

# Ontwikkeling in de pedagogisch didactische vaardigheid van leraren in het basisonderwijs

W. van de Grift, M. van der Wal en M. Torenbeek

## Samenvatting

De activiteiten van leraren die een uitdrukking vormen van hun pedagogisch didactische vaardigheid zijn geobserveerd bij steekproeven basisscholen in Nederland, Vlaanderen, Duitsland, Slowakije, Kroatië en Schotland. In deze activiteiten blijkt een hiërarchie te bestaan die loopt van het creëren van een veilig en stimulerend onderwijsleerklimaat naar efficiënt klasmanagement, gevolgd door duidelijke en gestructureerde instructie, het intensiveren van de les, het afstemmen van het onderwijs op verschillen tussen leerlingen en het leerlingen aanleren hoe ze iets moeten leren. Uit kleinschalig cross-sectioneel onderzoek onder Duitse leraren blijkt dat de hoeveelheid arbeidservaringjaren aanvankelijk gelijk oploopt met het niveau in pedagogisch didactische vaardigheid. Na ongeveer het 20e beroepsveringjaar treedt bij de gemiddelde leraar een geleidelijke daling op in het niveau van pedagogisch didactische vaardigheid.

## 1 Inleiding

Hanushek en Zhang (2006) analyseerden data van 13 landen uit de 'International Adult Literacy Survey'. Een toename van een standaarddeviatie op geletterdheid blijkt gepaard te gaan met een stijging in jaarlijks inkomen van gemiddeld ruim 9%. Voor Nederland was dit zelfs bijna 15%. Hanushek (2011) beargumenteerde dat de gemiddelde leerling van een leraar waarvan de kwaliteit van het pedagogisch didactisch handelen een standaarddeviatie hoger ligt dan van zijn collega's, per jaar gemiddeld 20.000 dollar meer zal gaan verdienen. Goede leraren zijn daarmee een belangrijke economische factor. Dat alleen al maakt inzicht in de ontwikkeling van het professionele niveau van leraren belangrijk.

## 2 Theoretisch kader

Het theoretisch kader van dit onderzoek bestaat uit kennis over vaardigheden waarover leraren moeten beschikken om bij leerlingen betrokkenheid en leerwinst te bereiken en uit kennis over de ontwikkeling die leraren tijdens hun beroepsloopbaan doormaken.

Prestaties van leerlingen worden sterk bepaald door leerlingenkenmerken zoals intelligentie, socio-economische achtergrond en motivatie (Coleman, et al., 1966; Hattie, 2009). Dat neemt niet weg dat ook de kwaliteit van schoolleiding en leraren belangrijk is. Leraren in het basisonderwijs kunnen afhankelijk van de leerstof, 15-30% van de verschillen in prestaties van leerlingen beïnvloeden (Bosker & Witziers, 1996; Brandsma & Knuver, 1989; Hattie, 2009; 2007a; Houtveen & Van de Grift, 2007b; Wijnstra, Ouwens, & Béguin, 2003). In deze 15-30% ligt de beroepsruimte van de leraar. Deze beroepsruimte wordt niet altijd volledig benut. Sommige leraren boeken bij hun leerlingen in een jaar tijd anderhalf jaar leerwinst, andere niet meer dan een half jaar (Hanushek, 1992; Houtveen, Mijs, Vernooij, Van de Grift & Koekebacker, 2003).

Uit meer dan 40 jaar onderzoek naar de effectiviteit van leraren blijkt dat de volgende activiteiten van leraren effect sorteren als het gaat om prestaties of leerwinst van leerlingen: een veilig en stimulerend onderwijsleerklimaat, efficiënt klasmanagement, heldere en duidelijke instructie, alle leerlingen bij de les betrekken, afstemming van het onderwijs op verschillen tussen leerlingen, leerlingen leren hoe ze iets moeten leren, voldoende gelegenheid bieden tot leren, monitoren van leerprestaties, en speciale maatregelen voor leerlingen die achterop dreigen te komen (Voor reviews zie: Cotton, 1995; Creemers, 1994; Ellis & Worthington, 1994; Hattie, 2009; Levine & Lezotte, 1990; Levine & Lezotte, 1995; Marzano, 2008; Purkey, &

Smith, 1983; Sammons, Hillam, & Mortimore, 1995; Scheerens, 1992; Scheerens & Bosker, 1997; Walberg & Haertel, 1992). Van de Grift (1985, 2007) en Van de Grift en Lam (1998) hebben gedetailleerd uitgewerkt welke van de bovengenoemde activiteiten van leraren in welke oorspronkelijke studies samenhang bleken te vertonen met meer betrokkenheid en/of leerwinst van leerlingen (al dan niet gecorrigeerd voor leerlingkenmerken).

Enkele observatiestudies brengen een hiërarchie in activiteiten van leraren aan het licht. Schaffer, Stringfield en Wolfe (1992) observeerden 45 leraren in het begin van hun beroepsuitoefening. Zij ontdekten dat leraren gedurende hun eerste jaar significante vooruitgang boekten in klasmanagement. De vooruitgang in het tweede jaar betrof meer complexe onderwijsactiviteiten. Een observatiestudie in Nederland (Van de Grift & Lam, 1998) toont een opeenvolging in het pedagogisch didactisch handelen die begint met het creëren van een veilig en stimulerend onderwijsleerklimaat. Het volgende niveau is efficiënt klasmanagement, gevolgd door duidelijke en gestructureerde instructie. Hogere niveaus zijn het intensiveren van de les, het activeren van leerlingen en het onderwijs goed afstemmen op verschillen tussen hun leerlingen. Van de Grift en Lam (1998) vonden aanwijzingen dat leerlingen van leraren waarvan de vaardigheid zich op een hoger punt in de hiërarchie bevond, een grotere betrokkenheid bij de les vertoonden en ook iets hogere prestaties hadden bij lezen, rekenen en wereldoriëntatie. Een toename in complexiteit in pedagogisch didactische vaardigheid gaat kennelijk gepaard met meer betrokkenheid en hogere prestaties van leerlingen. Een observatiestudie van Kyriakides, Creemers en Antoniou (2009) doet eveneens vermoeden dat er een opeenvolging bestaat in de vaardigheid van leraren die loopt van beheersing van zeer basale vaardigheden zoals zorgen voor een veilig en stimulerend onderwijsleerklimaat, efficiënte lesorganisatie en duidelijke gestructureerde instructie naar meer complexe vaardigheden, zoals het intensiveren van de les, het onderwijs afstemmen op verschillen en het leerlingen leerstrategieën aanleren. Deze drie studies laten ook zien dat sommige activiteiten moeilijker of

complexer zijn dan andere, en dat leraren die moeilijker of complexer activiteiten vertonen, doorgaans ook eenvoudiger of minder complexe activiteiten vertonen.

Het is de vraag of deze hiërarchie in activiteiten van leraren stand houdt bij grote steekproeven van leraren uit verschillende landen en of deze hiërarchie verband houdt met de ontwikkeling die leraren in hun beroepsloopbaan doormaken. Er zijn theorieën die de ontwikkeling van leraren beschrijven in termen van lineaire groei en er zijn theorieën waarin de ontwikkeling van leraren geen monotone stijging of zelfs een parabolisch verloop lijkt te hebben.

Theorievorming naar de ontwikkeling in de beroepsloopbaan van leraren startte met een studie van Fuller (1969, 1970) over wat zij noemde de ontwikkeling in 'concerns' van leraren. Volgens Fuller zouden leraren een ontwikkeling doormaken die startend bij zorgen over zichzelf, via zorgen over hun taakvervulling, overgaat naar zorgen over de impact die zij hebben op hun leerlingen. In de eerste fase gaat het nog niet zozeer om zorgen over het pedagogisch didactisch handelen als wel om concerns van het type: 'Is het leraarschap wel wat voor mij?' en 'Kan ik wat voor het onderwijs betekenen?'. Deze 'self concerns' fase maken de studenten van de lerarenopleidingen door en leiden doorgaans tot een besluit om hun opleiding al dan niet te vervolgen. Geleidelijk ontstaan concerns over het beroep. De tweede fase betreft zorgen over kennis van de leerstof en vooral ook over vaardigheden zoals klasmanagement. Deze 'task concerns' blijven gedurende de opleiding en tijdens het beginnend leraarschap een rol spelen. In de derde fase gaat het om zorgen over de relatie met de leerlingen en over wat de leerlingen geleerd wordt: 'Wat hebben leerlingen nodig en wat moet gedaan worden om leerlingen tot ontwikkeling te brengen?' Deze 'impact concerns' gaan pas een rol spelen als de 'task concerns' achter de rug zijn.

De observatie van Schaffer et al. (1992) dat de ontwikkeling van klasmanagementvaardigheden vooraf gaat aan het ontwikkelen van complexere vaardigheden is consistent met de theorie van Fuller. Dat geldt ook voor de bevindingen van Van de Grift en

Lam (1998) en van Kyriakides et al. (2009). In Fuller's idee van task concerns zien we het onderwijsleerklimaat, het klasmanagement en de basale instructievaardigheden van Van de Grift e.a. en Kyriakides e.a. terug en in Fuller's idee van impact concerns kunnen we het activeren van leerlingen, het inspelen op verschillen tussen leerlingen en het leerlingen leren hoe ze iets moeten leren van Van de Grift e.a. en Kyriakides e.a. herkennen.

Andere inspiratiebronnen voor theorieontwikkeling over een lineaire ontwikkeling in de beroepsloopbaan van leraren zijn te vinden in het model van Dreyfus en Dreyfus (1986). Zij onderscheiden vijf niveaus, lopend van 'novice', via 'advanced beginner', 'competent' en 'proficient' naar 'expert'. Bij Berliner (1994, 2001) zien we een vergelijkbaar model dat loopt van novice, advanced beginner, competent performer, proficient performer naar expert. Een andere variant is het zesfasenmodel van Steffy en Wolfe (2001) dat loopt van 'novice', 'apprentice', 'professional', 'expert', 'distinguished' naar 'emeritus'. Er is maar weinig empirisch materiaal waaruit blijkt dat de ontwikkeling in de beroepsloopbaan van leraren daadwerkelijk zo lineair en zo voorspelbaar verloopt. Ook is het maar de vraag of het aantal ervaringsjaren de enige motor is van een dergelijke ontwikkeling of dat ook andere factoren inwerken op de ontwikkeling van de leraar. Steffy en Wolfe (2001) veronderstellen dat de ontwikkeling van leraren wordt gestimuleerd door reflectie, gevolgd door het invoeren van vernieuwingen die, wanneer het verbeteringen blijken, doorgezet worden, en als dat niet het geval is verlaten worden. Dit inspirerende idee over de 'reflective practitioner' zagen we al eerder bij Schön (1983, 1987). Joyce en Showers (2002) beschrijven een hierop lijkend iteratief proces van leren en experimenteren bij het ontwikkelen van nieuwe vaardigheden door leraren.

Naast reflectie lijkt ook de identiteit van leraren een rol te spelen in de ontwikkeling van de beroepsloopbaan van leraren. Leraren moeten niet alleen kennis en vaardigheden ontwikkelen, maar moeten ten opzichte van hun leerlingen ook een pedagogische rol vervullen. Dat vereist de ontwikkeling van een persoonlijke identiteit (Canrinus, Helms-

Lorenz, Beijaard, Buitink & Hofman, 2011; Dreyfus & Dreyfus, 1986; Eraut, 1994; Feiman-Nemser, 1983; 2001; Hammerness, Darling-Hammond & Bransford, 2005; Huberman, 1989; Richardson & Placier, 2001; Sprinthall, Reiman & Theis-Sprinthall, 1996).

Er zijn ook theorieën die een niet-monotoon stijgende of zelfs parabolische ontwikkeling van leraren beschrijven. Burke, Christensen, Fessler en Price (1987) beschrijven de ontwikkeling van leraren met een parabool. Zij zien een groei van 'pre-service, induction, competency building, enthusiastic and growing', waarna een neerwaartse ontwikkeling volgt met 'career frustration, stable, stagnant and career exit'. Lynn (2002) heeft dit model uitgebreid met twee verklarende cycli die respectievelijk de persoonlijke en organisatorische omgeving betreffen. Day (2008) beschrijft zes stadia in het professionele leven van leraren. In het eerste stadium (0-3 jaar ervaring) ligt de nadruk op het ontwikkelen van een gevoel van doelmatigheid in de klas. De betrokkenheid van de leraar is groot in deze fase. Het tweede stadium (4-7 jaar) wordt gekenmerkt door toename van zelfvertrouwen om een effectieve leraar te worden. In stadium drie (8-15 jaar) vinden veranderingen plaats in de rollen en identiteit van de leraar. In het vierde stadium (16-23 jaar) krijgen de leraren te maken met spanningen tussen werk en privé. In het vijfde stadium (24-30 jaar) wordt de motivatie van leraren sterker op de proef gesteld. Leraren gaan extern beleid en externe initiatieven als negatief ervaren. In het zesde stadium (na het 31<sup>e</sup> jaar) komt het pensioen dichterbij. In de stadia twee tot en met zes weet de meerderheid van de leraren een sterke mate van identiteit, 'self-efficacy', motivatie en betrokkenheid te behouden, maar een deel lukt dit niet en kampt met desinteresse en verminderde betrokkenheid. Gonçalves (2009) ziet vergelijkbaar hiermee vijf stadia in de loopbaan van leraren. De eerste fase duurt ongeveer vier jaar en betreft de inductie in het beroep. De tweede fase (5-7 jaar ervaring) is de fase waarin de leraar vertrouwen en plezier heeft en tevreden is. In fase drie (8-14 jaar) zijn er, vergelijkbaar met de studie van Day (2008), leraren die toegewijd en enthousiast blijven,

en leraren die dit niet lukt en die routinematig gaan werken. De vierde fase (15-25) kenmerkt zich door gevoelens van tevredenheid en overtuigd zijn van zichzelf. In de laatste fase zijn opnieuw twee groepen leraren te onderscheiden; leraren die enthousiast blijven en bereidwillig blijven om te leren en leraren die niet langer enthousiasme vertonen en uitzien naar hun pensioen.

Afgezien van het tijdsverloop, zijn leren van de ervaring en identiteitsontwikkeling kennelijk de verklaring voor een lineair stijgende ontwikkeling. In het denken over de niet-monotoon stijgende ontwikkeling van leraren zien we aandacht ontstaan voor problemen van leraren in of na het midden van hun carrière (Hargreaves, 2005; Holloway, 2003; Huberman, 1989; Rosenholtz & Simpson, 1990). Deze verklaringen vallen voor een deel onder de 'mid life crisis' (Kompier & Cooper, 1999) die onafhankelijk is van het leraarsberoep. Echter rondom het 20<sup>ste</sup> beroepservaringjaar vinden niet alleen in het persoonlijke maar ook in het professionele leven van leraren tamelijk ingrijpende veranderingen plaats. Sommige leraren krijgen meer lesoverstijgende taken zoals mentoraat, team- en of sectieleiderschap en dergelijke. Ze komen voor nieuwe uitdagingen te staan, die het zelfbeeld over hun functioneren tijdelijk uit het evenwicht kunnen brengen. De veilige zekerheid over de verworven vaardigheid in het lesgeven wordt ingeruild voor onzekerheid over vaardigheden zoals het kunnen beïnvloeden van collega's na het accepteren van een team-, of sectieleiderschap.

Ook de betrokkenheid of motivatie van leraren lijkt een rol te spelen in hun ontwikkeling. Rosenholtz en Simpson (1990) vonden dat de betrokkenheid van leraren groter is naarmate ze meer ervaring hebben (zes tot tien jaar), maar dat de betrokkenheid van leraren met meer dan tien jaar ervaring juist weer lager is. Fokkens-Bruinsma en Canrinus (2011) vonden echter een hoge mate van betrokkenheid bij ervaren leraren in vergelijking met leraren die minder ervaring hebben. Deze discrepantie is in overeenstemming met de overwegingen van zowel Day als Gonçalves dat een deel van de ervaren leraren een sterke mate van identiteit, 'self-efficacy' en motivatie weet te behouden, terwijl een ander

deel kampt met desinteresse en verminderde betrokkenheid. Behalve motivatie noemen Rosenholtz en Simpson (1990) ook de mate van autonomie en de vraag of leraren voldoende waardering ontvangen van collega's en schoolleiders als verklaring voor een neergaande ontwikkeling. Ook Day (2008) noemt de mate van steun vanuit de school, veranderingen in onderwijsbeleid, leerlinggedrag, werkdruk en familiale verplichtingen als verklaring voor stagnatie en neergaande ontwikkeling.

### 3 Onderzoeksopzet

Om de opleiding en verdere professionalisering van leraren te kunnen richten op hun zone van naaste ontwikkeling is het van groot belang om te weten of er een (wellicht universele) hiërarchie bestaat in de activiteiten van leraren die samenhangt met hun ontwikkeling. Dit brengt ons tot de volgende onderzoeksvragen:

- a. Bestaat er een hiërarchie in activiteiten van leraren die oploopt van het zorgen voor een veilig en stimulerend onderwijsleerklimaat naar het efficiënt klasmanagement, gevolgd door duidelijke en gestructureerde instructie, die voortgaat met het intensiveren van de les en die eindigt bij goed afstemmen van het onderwijs op verschillen tussen leerlingen en het leerlingen aanleren hoe ze iets moeten leren?
- b. Doet deze hiërarchie in pedagogisch didactische activiteiten zich voor bij leraren uit verschillende landen?
- c. Verloopt de ontwikkeling van de pedagogisch didactische vaardigheid van leraren in overeenstemming met deze hiërarchie?

Voor het beantwoorden van de eerste onderzoeksvraag construeren we een op het Rasch-model gebaseerd onderzoeksinstrument dat geschikt is om de genoemde deelvaardigheden te meten. De oorspronkelijk Nederlandse versie van het instrument is door onszelf vertaald in het Duits en het Engels. Deze vertalingen zijn gecorrigeerd door een 'native speaker', die ook over de Nederlandstalige versie beschikte. Vervolgens heeft tussen onderzoeker en vertaler overleg plaatsgevon-

den over betekenisvolle verschillen. De Slovaakse en Kroatische versie zijn door aan het onderzoek deelnemende inspecteurs die de Nederlandse taal machtig waren uit het Nederlands vertaald. Tijdens de trainingen in elk van de landen (zie hieronder) zijn nog enkele aanpassingen doorgevoerd. De items van het instrument worden gescoord op vier antwoordmodaliteiten ('zwak'; 'meer zwak dan sterk'; 'meer sterk dan zwak'; 'sterk'). De cesuur voor 'voldoende/onvoldoende' ligt voor elk item in het midden.

De psychometrische kwaliteit van het instrument is onderzocht met factoranalyse, homogeniteitanalyse en enkele specifieke op de fit van het Raschmodel gerichte toetsen. Tevens is de interbeoordelaarsbetrouwbaarheid onderzocht. Daarnaast is correlatoneel onderzoek verricht naar de constructvaliditeit en predictieve validiteit van het instrument. Ten behoeve van de constructvaliditeit zijn de geobserveerde leraren kort geïnterviewd en voor het onderzoek naar de predictieve validiteit is tijdens observaties bij de leerlingen vastgesteld hoe groot hun betrokkenheid bij de les was en in hoeverre zij verantwoordelijkheid namen voor hun eigen leerproces. Wanneer het onderzoeksinstrument voldoet aan de eisen van het Raschmodel dan is ook duidelijk of er een hiërarchie in de items bestaat en of deze hiërarchie in overeenstemming is met de gepostuleerde volgorde.

De tweede onderzoeksvraag vergt internationaal vergelijkend onderzoek. Verschillen tussen landen worden onderzocht op significantie met de Scheffétoets en op effectgrootte met Cohen's  $\delta$ . Verschillen tussen het gemiddelde van de landen worden met ANOVA gecorrigeerd voor verschillen in nationale context. Voor de derde onderzoeksvraag wordt in een cross-sectioneel onderzoek de samenhang geëxploreerd tussen de niveaustuur in het pedagogisch didactisch handelen en de beroepservaring van leraren.

Voor het beantwoorden van de beide onderzoeksvragen zijn door speciaal getrainde inspecteurs observaties uitgevoerd bij grote representatieve steekproeven van scholen in Nederland ( $n = 332$ ), Nedersaksen ( $n = 286$ ), Slowakije ( $n = 324$ ) en Vlaanderen ( $n = 240$ ). Voor Nederland is uit de populatie een simple random steekproef getrokken. In Neder-

saksen, Slowakije en Vlaanderen wordt de aselektiviteit van de steekproeven gegarandeerd door een procedure die door de betreffende onderwijsinspecties getypeerd werd als 'just the next school to visit'. De steekproefgrootte in Nederland, Nedersaksen, Slowakije en Vlaanderen heeft een nauwkeurigheid van ongeveer  $e = 5\%$ . De steekproefnauwkeurigheid  $e$  wordt uitgedrukt in een proportie van de standaarddeviatie en geeft aan met welke nauwkeurigheid resultaten uit een steekproef gegeneraliseerd kunnen worden naar de populatie waarover men uitspraken wil doen (Cochran, 1977, p. 76.) In Schotland en Kroatië zijn observaties uitgevoerd bij kleine steekproeven van respectievelijk 94 en 43 scholen. Voor deze landen wordt het populatiegemiddelde met minder grote nauwkeurigheid ( $e > 10\%$ ) geschat.

Het basisonderwijs start en eindigt in verschillende landen op verschillende leeftijden. Om onnodige contextverschillen te vermijden is ervoor gekozen de observaties uit te voeren bij leraren in rekenlessen van klassen waar leerlingen bij de start van het schooljaar ongeveer 9 jaar oud zijn. Onvermijdbare contextverschillen worden met covariantieanalyse gecontroleerd. Tabel 1 biedt inzicht in deze contextverschillen.

Voor het beantwoorden van de derde onderzoeksvraag zijn door onderwijsinspecteurs observaties uitgevoerd bij een aselekt getrokken steekproef van 402 Duitse leraren (steekproefnauwkeurigheid  $e = 5\%$ ) die rekenen gaven aan leerlingen die aan het begin van het schooljaar ongeveer 9 jaar oud waren.

## 4 Resultaten

### 4.1 Psychometrische kwaliteit

Bestudering van bovengenoemde reviews naar de leerling-effecten van het pedagogisch didactisch handelen van leraren leverde een grote pool van activiteiten op. Eerst zijn dubbele formuleringen en activiteiten die niet in elke les vatbaar zijn voor observatie verwijderd. De laatstgenoemde activiteiten zijn als items verwerkt in een interviewleidraad, waarmee elke leraar na de observatie is geïnterviewd over:

Tabel 1

Contextverschillen in deelnemende landen: gemiddelde (standaarddeviatie)

	Totaal	Vlaanderen	Neder- saksen	Slowa- kije <sup>1)</sup>	Neder- land	Kroatië	Schot- land
% Leerlingen uit sociaal zwakke gezinnen	16 (17)	21 (19)	16 (19)	-	12 (12)	7 (8)	17 (17)
% Leerlingen voor wie de schooltaal geen thuistaal is	11 (19)	14 (21)	15 (22)	-	9 (18)	2 (3)	2 (3)
Aantal leerlingen in de klas	21 (5)	19 (4)	20 (4)	22	21 (6)	23 (6)	21 (6)
% Combinatieklassen	20 (40)	10 (30)	31 (46)	0	32 (47)	35 (48)	40 (49)
% Beginnende leraren	11 (32)	10 (29)	2 (15)	-	16 (38)	12 (32)	22 (42)
% Mannelijke leraren	18 (39)	20 (40)	20 (40)	19	21 (40)	9 (29)	1 (11)
Aantal leerlingen op school	222 (124)	205 (106)	227 (98)	192	225 (131)	533 (256)	202 (157)

<sup>1)</sup>Voor Slowakije is gebruik gemaakt van de nationale statistieken

- De tijd die wekelijks besteed wordt aan rekenen.
- De frequentie waarmee leerlingprestaties gepeild worden met landelijk genormeerde toetsen.
- Maatregelen voor leerlingen die dreigen achterop te komen.

De resterende 32 activiteiten zijn gegroepeerd in de volgende domeinen van het observatie-instrument:

- Creëren van een veilig en stimulerend onderwijsleerklimaat.
- Efficiënt klasmanagement.
- Duidelijke en gestructureerde instructie.
- Het geven van een intensieve les die leerlingen activeert.
- Afstemmen van instructie en verwerking op verschillen.
- Leerlingen leerstrategieën aanleren.

De interbeoordelaarsbetrouwbaarheid van het observatie-instrument is onderzocht tijdens tweedaagse trainingssessies voor de onderwijsinspecteurs die aan het onderzoek deelnamen. Bij de training zijn drie gedigitaliseerde lessen gebruikt, die voorafgaand aan de trainingen waren geobserveerd door ongeveer 20 willekeurige Nederlandse inspecteurs. Hieraan zijn normen ontleend over het vaardigheidsniveau van de drie leraren op de gedigitaliseerde lessen en over de consensus die bereikt kon worden op de cesuur 'voldoende/onvoldoende'. Het vaardigheidsniveau dat voor de items als norm werd gebruikt was de gemiddelde score op de vierpuntschaal van elk item. De observatoren werden geacht om bij de verschillende items niet meer dan

een punt boven of onder dit gemiddelde te zitten. Hiermee werden de extreem soepele en de extreem strenge observatoren opgespoord. Van de onderlinge consensus werd verlangd dat deze gemiddeld over alle items op de cesuur voldoende/onvoldoende tenminste 80% was. Na de observatie van de eerste gedigitaliseerde les kregen de observatoren bij afwijking van de normscore en afwijking van de meerderheidsscore feedback in een groepsdiscussie. De te soepel of te streng oordelende observatoren en de observatoren die bij de cesuur voldoende/onvoldoende in de minderheid waren, werd gevraagd hun oordeel te verklaren aan hun collega's. Bij de tweede gedigitaliseerde les werd dezelfde procedure gevolgd. Daarna volgde de derde gedigitaliseerde les die als toets is gebruikt. Bij deze derde les lag de consensus met het criterium bij de verschillende trainingssessies gemiddeld tussen 85% en 95%.

Met het oog op de inhoudvaliditeit is met een toetsende factoranalyse (LISREL) vastgesteld of de zesfactorstructuur in het observatie-instrument, terug gevonden kon worden (Tabel 2). Het LISREL-model is zo opgesteld dat de schattingen van de ladingen van de items alleen zijn vrijgelaten voor de factor waar de items in het theoretische model bij horen. De overige parameters zijn vooraf op 0 gesteld.

Het zesfactormodel blijkt goed te passen op de data ( $\chi^2_{449} = 419,26$   $p = 0,84$   $GFI = 0,88$   $AGFI = 0,86$   $RMSR = 0,05$ ). Aangezien het vooraf opgestelde LISREL-model goed fit, is er geen reden om andere modellen te toetsen.

Tabel 2

Factorloadingen ( $\lambda$ ) en samenhang tussenfactoren ( $\phi$ ) van het LISRELmodel met zes factoren

De leraar...	Strategieën	Afstemmen	Activeren	Instructie	Organisatie	Klimaat
Vraagt leerlingen na te denken over oplossingsstrategieën	,78					
Moedigt kritisch denken van leerlingen aan	,79					
Leert leerlingen hoe zij complexe problemen kunnen vereenvoudigen	,73					
Laat leerlingen hardop denken	,76					
Leert leerlingen oplossingen te checken	,70					
Stimuleert het gebruik van controle activiteiten	,68					
Bevordert het toepassen van het geleerde	,55					
Stelt vragen die leerlingen tot denken aanzetten	,78					
Vraagt leerlingen na te denken over strategieën bij de aanpak	,78					
Stemt verwerking van leerstof af op verschillen tussen leerlingen		,87				
Stemt instructie af op verschillen tussen leerlingen		,86				
Biedt zwakke leerlingen extra leer- of instructietijd		,79				
Stimuleert het zelfvertrouwen van zwakke leerlingen		,48				
Zorgt voor interactieve instructie- en werkvormen			,71			
Hanteert werkvormen die leerlingen activeren			,72			
Betrekt alle leerlingen bij de les			,68			
Bevordert dat leerlingen hun best doen			,67			
Geeft duidelijke uitleg van de leerstof				,71		
Geeft duidelijke uitleg van het gebruik van didactische hulpmiddelen en opdrachten				,71		
Gaat tijdens de instructie na of leerlingen de leerstof hebben begrepen				,71		
Geeft feedback aan de leerlingen				,43		
Gaat na of de lesdoelen werden bereikt				,76		
Geeft goed gestructureerd les				,56		
Verduidelijkt bij de aanvang van de les de lesdoelen				,59		
Gebruikt de leertijd efficiënt					,74	
Zorgt voor een doelmatig klasmanagement					,74	
Gaat tijdens de verwerking na of leerlingen de opdrachten op een juiste manier uitvoeren					,65	
Zorgt voor een ordelijk verloop van de les					,66	
Zorgt voor wederzijds respect						,78
Ondersteunt het zelfvertrouwen van leerlingen						,79
Zorgt voor een ontspannen sfeer						,79
Toont in gedrag en taalgebruik respect voor leerlingen						,81
<b>Samenhang tussen factoren (<math>\phi</math>)</b>						
Leerlingen leerstrategieën aanleren	1,00					
Afstemmen van instructie en verwerking op verschillen	,66	1,00				
Zorgen voor een intensieve les die leerlingen activeert	,88	,71	1,00			
Duidelijke en gestructureerde instructie	,79	,66	,88	1,00		
Efficiënte lesorganisatie	,70	,62	,85	,92	1,00	
Creëren van veilig en stimulerend onderwijsleerklimaat	,63	,55	,82	,76	,76	1,00

De factorladingen ( $\lambda$ ) variëren tussen 0,55 en 0,87. Ze zijn hoog genoeg om een zinnvolle plaats te hebben bij de betreffende factor. De zes factoren blijken onderling sterk samen te hangen: De met LISREL berekende  $\phi$ -coëfficiënten variëren tussen 0,55 en 0,92. Ze zijn zo hoog dat een éénfactoroplossing eveneens voor de hand ligt.

De paralleliteit van de itemkarakteristieke curves is getoetst met een  $\chi^2$ -toets van Martin Löf en met een door Andersen (1973) ontworpen  $\chi^2$ -toets voor het vergelijken van de itemparameters bij personen met een hoge en lage vaardigheid. Zowel het resultaat van de Martin Löftoets,  $\chi^2_{930} = 1304,17$ , als dat van de Andersentoets,  $\chi^2_{31} = 88$ , bleek te hoog ten opzichte van het aantal vrijheidsgraden. De volgende items veroorzaakten de modelschending: 'verduidelijkt bij de aanvang van de les de lesdoelen', 'gaat na of de lesdoelen werden bereikt', 'geeft goed gestructureerd les', 'gaat tijdens de instructie na of leerlingen de leerstof hebben begrepen', 'bevordert het toepassen van het geleerde', 'stimuleert het gebruik van controle activiteiten', 'leert leerlingen oplossingen te checken', en 'moedigt kritisch denken van leerlingen aan'. De eerstgenoemde uitvallende items komen uit het domein instructie, de overige uit het domein leerstrategieën. De Andersentoets voor hoog- en laagscorders is bij de internationale steekproef na verwijdering van deze items in orde ( $\chi^2_{23} = 23,56; p = 0,429$ ) en de Martin Löftoets is net op de grens ( $\chi^2_{506} = 602,76; p = 0,002$ ). Ook voor de afzonderlijke landen zijn resultaten voor beide toetsen in orde ( $p$ -waarden: Vlaanderen: 0,089; 0,580; Nedersaksen: 0,993; 0,886; Nederland: 0,030; 0,224; Slowakije: 0,008; 0,668; Kroatië: 0,006; 0,668 en Schotland: 0,561; 0,667). Vervolgens is voor de 24 items de  $P$ -waarde, de moeilijkheidsgraad (itemparameter  $\sigma_i$ ) en de standaard error berekend. Tabel 3 biedt een overzicht.

In Tabel 3 constateren we dat de veronderstelde hiërarchie in activiteiten die loopt van het creëren van een veilig en stimulerend onderwijsleerklimaat naar het efficiënt organiseren van de les, gevolgd door duidelijke en gestructureerde instructie, het intensiveren van de les, en het goed afstemmen van het onderwijs op verschillen tussen

leerlingen en het leerlingen aanleren hoe ze iets moeten leren, inderdaad bestaat. Daarmee is het antwoord op de eerste onderzoeksvraag bevestigend.

Uit Tabel 3 blijkt verder dat de moeilijkheidsgraad van de itemparen: 2 en 3; 4 en 5; 8 en 9; 10 en 11; 12 en 13; 15 en 16; 20 en 21 en 23 en 24 zeer dicht bijeen ligt. Wanneer we uitsluitend in de meettechnische precisie van het observatie-instrument geïnteresseerd zouden zijn, zouden we van elk van deze acht itemparen een item kunnen missen. We zijn echter vooral geïnteresseerd in de structuur van het pedagogisch didactisch handelen van leraren. Daarom behouden we deze items in de schaal. Het resultaat is een toets die nu voor elk van de zes domeinen vier items telt. Dat is een qua inhoudsvaliditeit evenwichtiger verdeling van items dan in de itempool waarmee wij de analyses aanvingen.

Van een (Rasch)schaal mag verwacht worden dat de items samen op homogene wijze dezelfde latente vaardigheid meten. Eerste aanwijzingen voor de unidimensionaliteit van de schaal vonden we in de LISRELanalyse waaruit bleek dat de inter-correlaties tussen de zes domeinen bijzonder hoog zijn. Ook de KR-20 van 0,93 duidt op een sterk homogene schaal. Andere aanwijzingen kunnen we ontlenen aan onderzoek waarin we nagaan of de itemparameters ( $\sigma_i$ ) verschillen voor verschillende subgroepen van respondenten: mannelijke en vrouwelijke leraren, beginnende en ervaren leraren, en leraren uit verschillende landen. Hiervoor gebruiken we opnieuw de  $\chi^2$ -toets van Andersen. De resultaten staan in Tabel 4.

De unidimensionaliteit van de schaal wordt niet geschonden door sekse of beginnend leraarschap. Dat laatste is belangrijk omdat we het instrument willen inzetten in onderzoek naar de ontwikkeling van leraren. De nationaliteit van de leraren blijkt wel een schending van de unidimensionaliteit op te leveren. In Nedersaksen en Slowakije ligt de moeilijkheidsgraad van de items voor het afstemmen op verschillen iets hoger dan de moeilijkheidsgraad van de items voor het leerlingen aanleren van leerstrategieën. In Nederland, Vlaanderen, Schotland en Kroatië is dat precies andersom. We moeten ons hierbij wel realiseren dat de gemiddelde moei-



Tabel 3

Calibratie van 24 items in de internationale steekproef

nr	De leraar...	P	$\sigma_i$	s.e.
24	vraagt leerlingen na te denken over oplossingsstrategieën	,50	2,07	,09
23	moedigt kritisch denken van leerlingen aan	,50	2,06	,09
22	stemt verwerking van leerstof af op verschillen tussen leerlingen	,53	1,87	,09
21	stemt instructie af op verschillen tussen leerlingen	,56	1,65	,09
20	biedt zwakke leerlingen extra leer- of instructietijd	,56	1,65	,09
19	leert leerlingen hoe zij complexe problemen kunnen vereenvoudigen	,62	1,27	,09
18	zorgt voor interactieve instructie- en werkvormen	,64	1,13	,09
17	laat leerlingen hardop denken	,68	,82	,09
16	hanteert werkvormen die leerlingen activeren	,71	,60	,09
15	stimuleert het zelfvertrouwen van zwakke leerlingen	,71	,57	,09
14	gebruikt de leertijd efficiënt	,80	-,16	,09
13	geeft duidelijke uitleg van het gebruik van didactische hulpmiddelen en opdrachten	,82	-,31	,10
12	gaat tijdens de instructie na of leerlingen de leerstof hebben begrepen	,82	-,33	,11
11	geeft feedback aan de leerlingen	,84	-,49	,11
10	betreft alle leerlingen bij de les	,84	-,53	,11
9	bevordert dat leerlingen hun best doen	,86	-,72	,11
8	zorgt voor een doelmatig klasmanagement	,86	-,75	,12
7	geeft duidelijke uitleg van leerstof en opdrachten	,87	-,89	,12
6	gaat tijdens de verwerking na of leerlingen de opdrachten op een juiste manier uitvoeren	,88	-1,04	,12
5	zorgt voor wederzijds respect	,90	-1,32	,12
4	ondersteunt het zelfvertrouwen van leerlingen	,90	-1,33	,13
3	zorgt voor een ordelijk verloop van de les	,93	-1,82	,15
2	zorgt voor een ontspannen sfeer	,93	-1,87	,15
1	toont in gedrag en taalgebruik respect voor leerlingen	,94	-2,14	,17

lijkheidsgraad van de items over het leerlingen aanleren van leerstrategieën (gemiddelde  $\sigma = 1,56$ ) vrij dicht ligt bij de gemiddelde moeilijkheidsgraad van de items over het afstemmen van het onderwijs op verschillen tussen leerlingen (gemiddelde  $\sigma = 1,44$ ). Nader onderzoek moet uitwijzen of de nabijheid van de itemparameters, dan wel diversiteit in onderwijsculturen en opleiding hier de hoofdrol speelt.

De verzamelde gegevens over de tijdbesteding aan rekenen, de toetsfrequentie en de maatregelen voor zwakke leerlingen bieden de mogelijkheid de constructvaliditeit te onderzoeken. Tabel 5 geeft de resultaten.

Het blijkt dat de correlatie van het observatie-instrument met een interventie, die binnen de autonomie van de leraar valt (het treffen van maatregelen voor zwakke leerlingen) hoog en significant is, terwijl de correlaties met zaken die het schoolbeleid betreffen (rooster- en toetsplanning) zeer laag of nul zijn. Dat zijn aanwijzingen voor de constructvaliditeit. Leraren met een hoge pedagogisch didactische vaardigheid treffen vaker

en meer maatregelen voor leerlingen die achterop dreigen te komen en zaken die het schoolbeleid betreffen, hangen niet of nauwelijks met hun vaardigheid samen.

Uit Tabel 5 blijkt verder dat de correlaties van het observatie-instrument met de betrokkenheid van de leerlingen bij de les (0,62) en de mate waarin leerlingen verantwoordelijkheid nemen voor hun eigen leerproces (0,73) zeer hoog zijn. Een toename in complexiteit van pedagogisch didactisch handelen van de leraren gaat dus gepaard met een toename in betrokkenheid en verantwoordelijkheidszin van leerlingen bij hun eigen leerproces. Dat is een indicatie voor de predictieve validiteit van het observatie-instrument. Hoewel de items van het observatie-instrument ontleend zijn aan onderwijs-effectiviteitsonderzoek is verder onderzoek naar het verband met leerwinst nodig.

#### 4.2 Verschillen tussen landen

Tabel 6 bevat het vaardigheidsniveau ( $\xi_p$ ) dat bij de ruwe score (aantal 'goed') hoort. De internationale frequentieverdeling is links

Tabel 4

Unidimensionaliteit van de observatieschaal

	$\chi^2$	Degrees of freedom	p
Sekse	22,71	23	,478
Beginnende leraren / ervaren leraren	32,22	23	,096
Land	680,26	115	,000

Tabel 5

Constructvaliditeit en predictieve validiteit van de observatieschaal

	Cronbach's $\alpha$	M	sd	Correlatie met observatieschaal	
Aantal minuten rekenen per week		282	45	-,02	ns
Aantal landelijk genormeerde toetsen per jaar		1,3	1,5	,09	**
Aantal maatregelen voor zwakke leerlingen	,87	,64	,41	,42	**
Mate van betrokkenheid van leerlingen	,83	,59	,18	,62	**
Mate waarin leerlingen verantwoordelijkheid nemen voor hun eigen leerproces	,84	,72	,14	,73	**

scheef, wat erop duidt dat veel leraren van de onderzochte landen over hoge vaardigheden beschikken. Ruim een kwart van de leraren scoort zelfs hooguit één item negatief. Zij scoren ook op de moeilijkste onderdelen van het leraarsvak, zoals het afstemmen van het onderwijs op verschillen tussen leerlingen en het leerlingen kunnen uitleggen hoe ze iets moeten leren, positief. Ruim de helft van de leraren scoort rond het midden van de schaal. Ongeveer 18% van de leraren scoort een volle standaarddeviatie of meer onder het gemiddelde. Onder hen is een kleine groep leraren, het gaat om 3%, die er niet in slaagt om meer dan 5 items positief te scoren.

Tabel 7 biedt een overzicht van de gemiddelde ruwe scores, vaardigheidsscores ( $\xi$ ) en standaard deviaties van de deelnemende landen. Vanwege de bescheiden omvang van de steekproeven zijn Kroatië en Schotland tussen haken geplaatst.

De hoogste gemiddelde scores worden in Schotland en Nederland gevonden. De voor Nederland en Schotland positieve afwijkingen met Vlaanderen, Nedersaksen en Slowakije zijn significant. De laagste gemiddelde scores worden in Nedersaksen gevonden. De negatieve afwijking van Nedersaksen ten opzichte van Slowakije, Nederland en Schotland is significant. De gemiddelde scores in Slowakije en Vlaanderen nemen een middenpositie in.

In de laatste kolom wordt in *effect size* (i.e. percentage van de internationale standaarddeviatie) aangegeven hoe groot de ver-

schillen van de nationale gemiddeldes zijn ten opzichte van het internationaal gemiddelde. Het gemiddelde van Slowakije ligt op het internationaal gemiddelde. De gemiddelden van Vlaanderen en Kroatië liggen respectievelijk ongeveer een vijfde standaarddeviatie onder dan wel boven het internationaal gemiddelde. Dit zijn kleine verschillen. De gemiddelden van Nedersaksen en Nederland liggen respectievelijk 41% van een standaarddeviatie onder dan wel boven het internationaal gemiddelde. Dat is een middelgroot verschil. Het Schotse gemiddelde ligt 58% boven het internationaal gemiddelde. We memoreren dat de steekproeven uit Kroatië en Schotland relatief klein zijn en daardoor een weinig nauwkeurige schatting ( $e > 10\%$ ) bieden van de gemiddelde scores van de populaties in deze landen.

In Tabel 8 zijn de vaardigheidsscores ( $\xi$ ) van de verschillende landen gecorrigeerd voor respectievelijk klaskenmerken alléén, en leerling-, klas- en schoolkenmerken samen. Dat biedt inzicht in de doorwerking van internationale contextverschillen op de kwaliteit van het pedagogisch didactisch handelen van leraren. Voor Slowakije en Kroatië waren de gegevens over de contextvariabelen onvolledig. De afwijkingen van de voor klaskenmerken gecorrigeerde scores voor Vlaanderen, Nedersaksen, Nederland, Kroatië en Schotland zijn verwaarloosbaar. We kunnen aannemen dat verschillen in gemiddelde klasgrootte, percentage combinatieklassen, percentage beginnende leraren en sekse van

Tabel 6

Vaardigheid van leraren in de internationale steekproef

Ruwe score	Vaardigheid ( $\xi$ )	s.e.	% respondenten
24	>3,84 = 4	-	17,2
23	3,84	1,05	8,1
22	3,04	,78	8,2
21	2,53	,67	7,7
20	2,13	,61	7,3
19	1,79	,57	6,1
18	1,48	,54	5,3
17	1,20	,52	5,0
16	,94	,51	4,7
15	,69	,50	4,6
14	,44	,49	3,8
13	,21	,48	3,6
12	-,03	,48	3,9
11	-,26	,48	3,5
10	-,49	,48	2,3
9	-,72	,49	2,0
8	-,97	,50	1,6
7	-1,22	,51	1,2
6	-1,48	,53	1,1
5	-1,77	,55	0,4
4	-2,10	,59	0,9
3	-2,49	,66	0,7
2	-3,00	,77	0,2
1	-3,78	1,05	0,4
0	<-3,78 = -4	-	0,4

leraren geen rol spelen bij verschillen in de gemiddelde kwaliteit van de vaardigheden van de leraren in deze landen. De afwijkingen worden een heel kleine fractie groter wanneer we ook schoolkenmerken (ligging in stedelijk gebieden, schoolgrootte) en leerlingenkenmerken (percentage sociaal zwakke leerlingen en percentage leerlingen waarvan de thuistaal ongelijk is aan de schooltaal) in de correctie meenemen. Maar dan nog is de grootste afwijking (voor Schotland) verwaarloosbaar klein (< 10% van de standaarddeviatie).

#### 4.3 Cross-sectioneel onderzoek

Op grond van theorieën over lineair stijgende ontwikkeling in de beroepsloopbaan van le-

raren verwachten we dat leraren met meer ervaring een hoger niveau in pedagogisch didactische vaardigheid vertonen. Theorieën over niet-monotoon stijgende ontwikkeling doen echter vermoeden dat de pedagogisch didactische vaardigheid van leraren alleen aan het begin van hun beroepsloopbaan geleidelijk toeneemt om na 15 à 20 ervaringsjaren een daling door te maken. In cross-sectioneel onderzoek is bij een steekproef van 402 leraren in het Duitse basisonderwijs geëxploreerd hoe deze samenhang tussen beroepservaringen en niveau van pedagogisch didactische vaardigheden eruit ziet. Figuur 1 bevat de resultaten. Op de y-as staat de vaardigheid  $\xi$ , op de x-as het aantal beroepservaringen.

Tabel 7

Vaardigheid van leraren in verschillende landen

Gemiddelde en standaard deviatie					Vaardigheid $\xi$ verschilt significant op .01 niveau van:						Verskil in e.s. ten opzichte van internationaal gemiddelde
	Ruwe score		Vaardigheid $\xi$		V	N	S	NI	K	S	
	M	sd	M	sd							
Internationaal	18,08	5,51	1,86	1,77							
Vlaanderen	17,23	4,94	1,52	1,55				*		*	-0,19
Nedersaksen	15,87	5,40	1,13	1,63			*	*		*	-0,41
Slowakije	17,69	6,41	1,77	2,04		*		*		*	-0,05
Nederland	20,40	3,79	2,59	1,34	*	*	*				0,41
(Kroatië)	19,35	4,56	2,26	1,61							0,23
(Schotland)	21,00	4,21	2,89	1,46	*	*	*				0,58

Tabel 8

Vaardigheid van leraren in verschillende landen gecorrigeerd voor leerling-, klas- en schoolkenmerken

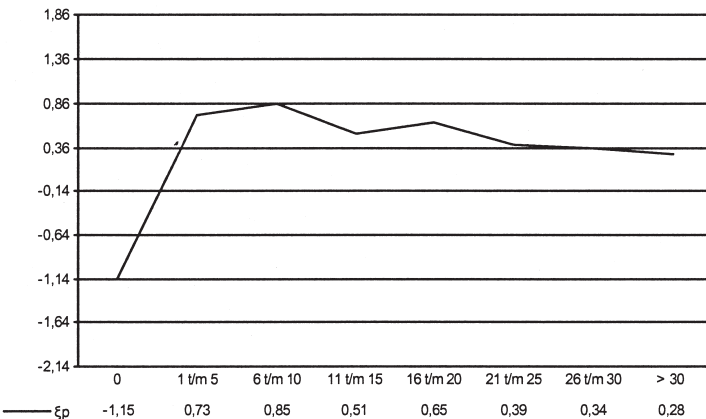
Vaardigheid $\xi$ gecorrigeerd voor:	Klaskenmerken	Leerling-, klas- en schoolkenmerken
	Klasgrootte, Combinatieklas, Beginnende leraar, Sekse leraar	% Sociaal zwakke leerlingen, % Leerlingen thuistaal ongelijk schooltaal, Klasgrootte, Combinatieklas, Beginnende leraar, Sekse leraar, Stedelijk gebied, Schoolgrootte
Vlaanderen	1,54	1,44
Nedersaksen	1,14	1,13
Slowakije	-	-
Nederland	2,51	2,55
(Kroatië)	2,22	-
(Schotland)	2,85	2,75

De net afgestudeerde Duitse leraar heeft nog een grote afstand tot de score van de gemiddelde leraar. Duitse leraren met 1-5 ervaringsjaren bereiken het niveau van de gemiddelde leraar. De Duitse leraar met 5-10 jaar ervaring komt een fractie boven het gemiddelde uit. Na het tiende ervaringsjaar gaat de grafiek afvlakken: De Duitse leraar met 11-20 jaar ervaring scoort rond het landelijk gemiddelde. De stijgende lijn die we bij de leraren met minder beroepservaringen zagen, is verdwenen. We zien hier hetzelfde niveau als bij de leraar met vijf beroepservaringen. De gemiddelde score van de Duitse leraren met meer dan 20 jaar ervaring ligt een kwart standaarddeviatie en het gemiddelde van leraren met meer dan 30 jaar ervaring ligt veertiende standaarddeviatie onder het landelijk gemiddelde. Het feit dat na het 20<sup>e</sup> arbeidservaringjaar de gemiddelde scores bij hogere arbeidservaringgroepen lager zijn dan bij lage ervaringgroepen betekent uiteraard niet, dat er geen hoogscorende oudere leraren zijn. Ongeveer 59% van de leraren

met 20 of meer arbeidservaringen scoort op of boven de hoogste score ( $\xi = .85$ ) in Figuur 1. Bij leraren met minder dan 20 jaar arbeidservaring ligt het percentage op of boven deze norm op ongeveer 72%. Dit is in overeenstemming met de hypothesen van Day en Gonçalves.

## 5 Discussie

De steekproef van het cross-sectionele onderzoek naar de ontwikkeling van beroepsvaardigheden bevat 402 leraren. Dat is veel voor een observatiestudie, maar weinig voor het doen van nauwkeurige uitspraken over acht beroepservaringgroepen, die qua omvang uiteenlopen van 17 tot 90 leraren. Voor nauwkeuriger analyses zijn grotere steekproeven nodig. Cross-sectionele resultaten kunnen niet longitudinaal geïnterpreteerd worden. Elke ervaringsjaren categorie bevat andere leraren, die op een ander moment hun opleiding hebben gehad en op een ander mo-



Figuur 1. Vaardigheid en aantal beroepservaringen

ment met hun beroepsloopbaan zijn gestart. Generatieverschillen kunnen belangrijke verschillen met zich meebrengen in beroepsvaardigheden, beroepsopvattingen, beroepsgewoonten en dergelijke. Ook kunnen op verschillende momenten in de beroepsloopbaan gebeurtenissen optreden die doorwerken op een bepaalde leeftijdsgroep zonder gevolgen te hebben voor een andere leeftijdsgroep. We moeten dus voorzichtig zijn met de interpretatie van deze gegevens. Dit neemt niet weg dat we in deze discussieparagraaf op grond van de gepresenteerde gegevens verder kunnen theoretiseren over de ontwikkeling in de pedagogisch didactische vaardigheden van leraren.

Duitse leraren die net de opleiding verlaten hebben, creëren voor hun leerlingen een veilig en stimulerend onderwijsleerklimaat. Ze organiseren hun klas efficiënt en sommigen kunnen duidelijk uitleggen en gestructureerd lesgeven. In termen van Fuller: Ze zijn hun 'task concerns' voorbij en klaar voor hun taak. Daarna moeten verdere persoonlijke rijping en vooral ook het systematisch leren van de ervaring ertoe leiden dat de pedagogisch didactische vaardigheden van jonge leraren zich verder ontwikkelen in de richting van steeds complexere vaardigheden die steeds sterker gericht raken op het beïnvloeden van de betrokkenheid en de prestaties van de leerlingen. Het gaat daarbij om het adequaat leren inspelen op verschillen tussen leerlingen en het leerlingen kunnen uitleggen hoe ze iets moeten leren. Dit zijn de moeilijkste dingen in het leraarsvak en ze kosten ook de meeste energie. Dit aanleren lukt niet in een enkele masterclass of eendaagse nascholingsconferentie. Uit diverse onderzoeken blijkt dat systematische klasobservatie gekoppeld aan gerichte coaching gedurende 1 à 2 jaar ertoe leidt dat leraren hun onderwijs beter leren afstemmen op verschillen tussen leerlingen. Uit deze onderzoeken blijkt ook dat de prestaties van de leerlingen onder invloed hiervan aanzienlijk vooruit gaan (Houtveen, Booi, De Jong & Van de Grift, 1999; Houtveen, Van de Grift & Creemers, 2004; Houtveen & Van de Grift, 2007a; 2011). Persoonlijke rijping en systematisch leren van de ervaring kunnen worden bevorderd door toenemende motivatie, betrokken-

heid, zelfvertrouwen en het gevoel van eigen doelmatigheid.

Leraren verschillen, maar bij velen blijkt hun vaardigheidsontwikkeling tussen het 10<sup>e</sup> en het 20<sup>e</sup> beroepservaringjaar op het hoogtepunt te komen. Eenmaal op dit punt aangekomen zullen zij door hun sterk toegenomen vaardigheden en het daarbij behorende toegenomen gevoel van eigen doelmatigheid, problemen op school en in het onderwijs niet meer ervaren als eigen vaardigheidsdeficiënties. Ze gaan zich bij problemen ergeren aan werkdruk, gebrek aan autonomie, tekort aan waardering van schoolleiding of collega's en aan veranderingen in het onderwijsbeleid die zich buiten hen om voltrekken (vergelijk ook Day, 2008.) Leraren komen gedurende deze jaren ook net als ieder ander in de 'midlife' overgang terecht. Motivatie- en betrokkenheidsproblemen zouden wel eens een sleutelrol kunnen vervullen (vergelijk ook Rosenholz en Simpson, 1990.)

Beginnende leraren maken, zolang ze problemen hebben met het onderwijsleerklimaat, ordehandhaving, of helder en gestructureerd uitleggen, zeer drukke en vermoeiende dagen door. Leraren die deze basale vaardigheden volledig beheersen 'komen veel minder moe thuis'. De meer complexe vaardigheden zoals het activeren van de leerlingen, het leerlingen uitleggen hoe ze iets moeten leren en vooral ook het adequaat inspelen op verschillen tussen leerlingen vragen echter weer veel voorbereiding en organisatie. Om de energie hiervoor te kunnen opbrengen is het belangrijk dat geen onnodige motivatie- en betrokkenheidsproblemen ontstaan.

Ook mechanismen op het vlak van zelfvertrouwen en eigen doelmatigheidsbeleving kunnen een dip veroorzaken (vergelijk Day, 2008 en Gonçalves, 2009.) Een aantal ervaren leraren komt in professioneel opzicht voor nieuwe taken te staan. Sommigen krijgen meer lesoverstijgende taken zoals mentoraat, team- en/of sectieleiderschap. De nieuwe uitdagingen kunnen het zelfbeeld over het eigen functioneren tijdelijk uit evenwicht brengen. De veilige zekerheid over de verworven vaardigheid in het lesgeven wordt ingeruild voor onzekerheid over vaardigheden zoals het kunnen beïnvloeden van col-

lega's bij het accepteren van een team-, of sectieleiderschap.

De studie onder Duitse leraren leidt tot nieuwe onderzoeksvragen. Een eerste vraag is: Ligt het startniveau van leraren die net de Nederlandse lerarenopleiding verlaten hebben net zo laag als bij Duitse leraren? Een tweede vraag gaat over de problematiek in de eerste jaren van beroepsuitoefening. Er wordt al jaren onderzoek gedaan naar de problemen van beginnende leraren (Corcoran, 1981; Créton & Wubbels, 1984; De Jonge & De Muijnck, 2002; Houtveen, Versloot & Groenen, 2006; Tabachnick, 1980; Van Kregten & Moerkamp, 2004; Veenman, 1984; Vonk, 1982) maar als het gaat om de feitelijke vaardigheden van beginnende leraren, dan liggen nog vele vragen open. Duitse leraren met 5-10 ervaringsjaren lieten een opmerkelijk hoge beroepsvaardigheid ten opzichte van de startvaardigheid zien. De vraag is: Maken jonge leraren in het begin van hun beroepsloopbaan een spectaculaire ontwikkeling door, of wordt dit opmerkelijke effect veroorzaakt doordat zoveel beginnende leraren in de eerste vijf jaar het beroep verlaten, waardoor alleen de besten overblijven? In Nederland verlaat ongeveer een kwart van de beginnende leraren gedurende de eerste vijf jaar het beroep (De Jonge & De Muijnck, 2002; Ministerie OCW, 2003; Van Kregten & Moerkamp, 2004).

Een derde vraag betreft leraren met 11-20 jaar beroepservaring. Het niveau van Duitse leraren met 11-20 jaar ervaring reikt niet verder dan een veilig en stimulerend onderwijsleerklimaat, een ordelijk georganiseerde les en duidelijke en gestructureerde instructie. Waarom zet de ontwikkeling van Duitse leraren zich niet door in de richting van intensieve feedback en interactie met hun leerlingen en een goede afstemming op verschillen tussen leerlingen? Komt het niveau van de Nederlandse leraar met 11-20 jaar ervaring uit boven het niveau van de Duitse leraren dat niet verder reikt dan de meest basale vaardigheden? Stagneert ook in het Nederlandse onderwijs de ontwikkeling bij veel leraren met meer dan 20 jaar ervaring?

Een vierde vraag betreft de oudere leraar. We hebben gezien dat de vaardigheidsscores van de Duitse leraren met meer dan 20 jaar

ervaring in het basisonderwijs een kwart standaarddeviatie onder het nationaal gemiddelde liggen. Bij leraren die meer dan 30 jaar voor de klas staan, ligt de gemiddelde vaardigheidsscore viertiende standaarddeviatie onder het landelijk gemiddelde. Doet dit zich ook voor in het Nederlandse onderwijs?

## Literatuur

- Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, 38, 1.
- Berliner, D. C. (1994). Expertise: The wonder of exemplary performances. In J. Mangieri, & C.C. Block (red.), *Creating powerful thinking in teachers and students: Diverse perspectives* (pp. 161-186). Fort Worth, Texas: Harcourt Brace College Publishers.
- Berliner, D. (2001). Learning about learning from expert teachers. *International Journal of Educational Research*, 35, 463-483.
- Bosker, R. J., & Witziers, B. (1996, april). *The magnitude of school effects, or: does it really matter which school a student attends?* Paper gepresenteerd op de jaarlijkse bijeenkomst van de American Educational Research Association, New York.
- Brandsma, H. P., & Knuver, J. W. M. (1989). Effects of school and classroom characteristics on pupil progress in language and arithmetic's. *International Journal of Educational Research*, 13, 777-788.
- Burke, P. J., Christensen, J. C., Fessler, R., & Price, J. R. (1987, april). The teacher career cycle: Model development and research report. Paper gepresenteerd op de jaarlijkse bijeenkomst van de American Educational Research Association, Washington D.C..
- Canrinus, E.T., Helms-Lorenz, M., Beijgaard, D., Buitink, J. & Hofman, W.H.A. (2011). Profiling teachers' sense of professional identity. *Educational Studies*, 37 (5), 593-608.
- Cochran, W. G. (1977). *Sampling techniques*. New York: John Wiley and Sons.
- Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weifeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.
- Corcoran, E. (1981). Transition shock: teacher's paradox. *Journal of Teacher Education*, 32 (3), 19-23.

- Cotton, K. (1995). *Effective schooling practices: a research synthesis*. Portland: Northwest Regional Educational Laboratory.
- Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
- Creton, H. A., & Wubbels, T. (1984). *Ordeproblemen bij beginnende leraren*. Utrecht: Uitgeverij W.C.C.
- Day, C. (2008). Committed for life? Variations in teachers' work, lives and effectiveness. *Journal of Educational Change*, 9, 243-260.
- Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. Oxford: Basil Blackwell.
- 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, National Center to Improve the Tools of Educators.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London: Routledge Falmer.
- Feiman-Nemser, S. (1983). Learning to teach. In L.S. Shulman, & G. Sykes (red.), *Handbook of teaching and policy* (pp. 150-171). New York, VS: Longman.
- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers' College Record*, 103, 1013-1055.
- Fokkens-Bruinsma, M., & Cannrinus, E. T. (2011). Wat motiveerde je om docent te worden? Relaties met professionele betrokkenheid en vertrouwen in eigen kunnen bij beginnende en ervaren docenten in het voortgezet onderwijs. *Pedagogische Studiën*, 88, 41-56.
- Fuller, F. F. (1969). Concerns of teachers: a developmental conceptualization. *American Educational Research Journal*, 6, 207-226.
- Fuller, F. (1970). *Personalized education for teachers: One application of the teachers concerns model*. Austin: R&D Center for Teacher Education.
- Gonçalves, J. A. (2009). Professional development and teaching career. Career phases, curriculum and supervision. *Sísifo. Educational Sciences Journal*, 8, 21-32.
- Grift, W. van de. (1985). Onderwijsleerklimaat en leerlingprestaties. *Pedagogische Studiën*, 62, 401-414.
- Grift, W. van de. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument *Educational Research*, 49, 127-152.
- Grift, W. van de, & Lam, J. F. (1998). Het didactisch handelen in het basisonderwijs. *Tijdschrift voor Onderwijsresearch*, 23, 224-241.
- Hammerness, K., Darling-Hammond, L., & Bransford, J. (2005). How teachers learn and develop. In L. Darling-Hammond, & J. Bransford (red.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 358-389). San Francisco Californië: Jossey-Bass.
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy* 100, 84-117.
- Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review* 30, 466-479.
- Hanushek, E.A. & Zhang, L. (2006). *Quality-Consistent Estimates of International Returns to Skill*, NBER Working Papers12664, National Bureau of Economic Research, Inc.
- Hargreaves, A. (2005). Educational changes takes ages: Life, career and generational factors in teachers' emotional responses to educational change. *Teaching and Teacher Education*, 21, 967-983.
- Hattie, J. (2009). *Visible learning A synthesis of over 800 meta-analyses relating tot achievement*. London: Routledge.
- Holloway, J. H. (2003). Sustaining experienced teachers. *Educational Leadership*, 60 (8), 87-89.
- Houtveen, A. A. M., Booij, N., Jong R. de, & Grift, W. J. C. M. van de. (1999). Adaptive instruction and pupil achievement. *School Effectiveness and School Improvement*, 10, 172-192.
- Houtveen, A. A. M., Mijs, D., Vernooy, K., Grift, W. J. C. M., van de, & Koekebacker, E. (2003). *Risicoleerlingen bij technisch lezen*. Utrecht: ISOR.
- Houtveen, A. A. M., Grift, W. J. C. M, van de, & Creemers, B. P. M. (2004). Effective school improvement in mathematics. *School Effectiveness and School Improvement*, 15, 337-376.
- Houtveen, A. A. M., & Grift, W. J. C. M. van de. (2007a). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, 18, 173-190.
- Houtveen, A. A. M., & Grift, W. J. C. M. van de.

- (2007b). Reading instruction for struggling learners. *Journal of Education for Students Placed at Risk*, 12, 405-424.
- Houtveen, T., & Grift, W. van de. (2011). Improving reading achievements of struggling learners. *School Effectiveness and School Improvement*. Online beschikbaar op <http://www.tandfonline.com/doi/pdf/10.1080/09243453.2011.600534>.
- Houtveen, A. A. M., Versloot, B., & Groenen (2006). *De begeleiding van startende leraren; in het voortgezet onderwijs en in het basis-onderwijs*. Den Haag: Sectorbestuur Onderwijsarbeidsmarkt (SBO).
- Huberman, M. (1989). The professional life cycle of teachers. *Teachers College Record* 91, 31-57.
- Jonge, J. F. M. de, & Muijnck, J. A. de. (2002). *Waarom leraren de sector verlaten. Onderzoek naar de uitstroom in het primair en secundair onderwijs*. Zoetermeer: EIM.
- Joyce, B., & Showers, B. (2002). *Student Achievement Through Staff Development* (3e druk). Alexandria: ASCD.
- Kompier, M. A. J., & Cooper, C. L. (1999). *Preventing stress, improving productivity: European case studies in the workplace*. London: Routledge.
- Kregten, A. van, & Moerkamp, T. (2004). *Startende leraren in het buitenland*. Den Haag: SBO.
- 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, 12-23.
- Levine, D. U., & Lezotte, L. W. (1990). Unusually effective schools: a review and analysis of research and practice. Madison: The 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* (pp. 525-547). New York, VS: Macmillan.
- Lynn, S. K. (2002). The winding path: Understanding the career cycle of teachers. *The Clearing House*, 75, 179-182.
- Marzano, R. (2008). *Wat werkt in de klas*. Middelburg: Bazalt.
- Ministerie van OCW. (2003). *Werken in het onderwijs (WIO-2003)*. 's-Gravenhage: OCW.
- Purkey, S. L., & Smith, M. S. (1983). Effective schools: a review. *Elementary School Journal*, 83, 427-452.
- Richardson, V., & Placier, P. (2001). Teacher change. In V. Richardson (red.), *Handbook of research on teaching* (pp. 905-947). Washington, DC: American Educational Research Association.
- Rosenholtz, S. J., & Simpson, C. (1990). Workplace conditions and the rise and fall of teacher commitment. *Sociology of Education*, 63, 241-257.
- Sammons, P. Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools: a review of school effectiveness research*. London: Office for Standards in Education.
- Schaffer, E. C., & Nesselrodt, P. S. (1992, april). *The development and testing of the special strategies observation system*. Paper gepresenteerd op de jaarlijkse bijeenkomst van de American Educational Research Association, San Francisco.
- Schaffer, E., Stringfield, S., & Wolfe, D. (1992). Two-year effects of a sustained beginning teacher induction program on classroom interactions. *Journal of Teacher Education*, 43 (3), 203-214.
- Scheerens, J. (1992). *Effective schooling: research, theory and practice*. London: Cassell.
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schön, D. A. (1987). *Educating the reflective practitioner*. San Fransico: Jossey-Bass.
- Sprinthall, N. A., Reiman, A. J., & Thies-Sprinthall, L. (1996). Teacher professional development. In J. Sikula (red.), *Handbook of research on teacher education* ( pp. 666-703). New York: Macmillan.
- Steffy, B. E., & Wolfe, M. P. (2001). A life cycle model for career teachers. *Kappa Delta Pi Record*, 38, 16-19.
- Tabachnick, B. R. (1980). Intern-teacher roles: Illusions, disillusion, and reality. *Journal of Education*, 162, 122-137.
- Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*. 54, 143-178.
- Vonk, J. H. C. (1982). *Opleiding en praktijk*. Am-



sterdam: VU-boekhandel/uitgeverij.

Walberg, H. J., & Haertel, G. D. (1992). Educational psychology's first century. *Journal of Educational Psychology*, 84, 6-19.

Wijnstra, J., Ouwens, M., & Béguin, A. (2003). *De toegevoegde waarde van de basisschool*. Arnhem: CITOgroep.

Manuscript aanvaard op: 11 oktober 2011

## Auteurs

**Wim van de Grift** is Hoogleraar Onderwijskunde en directeur Lerarenopleiding, **Marieke van der Wal** is als universitair docent en **Marjolein Torenbeek** als onderzoeker verbonden aan de Lerarenopleiding van de Rijksuniversiteit Groningen.

*Correspondentieadres:* W. van de Grift, Postbus 800, 9700 AV Groningen, e-mail:w.j.c.m.van.de.grift@rug.nl

## Abstract

### Development in teaching skills

Teaching skills are observed in samples of primary schools in the Netherlands, Belgium, Germany, Slovakia, Croatia and Scotland. A sequential level of difficulty in teaching skills has been observed. Activities in the domains 'ensuring a safe and stimulating environment' and 'efficient lesson organization' constitute the most easy competence domains, while those in the domains of 'clear and structured instruction' and 'intensifying the lesson and activating students' are of intermediate difficulty. Activities in the domains of 'adapting instruction to student differences' and 'teaching students thinking and learning strategies' are the most difficult ones. A cross-sectional study reveals that teachers, during the first 20 years of experience seem to develop their teaching skills according this sequential level of difficulty. After about 20 years of experience the level of teaching of the average teacher seems to decrease slowly.