de millimeter nauwkeurig met elkaar te kunnen vergelijken (zoals bijvoorbeeld in prestatierangordes wordt gedaan) is niet te verdedigen.

Wat voor leerkrachtprofessionalisering ook belangrijk is is dat de maten voor de toegevoegde waarde geen *aangrijpingspunten* bieden voor het optimaliseren van leerkrachtvaardigheden, ze geven slechts aan hoe de leerlingen van leerkrachten presteren in vergelijking met de leerlingen van andere leerkrachten. Wát je bij tegenvallende prestaties als leerkracht niet goed deed (als dat al het geval was) en hoe je een betere leerkracht zou kunnen worden blijft in het geval van de maten voor toegevoegde ongewis.

Procesmaten van leerkrachten

Andere maten voor het functioneren van leerkrachten betreffen procesmaten, ze bieden informatie over de kwaliteit van bepaalde aspecten van het handelen van leerkrachten in de klas (sommige ook over de kwaliteit van de lesvoorbereiding). Die informatie kan afkomstig zijn van een observator (iemand van buiten de school dan wel een collega), die met een lesobservatie-instrument de lessen van leerkrachten beoordeelt. Ook kunnen we de percepties van *leerlingen* gebruiken om een beeld te krijgen van hun perspectief op de kwaliteit van de lessen die ze aangeboden krijgen.

Leerkrachten krijgen doorgaans erg weinig feedback over hun lessen. Het zou goed zijn om een instrumentarium te ontwikkelen waarmee leerkrachten regelmatig een quick scan kunnen krijgen van de diverse aspecten van hun lesgeven. De formatieve meting van proceskenmerken en het geven van feedback daarover aan leerkrachten kan een basis bieden voor de doorontwikkeling van professionele leerkrachtcompetenties. Dergelijke feedback moet tijdens het werken aan verbetering frequent zijn, anders is de mate van progressie niet te monitoren en is ook bijsturing tijdens het werken aan verbetering onmogelijk.

Met de feedback zou een beeld moeten worden gegeven van de *algemene pedagogisch-didactische vaardigheden*, zoals het zorgen voor een productief klassenmanagement, het duidelijk uitleggen van de lesstof, het afstemmen van de instructie op verschillen tussen leerlingen, het bevorderen van zelfregulatie door leerlingen en het bij leerlingen stimuleren van het inzicht in de verbanden tussen de centrale concepten binnen een vak. Voor het meten van de algemene pedagogisch-didactische *basisvaardigheden* beschikken we al over instrumenten, zoals het in Nederland veel gebruikte ICALT-instrument, dat beschikbaar is als lesobservatie-instrument en als instrument voor het meten van de percepties van leerlingen. Het is vreemd dat het belang van de *complexe* pedagogisch-didactische leerkrachtvaardigheden zo vaak en sterk benadrukt wordt, maar dat het ontbreekt aan instrumenten waarmee we valide kunnen bepalen hoe goed leerkrachten deze vaardigheden beheersen.

Lesobservaties

Lesobservaties worden steeds populairder in het Nederlandse onderwijs. Als ze goed worden uitgevoerd zijn ze arbeidsintensief en duur, onder andere doordat een valide oordeel over de didactische capaciteiten van een leerkracht een behoorlijk aantal beoordeelde lessen en beoordelaars vereist (3 á 4 van elk; Hill, Charalambous, Kraft, 2012). Ook zijn een degelijke training van de beoordelaars en het regelmatig kalibreren van reeds gecertificeerde beoordelaars vereist. Je wilt dat beoordelaars zoveel mogelijk overeenkomstig oordelen bij het zien van dezelfde les. Dat is minder eenvoudig dan het lijkt. Zelfs na een gedegen training vormen beoordelaars een bron van variantie in de scores die aan leerkrachten worden toegekend. Desalniettemin kunnen observaties leerkrachten wel waardevolle feedback bieden over waar zij vooral nog kunnen verbeteren. En idealiter kan een observator ook goed aangeven *hoe* de beoordeelde leerkracht kan verbeteren. Over wat effectieve manieren van coaching zijn moeten we echter nog veel leren.

Als de middelen beschikbaar zijn en we de mogelijkheden van lesobservaties willen benutten, dan is het werken met externe professionele beoordelaars denk ik kansrijker, dan het binnen scholen beoordelen van leerkrachten door collega-leerkrachten, of door de schoolleider. Dit in verband met de complexiteit van de beoordelings- en coachingactiviteit (en de training die daarvoor vereist is) en de kans dat een extern observator waarschijnlijk objectiever zal oordelen. Veel van de landelijke initiatieven op dit terrein leunen echter nogal op collegiale lesobservaties. Het zou interessant zijn om te onderzoeken of met de samenwerking tussen extern opgeleide observatoren en leerkrachten binnen scholen twee doelen bereikt kunnen worden: de opleiding van docenten binnen scholen tot gecertificeerde lesbeoordelaars én de verbetering van het lesgeven van leerkrachten. Leerkrachten zouden dan misschien op andere scholen dan hun eigen scholen leerkrachten moeten beoordelen en coachen.

Lesobservatie-instrumenten verschillen: ze kunnen algemeen zijn (d.w.z. niet specifiek voor het gebruik bij één vak) of vakspecifiek. Ze kunnen op de beoordeling van uiteenlopende aspecten van het lesgeven gericht zijn (zie hierboven voor voorbeelden), op duidelijk zichtbare en beoordeelbare aspecten van het lesgeven gericht zijn (low inference, bijvoorbeeld het klassenmanagement door de leerkracht) dan wel veel aan de interpretatie door de beoordelaar overlaten (high inference, bijvoorbeeld de mate waarin leerlingen de uitleg snappen). Marjoleine Dobbelaer heeft voor haar binnen de Academische Werkplaats van de Inspectie van het Onderwijs

uitgevoerde dissertatiestudie wereldwijd alle lesobservatie-instrumenten die aan bepaalde criteria voldeden opgespoord en gaat die onderwerpen aan een evaluatie van de mate waarin ze aan uiteenlopende validiteitscriteria voldoen.

Leerlingperceptie-instrumenten

Het meten van leerlingpercepties van leskwaliteit is een werkwijze die ook aan populariteit wint. Leerlingen kunnen in *algemene zin* een oordeel geven over de mate waarin de lessen van hun leerkrachten doorgaans voldoen aan bepaalde *algemeen pedagogisch-didactische* kwaliteitsaspecten (leerlingen kunnen dit ook voor *één* les doen; daarover straks meer). Leerlingen zijn er immers (bijna) altijd wanneer de leerkracht lesgeeft. Dus niet slechts drie á vier lesobservaties, maar veel observaties door veel beoordelaars, die weinig middelen vereisen. Dat zijn de grote voordelen van het werken met leerlingpercepties. De oordelen van de leerlingen in de klas kunnen wel verschillen, bijvoorbeeld als gevolg van verschillen in onderwijsbehoeften die leerlingen hebben (bij lesobservaties streven we vooral naar overeenstemming tussen beoordelaars).

Een element dat naast de algemeen pedagogisch-didactische kwaliteitsaspecten niet mag ontbreken in een quick scan instrumentarium is de meting van de percepties van leerlingen m.b.t. de *interactie tussen hen en hun leerkracht*. Het werk van Wubbels c.s. (de VIL, de Vragenlijst Interpersoonlijk Leraargedrag; Brekelmans, 2010) heeft geleid tot een mooi instrumentarium voor het meten van dit belangrijke aspect van lesgeven waarvan de samenhang is aangetoond met hoeveel leerlingen leren.

De mogelijkheden die de technologie biedt

Technologie kan tegenwoordig een belangrijke rol vervullen bij het bieden van feedback aan leerkrachten over de kwaliteit van hun lesgeven. Voor het maken van lesobservaties zijn opname-kits beschikbaar waarmee gemakkelijk lesopnames kunnen worden gemaakt (in principe kan de leerkracht het zelf doen). De opnames kunnen gemakkelijk in een beveiligde internet-omgeving worden opgeslagen, zodat leerkrachten zelf de opnames van hun lessen kunnen bestuderen, beoordelaars op een moment dat hen schikt de opgenomen lessen kunnen beoordelen en coach en leerkracht, indien wenselijk op afstand, samen de lessen kunnen bekijken en bespreken.

De technologie heeft nog meer te bieden. Zo hebben we voor het meten van leerlingpercepties recentelijk de IMPACT! app ontwikkeld waarmee leerlingen aan het eind van een les in enkele minuten op hun smartphone kunnen aangeven in hoeverre de les naar hun mening voldeed aan die leskenmerken waarvan wetenschappelijk is vastgesteld dat ze bijdragen aan het leren van leerlingen: werd de lesstof volgens hen duidelijk uitgelegd, zorgde de leerkracht voor een goede werksfeer en een sfeer waarin leerlingen het durfden te zeggen wanneer ze de stof niet snapten, hield de leerkracht in de gaten of de leerlingen de lesstof goed begrepen, zette de leerkracht leerlingen met vragen aan tot nadenken over de lesstof, zorgde hij/zij ervoor dat leerlingen die de stof niet snapten het daarna alsnog gingen snappen? et cetera. In totaal zestien vragen waarbij de leerkracht ook zelf twee vragen kan toevoegen, bijvoorbeeld vragen over de lesstof, om te checken of de leerlingen de kern van de les goed begrepen hebben. Bij de IMPACT! app gaat het dus niet om de algemene kwaliteit van de lessen van een leerkracht, maar om de kwaliteit van één specifieke les. Het voordeel daarvan is dat de leerkracht direct feedback krijgt en die gemakkelijker kan relateren aan wat hij/zij precies deed in de les, dan wanneer leerlingen gevraagd worden om een algemeen oordeel te vellen over hoe een leerkracht zich meestal in de klas gedraagt. De feedback die de IMPACT! app aan leerkrachten aanbiedt wordt bovendien uitgesplitst naar drie groepen leerlingen: de oordelen van leerlingen die op het moment van dataverzameling benedengemiddeld, dan wel gemiddeld, of bovengemiddeld presteren voor het vak in kwestie (de leerkracht deelt de leerlingen in op basis van de cijfers van de leerlingen voor het vak in kwestie). De drie leerlingengroepen kunnen sterk van elkaar verschillen in hun oordeel over de mate waarin de les aansloot bij wat ze nodig hadden. De uitgesplitste oordelen zijn een stuk informatiever, dan een gemiddeld oordeel van de hele klas; dezelfde gemiddelden kunnen immers hele verschillende situaties representeren. Hoewel de IMPACT! app leerlingen wel de mogelijkheid biedt om een verbeterpunt aan de leerkracht te geven zal de feedback van leerlingen vaak onvoldoende zijn om te verbeteren, met name wanneer de leerkracht laag scoort op complexe, moeilijk te leren didactische vaardigheden.

De menselijke factor is bij de benutting van de technologie van grote betekenis voor de effecten van wat we technologisch kunnen. Zo is het bijvoorbeeld de vraag hoe gemotiveerd leerlingen op langere termijn zullen blijven om hun leerkracht van feedback te voorzien. Mogelijk is die bereidheid groter naarmate ze meer impact van hun feedback zien. Als leerlingen m.a.w. niet ervaren dat de leerkracht op basis van de feedback iets verandert aan zijn lessen, dan is de kans groot dat hun bereidheid om mee te werken afneemt. Anderzijds kan hun betrokkenheid bij het onderwijs vergroten als leerlingen zien dat de leskwaliteit door hun feedback verbetert. De acceptatie door leerkrachten van de feedback die leerlingen geven, of de feedback die een lesobservator geeft, die werkt met een digitale tool voor lesopnames is een andere voorwaarde voor verbetering. Hetzelfde geldt voor de motivatie om aan verbetering te willen werken. En als een leerkracht graag wil verbeteren kan dat gewoon moeilijk zijn, omdat de te verwerven vaardigheid complex is. Zo zal goed leren differentiëren in de klas onmogelijk zijn zonder een gedegen training. Feedback is m.a.w. goed om erachter te komen waar je kunt verbeteren, maar daarna moet nog aan diverse voorwaarden worden voldaan om daadwerkelijk te kunnen verbeteren.

We bestuderen in het IMPACT! app project nu of leerlingen door de app meer betrokken raken bij hun onderwijs, of de IMPACT! feedback leerkrachten aanzet tot reflectie over hun lesgeven en tot verbeteracties en natuurlijk ook of de leskwaliteit door het gebruik van de app verbetert.

Zoals bij elke meting zijn we ook bij het werken met leerlingpercepties niet gevrijwaard van potentiële validiteitsbedreigers, zoals de subjectieve voorkeuren van leerlingen. Die voorkeuren kunnen samenhangen met de persoon en het geslacht van de leerkracht evenals met de achtergrondkenmerken van leerlingen (hun geslacht, sociaaleconomische status, hoe goed ze in een vak zijn e.d.). We onderzoeken daarom nu in hoeverre de lesoordelen die leerlingen geven samenhangen met achtergrondkenmerken van leerkrachten en leerlingen. Tevens analyseren we in hoeverre verschillende perspectieven (leerkrachtzelfevaluaties, leerlingpercepties d.m.v. de IMPACT! app, lesobservaties door externe observatoren) tot vergelijkbare resultaten leiden wanneer daarbij dezelfde items gebruikt worden (de items zijn qua formulering alleen aangepast aan wie het item moet beantwoorden).

De combinatie van lesobservaties en leerlingpercepties

Ik denk dat de combinatie van het leerlingperspectief en het perspectief van de getrainde externe observator de leerkracht waardevolle feedback kan bieden over waar zijn/haar verbeteringsruimte zich vooral bevindt en, na interventies, over in hoeverre ondernomen verbeteringsactiviteiten effectief zijn. Beide perspectieven kunnen elkaar aanvullen. De leerling vertegenwoordigt het oordeel van de doelgroep (snapt en leerlingen bijvoorbeeld de uitleg?), de observator evalueert een leerkracht in het licht van professioneel standaarden. Voor bepaalde aspecten van het lesgeven zijn beide perspectieven interessant (bijvoorbeeld als het gaat om de kwaliteit van de uitleg van leerstof), voor andere is het perspectief van de leerling relevanter (bijvoorbeeld of leerlingen vinden dat er sprake is van een veilig klasseklimaat en of ze het gevoel hebben dat de leerkracht hoge verwachtingen van hen heeft), of juist het perspectief van de observator (beheerst de leerkracht bijvoorbeeld de vakdidactiek van een vak en zet hij leerlingen aan tot zelfregulering?).

De balans opmakend

Aan het eind van deze oratietekst concludeer ik dat we moeten blijven zoeken naar manieren om leerkrachten nog beter toe te rusten voor hun zeer complexe taak. Leerkrachten zo opleiden, dat ze sneller de basisvaardigheden onder de knie kunnen krijgen en ook sneller effectief kunnen zijn. Tevens is het van belang dat leerkrachten minder gaan variëren in hun impact op leerlingen en dat ze kunnen doorgroeien, daar waar ze nu na een aantal jaren nog tegen hun natuurlijke plafond aanlopen. Het zou fantastisch zijn als de beheersing van de complexe didactische vaardigheden zo binnen het bereik van alle leerkrachten komt.

Naar mijn mening is voor het bereiken van dat doel het volgende belangrijk:

1. De ontwikkeling van een *instrumentarium* waarmee leerkrachten snel en regelmatig valide formatieve feedback over hun professionele vaardigheden kunnen krijgen. De mogelijkheden die de technologie biedt moeten we daarbij zeker benutten. Op basis van de feedback kunnen leerkrachten zelf beslissen waaraan ze gericht gaan werken.
2. Aandacht voor de *menselijke factor* in termen van de vereiste kennis, vaardigheden en attitudes voor de benutting van de technologische mogelijkheden.

3. Een grondige empirische *analyse en definitie van de kerncompetenties van leerkrachten*, als basis voor het ontwerp van een *doorlopende leerlijn* voor leerkrachten én het ontwerp van de *training* in die kerncompetenties op basis van solide *leertheoretische* uitgangspunten. Binnen die leerlijn dienen de complexe didactische vaardigheden de aandacht te krijgen die ze verdienen. Ik nodig alle Nederlandse lerarenopleidingen uit voor zo'n exercitie op basis van het 4C/ID-model.
4. Een *schoolorganisatie* die het gericht ontwikkelen van leerkrachtkwaliteiten stimuleert en faciliteert en een *schoolcultuur* waarin de continue professionele ontwikkeling van leerkrachten geen signaal is van disfunctioneren, maar een vanzelfsprekendheid in verband met het belang van de leerling.
5. Een *intelligente overheid* overheid die het belang van het hier voorgestelde voor leerling, leerkracht en maatschappij inziet en om die reden voorziet in de daarvoor benodigde financiële middelen.

TOT SLOT

Mijn dank gaat uit naar het College van Bestuur van de Universiteit Twente voor mijn benoeming. Professor Theo Toonen, de decaan van de faculteit BMS heeft daarbij een belangrijke rol vervuld, door kort na zijn eigen start hier, ruim twee jaar geleden, de vacature zeer onconventioneel in te vullen met twee kandidaten daarvoor. Volgens mij doen Susan en ik ons best om te laten zien dat dat een verstandige beslissing was!

Susan, ik had nooit verwacht dat ik het zo fijn zou vinden dat mijn concurrent in een sollicitatieprocedure benoemd zou worden, nou, ..., laat ik niet overdrijven, dat die concurrent óók benoemd zou worden. Vijfentwintig jaar geleden was jij een van de eerste masterstudenten hier en ik nog een jonge docent. Niemand had toen durven voorspellen dat we hier nu samen onze oraties zouden uitspreken. We zijn heel verschillend, maar samen en met de andere leden van het management team van ELAN vormen we een goed team. Je persoon, je drive en professionele competenties maken het fijn om met je samen te werken. Ik zie de komende jaren dan ook met plezier en vertrouwen tegemoet.

Ik wil Cees Glas en de andere medewerkers van de vakgroep OMD hartelijk bedanken voor de opname in hun vakgroep toen we onderdak zochten. Daardoor konden we toen verder met de Focus-projecten en andere projecten, werd de data-analyse van die projecten sophisticated en werken we nu als vakgroepen gelukkig nog steeds veel samen.

Aan de jarenlange samenwerking met Roel Bosker en Jaap Scheerens dank ik het effectiviteitsperspectief op onderwijs. Of het nu om het handelen van schoolleiders, schoolteams of individuele leerkrachten gaat, de vraag onder welke condities *leerlingen* daarvan het meest profiteren is de belangrijkste. Dat uitgangspunt is minder vanzelfsprekend dan het lijkt. Gelukkig kunnen we er hier niet over discussiëren.

Bij ELAN heb ik me vanaf het begin af aan welkom gevoeld, gelukkig maar, stel je voor... Ik ben trots op ELAN als een dynamische, productieve vakgroep met een prettig sociaal klimaat. De combinatie in één vakgroep van lerarenopleiders en onderzoekers is er eentje die

ik elke onderwijskundevakgroep toewens. De kruisbestuiving tussen lerarenopleiders en onderzoekers is een vruchtbare, we hebben veel aan en leren veel van elkaar. Onderzoeksresultaten vinden snel hun weg naar de lerarenopleiding en onderwijspraktijk en de onderzoekers doen beter onderzoek door de inbreng van de lerarenopleiders.

De collega's van het Pre-University programme en ELAN zitten fysiek dicht bij elkaar en zo hoort het ook. Onze contacten met 'Den Haag' zijn uitstekend en we weten elkaar gemakkelijk te vinden wanneer het inhoudelijk belangrijk is dat we samen optrekken, of als Pieter Boerman of ik menen dat we elkaar iets grappigs te vertellen hebben.

ELAN zit dan wel op de hoogste verdieping van het Ravelijn gebouw, maar niet in een ivoren toren. Toen Susan en ik in december 2015 startten hebben we een kennismakingsrondje langs de bètafaculteiten gemaakt. We waren beiden onder de indruk van hoe goed de onderlinge contacten met ELAN waren en hoe de bètafaculteiten het belang van ELAN onderstreepten.

Vooraf door de lerarenopleiding zijn de contacten en samenwerking van ELAN met de scholen in de regio uitstekend. Speciaal noemenswaardig is hier de samenwerking met de Stichting Carmel waarmee ELAN al lang intensief samenwerkt en waarmee de samenwerking recentelijk is geformaliseerd. We hebben inmiddels een mooie kennisagenda geformuleerd waaraan we de komende jaren hard gaan werken.

In het hedendaagse hoger onderwijs krijgt hoger *onderzoek* vaak meer aandacht, dan hoger *onderwijs*. Een goede balans in de aandacht en waardering voor beiden is echter van groot belang. Via het bij ELAN ondergebrachte 4TU.Centre for Engineering Education werken we met nationale en internationale partners daarom ook aan de versterking van de kwaliteit van het hoger onderwijs.

Ik hoop dat u vandaag in het ochtend- en middagprogramma een indruk heeft gekregen van onze ELAN-projecten. We hebben echter nog meer, onder andere ruim 25 promovendi die hard aan hun dissertaties werken. Samen met scholen en de Inspectie van het Onderwijs werken we in een Academische werkplaats aan projecten op het gebied van differentiatie en lesobservatie.

En met de vakgroepen Onderzoeksmethodologie Meetmethoden en Data-Analyse, Instructietechnologie en Cognitieve Psychologie & Ergonomie vormen we, onder andere blijkend uit recente internationale rankings en andere beoordelingen, een sterk onderzoekscluster. We doen daarbij onderzoek op het raakvlak van leren en technologie en houden u daarvan de komende jaren graag op de hoogte.

Ik wil u allen tot slot hartelijk bedanken voor uw komst. Het was een hele zit vandaag en een behoorlijk aantal van u was 1,5 jaar geleden ook al in Groningen aanwezig. Ik beloof het u, het zal niet weer gebeuren!

Ik heb gezegd.

Fer Coenders, Anja Knuver, Hans Luijten, Jan van der Meij en Jan van der Veen wil ik hartelijk bedanken voor hun feedback op de conceptversie van (een deel van) deze oratie!

REFERENTIES

- Algemene Rekenkamer (2016). Aanpak van laaggeletterdheid. Den Haag: Algemene Rekenkamer.
- Berliner, D.C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research* 35, 463-482.
- Brekelmans, M. (2010). *Klimaatverandering in de klas*. Oratie. Utrecht: Universiteit Utrecht.
- Darling-Hammond, L. (2015). Can Value Added Add Value to Teacher Evaluation? *Educational Researcher*, Vol. 44, No. 2, pp. 132-137.
- Day, C. (2008). Committed for life? Variations in teachers' work, lives and effectiveness. *Journal of Educational Change*, (9) 243-260.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge handbook of expertise and expert performance*, 38, 685-705.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363.
- Ericsson, A. & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.
- Eva, K. W., & Regehr, G. (2005). Self-assessment in the health professions: a reformulation and research agenda. *Academic Medicine*, 80(10), S46-S54.
- Goldhaber, D. D., Brewer, D. J., & Anderson, D. J. (1999). A three-way error components analysis of educational productivity. *Education Economics*, 7(3), 199-208.

- Haertel, E.H. (2013). *Reliability and Validity of Inferences about Teachers Based on Student Test Scores*. Washington: ETS Research & Development & Princeton: Center for Research on Human Capital and Education.
- Hambrick, D.Z., Brooke N. Macnamara, G. Campitelli, F.U. & Mosing, M. A. (2016). Beyond born versus made: a new look at expertise. In: *Psychology of Learning and Motivation*, volume 64, pp. 1-55.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child development*, 76(5), 949-967.
- Hanushek, E. & Rivkin, S.G. (2006). Teaching Quality. *Handbook of the Economics of Education*, Volume 2, 1052-1078.
- Hanushek, E. & Rivkin, S.G. (2013). The distribution of teacher quality and implications for policy. *Annual review of Economics*, 4, p. 131-159.
- Hattie, J.A.C. (2003). Teachers make a difference: What is the research evidence? Paper presented at the Building Teacher Quality: What does the research tell us ACER Research Conference, Melbourne, Australia. Retrieved from http://research.acer.edu.au/research_conference_2003/4/
- 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, p. 178-204.
- Hendriks, M. A., & Scheerens, J. (2013) School leadership effects revisited: a review of empirical studies guided by indirect-effect models. *School leadership and management*, 33(4), p. 373-394.
- Hill HC, Charalambous CY, Kraft MA. When rater reliability is not enough: teacher observation systems and a case for the G-study. *Educational Researcher*, 41(2), p. 56-64.

- Huberman, M. (1993). *The lives of teachers*. London: Casell.
- Inspectie van het Onderwijs (2013). *Professionalisering als gerichte opgave. Verkennend onderzoek naar het leren van leraren*. De Meern: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2016). *De Staat van het Onderwijs 2014/2015*. De Meern: Inspectie van het Onderwijs.
- Jansen in de Wal, J. (2016). *Secondary school teachers' motivation for professional learning*. Proefschrift. Heerlen: Open Universiteit.
- Kini, T. & Podolsky, A. (2016). *Does Teaching Experience Increase Teacher Effectiveness? A Review of the Research*. Palo Alto: Learning Policy Institute.
- Kruger, J. & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, Vol 77(6), p. 1121-1134.
- Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review*, 29(3), 480-488.
- Luyten, J.W., Scheerens, J., Visscher, A.J., Maslowski, R., Witziers, B. & Steen, R. (2005). *School factors related to quality and equity. Results from PISA 2000*. Parijs: OECD.
- Luyten, H., Schildkamp, K., & Verachtert, P. (2009). Vooruitgang in technisch lezen gedurende het schooljaar en de zomervakantie. *Pedagogische Studiën*, 86(3), 196-213.
- Meelissen, M.R.M. & Luyten, H. (2011). Schooleffectiviteit en prestatieniveau natuuronderwijs in groep 6: secundaire analyses op TIMSS-2007 data. *Pedagogische Studiën* (88), p. 309-322
- Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice hall.

- Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges (2004). "How large are teacher effects?." *Educational Evaluation and Policy Analysis*, 26(3), p. 237-257.
- Platform Beleidsinformatie (2014). *Uitstroom van beginnende leraren*. Den Haag: Ministerie van Onderwijs Cultuur en Wetenschappen.
- Quint, J. C., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (2015). *Scaling Up the Success for All Model of School Reform*. Final Report from the Investing in Innovation Evaluation. MDRC: New York/Oakland.
- Van der Grift (2010). *Ontwikkeling in de beroepsvaardigheden van leraren*. Inaugurale Rede. Groningen: Rijksuniversiteit Groningen.
- Van Merriënboer, J.G. & P.A. Kirschner (2013). *Ten steps to complex learning: a systematic approach to four-component instructional design*. New York/London: Routledge.
- Verachtert, P., Van Damme, J., Onghena, P. & Ghesquière, P. (2009). A seasonal perspective on school effectiveness: Evidence from a Flemish longitudinal study in kindergarten and first grade. *School Effectiveness and School Improvement*, 20(2), 215-233.
- Voogt, J. & Pareja Roblin, N. (2010). *21st Century skills*. Enschede: Universiteit Twente.
- William, D. (2016). Teacher Learning: the key to improving the world. Retrieved on January 31, from: http://dylanwilliam.org/Dylan_Williams_website/Presentations.html

