

OPBRENGSTGERICHT WERKEN IN HET PRIMAIR ONDERWIJS:

COMPETENTIES, UITVOERING EN RESULTATEN

RAPPORTAGE

25-04-2015

OPBRENGSTGERICHT WERKEN IN HET PRIMAIR ONDERWIJS: COMPETENTIES, UITVOERING EN RESULTATEN

*J.M. Faber
A.J. Visscher
W.G.C. Schut*

UNIVERSITEIT TWENTE.

Faculty of Behavioral, Management & Social sciences (BMS)

Vakgroep Onderzoeksmethodologie, Meetmethoden en Data-analyse (OMD)

25-04-2015

INHOUDSOPGAVE

Samenvatting.....	4
1. Inleiding.....	7
2. Resultaten OGW-projecten primair onderwijs	10
3. Onderzoeksmethode	12
3.1. Werving en selectie.....	12
3.2. Dataverzameling en onderzoeksinstrumenten.....	13
3.3. Betrouwbaarheid observanten en leerkracht stabiliteit	16
3.4. Analyses	17
4. Resultaten	18
4.1. Beschrijvende Resultaten.....	18
4.2. Effecten op leerresultaten	19
5. Conclusies en discussie	23
5.1. Discussie.....	25
Literatuurlijst	27
Bijlage A: Focus-project.....	29
Bijlage B: Scoremodel protocol en groepsplan	30
Bijlage C: Format groepsplan Focus	34

SAMENVATTING

Het Ministerie van Onderwijs, Cultuur en Wetenschap ziet opbrengstgericht werken (OGW) als een belangrijk middel om taal- en rekenprestaties te verhogen in het primair onderwijs (Ministerie van Onderwijs, Cultuur en Wetenschap, 2007; 2011). Met OGW wordt over het algemeen een cyclische werkwijze met als doelstelling het realiseren van hogere opbrengsten bedoeld (Doolaard, 2013). Een OGW-cyclus bestaat doorgaans uit de vier volgende componenten; 1) het analyseren en evalueren van resultaten, 2) het SMART en uitdagende doelen stellen, 3) het bepalen van een strategie om gestelde doelen te bereiken, en 4) het uitvoeren van deze strategie (Keuning & van Geel, 2012). Uit het Onderwijsverslag blijkt dat basisscholen wel geïnvesteerd hebben in het beter analyseren van gegevens uit het leerlingvolgsysteem, maar vaak blijft het daarbij en worden geen concrete conclusies getrokken die leerkrachten vertalen naar een instructiestrategie (Inspectie van het Onderwijs, 2014). In het huidige onderzoek is de assumptie, dat een goede beheersing en uitvoering van de vier OGW-componenten resulteert in hogere leerresultaten in het primair onderwijs, nader onderzocht.

Voor het beantwoorden van de onderzoeksvragen is de OGW-attitude van leerkrachten in het primair onderwijs in beeld gebracht evenals de kennis die leerkrachten bezitten over de analysemogelijkheden binnen een leerlingvolgsysteem (LVS). De kwaliteit van 128 groepsplannen is geanalyseerd, en de mate van de afstemming van instructie en verwerking is gemeten met behulp van 212 lesopnames. Er participeerden 26 basisscholen met 77 leerkrachten in het onderzoek, alle deelnemende scholen hadden de Focus-interventie gevolgd, welke gericht is op OGW. Voor een goede spreiding in leeftijdsgroepen namen leerkrachten uit groep twee, vier en zeven deel aan het onderzoek.

De volgende onderzoeksvragen stonden centraal:

1. In hoeverre bezit men in de onderwijspraktijk de voor elk van de componenten van OGW vereiste attitudes, kennis en vaardigheden?

Leerkrachten in de steekproef bleken overwegend een positieve OGW-attitude te hebben, leerkrachten gaven gemiddeld een 3,5 score voor OGW op een schaal van 1 (laag) tot en met 4 (hoog). Leerkrachten beheersten daarnaast de benodigde kennis over de beschikbare analysemogelijkheden binnen een LVS, gemiddeld scoorden leerkrachten 78% goed op kennisvragen die hierop betrekking hadden. De kwaliteit van de geplande instructie in het groepsplan is echter vaak onvoldoende gedifferentieerd, en de uitkomsten van leerresultatenanalyses met een LVS zouden meer benut kunnen worden. Leerkrachten kregen gemiddeld lage scores op hun groepsplan, op de kwaliteit van de instructiebeslissingen, de evaluatie van de voortgang en de subgroep specifieke items werd niet meer dan 53% behaald van het totaal aantal punten. Met uitzondering van de kwaliteit van de doelen, hierop scoorden leerkrachten uit groep vier en zeven voldoende. Op het lesobservatie-onderdeel afstemmen van instructie en verwerking, scoren de leerkrachten uit groep vier en zeven vergelijkbaar. De gegeven scores van de leerkrachten uit groep vier en zeven liggen overwegend tussen de 'meer zwak dan sterk' en 'meer sterk dan zwak', terwijl de scores voor de leerkrachten uit groep twee iets lager liggen en voornamelijk als 'meer zwak dan sterk' zijn

beoordeeld. Leerkrachten kregen gemiddeld een voldoende beoordeling voor het creëren van een veilig en stimulerend leerklimaat, voor de efficiënte lesorganisatie, en voor het aanbieden van een duidelijke en gestructureerde instructie en een intensieve en activerende les. Na het afstemmen van de instructie en verwerking hadden leerkrachten het meest moeite met het aanleren van leerstrategieën.

2. In welke mate wordt elk van de OGW-componenten in de onderwijspraktijk uitgevoerd?

De uitvoering van OGW in de klas is vaak nog niet optimaal. Differentiëren is een complexe professionele vaardigheid (Inspectie van het Onderwijs, 2014) en bovendien bestaat er nog onduidelijkheid over hoe 'goed differentiëren' er exact uit ziet, en wanneer dit tot hogere leerresultaten leidt (evenals hoe leerkrachten daar effectief in getraind kunnen worden). De bevindingen uit dit onderzoek laten zien dat het maken van instructiegroepen alleen onvoldoende is, van belang lijkt een goede invulling van het afstemmen van de instructie en de verwerking, en het voorkomen van een te rigide aanpak in de lesorganisatie.

3. Hebben leerkrachten die meer over de OGW-competenties beschikken, én OGW meer in de praktijk brengen leerlingen die een sterkere groei in hun leerresultaten doormaken?

In het onderzoek is geen significant positief verband gevonden tussen enerzijds de lesobservatiescores en groepsplanscores, en anderzijds de leerresultaten. Een mogelijke verklaring voor deze bevinding zou zijn dat met het gebruikte instrument hoofdzakelijk de frequentie waarmee een aantal differentiatieactiviteiten worden uitgevoerd is gemeten, en onvoldoende de *kwaliteit* daarvan. Een hoge frequentie van differentiatieactiviteiten hoeft niet altijd samen te gaan met een goede differentiatiekwaliteit. Dit sluit aan bij de bevindingen van het beoordelen van de groepsplannen waar ook uit bleek dat weinig gedifferentieerd wordt in de instructiebeslissingen. Leerlingen in de instructiegevoelige groep lijken het minst profijt te hebben van een leerkracht met een hoge score op het afstemmen van de instructie en verwerking in vergelijking met leerlingen die een verlengde instructie volgen en leerlingen die een verkorte instructie volgen. Deze interactie is echter niet significant.

Concluderend kan gesteld worden dat er in het primair onderwijs meer aandacht nodig is voor het vertalen van leerlinggegevens naar goede en werkbare gedifferentieerde instructie-aanpakken. Dit zou bijvoorbeeld gedaan kunnen worden door de ontwikkeling van trainingen die specifiek gericht zijn op het opstellen van goed uitvoerbare groepsplannen, en de uitvoering van die plannen in de klas. Het gebruik van flexibele instructiegroepen is daarbij waarschijnlijk van belang, hierop wordt ook in differentiatie literatuur gewezen (Tomlinson et al., 2004). Bij de succesvolle uitvoering van differentiëren in de klas spelen complexe vaardigheden een rol (Inspectie van het Onderwijs, 2014). Uit de resultaten van het onderzoek blijkt dat leerkrachten het aanbieden van verschillende leerstrategieën en het afstemmen van de instructie en verwerking vaak nog onvoldoende beheersen. Het kunnen aanbieden van verschillende leerstrategieën om er zo voor te zorgen dat elke leerling actief kan participeren in de groep wordt gezien als een belangrijk onderdeel van een gedifferentieerde instructie (Tomlinson, et al., 2004). Hoe aan deze complexe vaardigheden in de pre-service opleiding van leerkrachten, en tijdens de in-service professionalisering aandacht besteed

kan worden is belangrijk om te onderzoeken, evenals hoe de kwaliteit van de differentiatie goed gemeten kan worden.

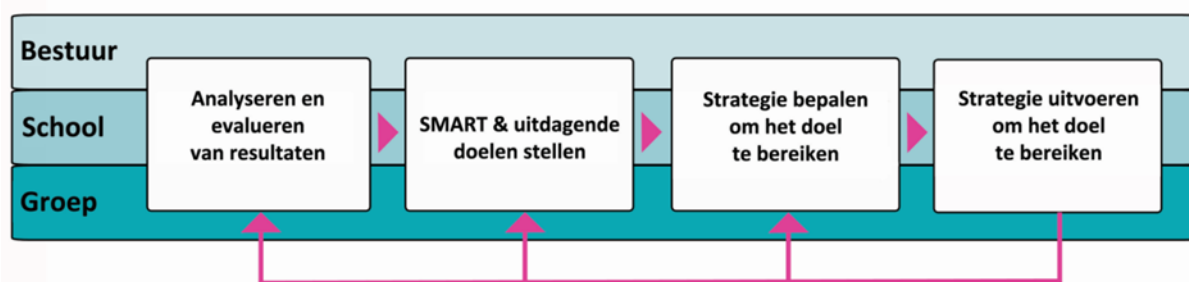
1. INLEIDING

Uit de nota 'Scholen voor Morgen' en het actieplan 'Basis voor presteren' volgt dat het Ministerie van Onderwijs, Cultuur en Wetenschap opbrengstgericht werken (OGW) ziet als een belangrijk middel om de taal- en rekenresultaten van leerlingen te verhogen (Ministerie van Onderwijs, Cultuur en Wetenschap, 2007; 2011). In 2015 zou 60% van de basisscholen moeten voldoen aan de indicatoren voor OGW, en in 2018 geldt dit doel voor 90% van de basisscholen. Uit het Onderwijsverslag over het schooljaar 2011-2012 blijkt echter dat nog slechts 28% van de scholen een voldoende beoordeling krijg op alle vijf inspectie-indicatoren voor OGW (Inspectie van het Onderwijs, 2013), en uit het daarop volgende Onderwijsverslag blijkt dat er in het schooljaar 2012-2013 weinig verbetering zichtbaar is (Inspectie van het Onderwijs, 2014).

Hoewel een algemeen aanvaarde definitie voor OGW (Blok, Ledoux & Roeleveld, 2013) niet beschikbaar is staat over het algemeen een cyclische werkwijze centraal met als doelstelling het realiseren van hogere opbrengsten (Doolaard, 2013). Centrale elementen binnen deze werkwijze zijn het hebben van hoge verwachtingen van leerlingen, het analyseren en in kaart brengen van de ontwikkeling en onderwijsbehoeften van leerlingen op basis van het afnemen van toetsen, het stellen van doelen, en op basis van dit alles, het zo goed mogelijk afstemmen van het onderwijs op wat leerlingen nodig hebben. OGW is veelal verbonden aan de kernvakken taal en rekenen, deze expliciete verbinding volgt echter niet altijd uit andere definities, zoals die van de Onderwijsinspectie en de Onderwijsraad (Blok, Ledoux & Roeleveld, 2013).

De genoemde centrale elementen van OGW komen terug in de OGW-cyclus (Figuur 1). Van een aantal van de componenten van OGW is aangetoond dat ze samengaan met hogere leerresultaten. Informatie uit toetsen kan beschouwd worden als prestatiefeedback, besturen, scholen en leerkrachten ontvangen immers informatie over hun resultaten in vergelijking met andere besturen en scholen en/of eigen voorgaande jaren. Onderzoek toont aan dat feedback een prestatieverhogend effect kan hebben (Hattie & Timperley, 2007). Hetzelfde geldt voor het werken met uitdagende SMART doelen (Locke & Latham, 2002). Dergelijke doelen zijn voor iedereen duidelijk en motiverend en geven bovendien richting aan de activiteiten die nodig zijn voor het realiseren van de doelen.

Op basis van de twee voorgaande stappen (feedback op resultaten, en het stellen van doelen) kan in de derde stap op een weloverwogen wijze een strategie bepaald worden om de huidige situatie in de gewenste situatie te veranderen (Visscher & Ehren, 2011). In stap drie in Figuur 1 gaat het daarbij nog om plannen en voornemens, in de vierde stap is het de bedoeling dat de voornemens in de praktijk worden gebracht. Op het niveau van de groep, zal dit resulteren in meer differentiatie tussen leerlingen, omdat een analyse van leerresultaten zal leiden tot het constateren van verschillende leerbehoeften bij leerlingen. Op die verschillen zal het onderwijs dan wel moeten worden afgestemd.



Figuur 1: Opbrengstgericht werken op verschillende niveaus (Keuning & van Geel, 2012)

Of OGW resulteert in hogere leerresultaten is, en wordt in verschillende projecten onderzocht (Inspectie van het Onderwijs, 2011; 2012; Doolaard, 2013; Staman, Visscher & Luyten, 2014). Uit de resultaten van de eerste projecten volgt dat leerkrachten door middel van training de benodigde vaardigheden en kennis ten aanzien van OGW kunnen verbeteren (Doolaard, 2013; Staman, Visscher & Luyten, 2014). De vraag is of leerkrachten de daarop volgende stappen van OGW ook daadwerkelijk en goed uitvoeren en, wanneer dit het geval is, of hogere leerresultaten worden behaald. De assumptie dat een goede beheersing en uitvoering van de vier OGW-stappen resulteert in hogere leerresultaten wordt in dit onderzoek nader onderzocht. Het onderzoek richt zich daarbij op de uitvoering van OGW op het niveau van de groep, dus de uitvoering van OGW door leerkrachten in de klas.

De attitude, kennis en vaardigheden van leerkrachten zal de uitvoering van de componenten van OGW beïnvloeden (Tabel 1). Leerkrachten zullen waarschijnlijk succesvoller zijn wanneer zij gemotiveerder zijn voor OGW, bereid zijn om kritisch te kijken naar de eigen onderwijsresultaten, en wanneer ze een link willen leggen met behaalde leerresultaten en het door hen zelf aangeboden onderwijs. Kennis en vaardigheden zijn nodig voor het analyseren en evalueren van resultaten, en voor het stellen van doelen. Het is bijvoorbeeld van belang om bekend te zijn met de verschillende analysemogelijkheden in leerlingvolgsystemen, om de resultaten uit deze analyses juist te kunnen interpreteren, en om werkbare en passende doelen te kunnen formuleren. Voor het bepalen van de instructiestrategie dient de leerkracht een goede verbinding te kunnen leggen tussen de 'data' en wat een leerling qua instructie nodig heeft. Voor een goede uitvoering van de geplande instructiestrategie zijn daarnaast ook algemene didactische vaardigheden voorwaardelijk, zonder een goed klassenmanagement zal het bijvoorbeeld lastig zijn om gedifferentieerd onderwijs te geven (Visscher & Ehren, 2011).

Tabel 1: Componenten van en competenties voor OGW

Componenten	Competenties		
	a. Attitude	b. Kennis	c. Vaardigheden
<ul style="list-style-type: none"> • Analyseren en evalueren van resultaten • SMART & uitdagende doelen stellen • Instructie strategie bepalen • Instructie uitvoeren 			

In aansluiting op het voorgaande zijn de volgende onderzoeksvragen onderzocht:

1. In hoeverre bezit men in de onderwijspraktijk de voor elk van de componenten van OGW vereiste attitudes, kennis en vaardigheden?
2. In welke mate wordt elk van de OGW-componenten in de onderwijspraktijk uitgevoerd?
3. Hebben leerkrachten die meer over de OGW-competenties beschikken, én OGW meer in de praktijk brengen leerlingen die een sterkere groei in hun leerresultaten doormaken?

De onderzoeksvragen worden beantwoord met onderzoeksgegevens uit de eerste drie Focus-projecten die uitgevoerd zijn binnen de Universiteit Twente, daarnaast is aanvullende data verzameld. In de Focus-projecten zijn 150 basisscholen met ongeveer 2800 teamleden twee jaar lang geschoold in, en begeleid bij het in de praktijk brengen van OGW. Bijlage A bevat een korte omschrijving van de inhoud van de Focus-projecten.

In het volgende hoofdstuk worden de bevindingen van Nederlandse onderzoeken naar OGW samengevat, in hoofdstuk drie wordt de onderzoeksmethode beschreven waarna in hoofdstuk vier en vijf de onderzoeksresultaten en conclusies van dit onderzoek gepresenteerd worden.

2. RESULTATEN OGW-PROJECTEN PRIMAIR ONDERWIJS

In de Focus-projecten van de Universiteit Twente zijn basisscholen intensief getraind in OGW. Het project betreft een twee jaar durende training gericht op het gebruiken van digitale leerlingvolgsystemen, het formuleren van doelen, het afstemmen van de instructie op de leerbehoeften, en het doorlopen van een OGW-cyclus. In Focus I werden leerkrachten van de onderbouw en bovenbouw apart getraind, steeds in combinatie met de intern begeleiders en de schoolleiders van de school. In Focus II was geen sprake meer van een splitsing tussen de onder- en bovenbouw en kreeg het gehele team gezamenlijk twee jaar training. Deze scholen konden ervoor kiezen om de interventie in het tweede jaar te verbreden naar OGW bij spelling en lezen, of om deze te verdiepen naar OGW bij rekenen. Er deden 43 scholen mee aan Focus I en 56 scholen aan Focus II, de resultaten uit deze eerste twee Focus-projecten zijn inmiddels (deels) gepubliceerd.

Uit de resultaten van Focus I blijkt dat OGW kennis en vaardigheden significant verbeteren door de training; de sterkste verandering werd gevonden bij het gebruik van de leerlingvolgsystemen en het interpreteren van de analyses uit deze leerlingvolgsystemen. Er was weinig verandering in de attitude ten aanzien van OGW. Respondenten hadden bij aanvang van de projecten al een positieve attitude en deze bleef positief gedurende de twee interventiejaren (Staman, Visscher & Luyten, 2013). In Focus II was wel sprake van een daling voor een aantal attitude-componenten, maar ook daar bleven de attitudescores hoog. Focus II had een positief effect op leerresultaten. Effecten waren bovendien groter op de scholen met een relatief hoog percentage leerlingen met een lage sociaaleconomische status (SES). Opvallend was dat de interventie een negatieve impact had op de resultaten van leerlingen met een gemiddelde SES die op scholen zaten met een gemiddeld hoog SES. Een mogelijke verklaring daarvoor zou kunnen zijn dat leerkrachten door OGW hun aandacht meer spreiden over alle leerlingen en daardoor minder tijd dan voorheen investeren in de leerlingen die in vergelijking met de rest van de groep lager presteren (Van Geel, Keuning, Fox & Visscher, submitted).

Een tweede project waarin scholen getraind zijn in OGW is het Streef project van het Gronings instituut voor onderzoek van onderwijs (GION-Rijksuniversiteit Groningen). Het Streefproject bestaat uit verschillende deelprojecten waaronder het project 'Gebruik maken van opbrengsten' en het project 'Streef Middenbouw'.

De doelstelling van het deelproject 'Gebruik maken van opbrengsten' was om scholen meer opbrengstgericht te laten werken door ze optimaal gebruik te laten maken van een leerlingvolgsysteem. Er participeerden 40 basisscholen in het project. De resultaten laten zien dat deelnemende leerkrachten in vergelijking met een controlegroep een uitgebreidere analyse van leerresultaten maakten, reflecteerden daar ook meer op, en ze maakten meer gebruik van groepsplannen waarin de geplande differentiatie werd beschreven. De effecten van het eerste project op de leerresultaten bleken niet significant te zijn. Wel was sprake van een positieve tendens, maar die werd niet vastgehouden in het volgende schooljaar. Doolaard (2013) geeft hiervoor twee verklaringen. Ten eerste dat het analyseren van leerresultaten een individuele activiteit bleef en vaak niet plaatsvond in teamverband waardoor een stok achter de deur ontbrak om na de interventie door te gaan met het analyseren van leerresultaten. Ten tweede doordat de stap van de analyse van leerlingvolgsysteemdata naar groepsplan en uitvoering van het plan in de groep nog te weinig gemaakt werd. Mogelijk zijn dan andere didactische vaardigheden van belang, zoals het zorgdragen

voor een goed klassenmanagement, die wellicht niet altijd voldoende beheerst worden om goed te kunnen differentiëren (Doolaard, 2013).

De doelstelling van het project 'Streef Middenbouw' was het verbeteren van begrijpend leesresultaten door een nascholingsprogramma bestaand uit drie componenten; leerstandaarden en prestatiedoelen voor iedere leerling, opbrengstgericht werken en vakspecifieke kennis en instructiepraktijken. In dit project werden positieve effecten gevonden op leerresultaten. Door het project stelden leerkrachten significant vaker gefundeerde doelen. Leerlingen die in het voorgaande leerjaar laag hadden gescoord profiteerden extra van een relatief hoog leerdoel. Aan de hand van lesobservaties kon niet vastgesteld worden dat de instructiepraktijken van leerkrachten waren verbeterd, de mate waarin leerkrachten differentieerden verschilde bijvoorbeeld niet veel tussen de voor- en nameting (Van Kuijk, 2014).

OGW is een terugkerend onderwerp in onderzoeken van de Inspectie van het Onderwijs. In de Onderwijsverslagen uit 2011 en 2012 benadrukt de inspectie dat OGW een belangrijke sleutel is tot kwaliteitsverbetering. De inspectie baseert dit op onderzoek waaruit blijkt dat zwakke en zeer zwakke basisscholen relatief laag scoren op de inspectie-indicatoren voor OGW (Inspectie van het Onderwijs, 2011; 2012). De vijf indicatoren die de inspectie voor OGW hanteert zijn:

- De school gebruikt een samenhangend systeem van instrumenten en procedures voor het volgen van prestaties en ontwikkeling van leerlingen.
- Leerkrachten volgen en analyseren systematisch de vorderingen van leerlingen.
- De school gaat de effecten van de zorg na.
- De school evalueert jaarlijks systematisch de kwaliteit van opbrengsten.
- De school evalueert regelmatig het onderwijsleerproces.

Op scholen die door onderwijsinspecteurs beter beoordeeld werden op genoemde OGW-indicatoren bleken de leerlingen hogere leerresultaten te behalen.

Uit het bovenstaande volgt dat leerkrachten met behulp van een training de eerste twee elementen uit de OGW-cyclus (Figuur 1) beter gaan beheersen. Over de uitvoering van de daarop volgende stappen is minder bekend, alleen in de tweede deelstudie van het Streefproject is hiernaar gekeken en daaruit volgde niet dat leerkrachten meer gingen differentiëren. Een mogelijke verklaring zou kunnen zijn dat leerkrachten de daarvoor voorwaardelijke didactische vaardigheden niet altijd voldoende beheersen. Uit het Onderwijsverslag van het schooljaar 2012-2013 blijkt dat tussen 2009 en 2012 ongeveer 50% van de leerkrachten geen voldoende beoordeling kreeg op alle vier indicatoren voor differentiëren (planmatige differentiatie, en differentiatie in tijd, instructie en verwerkingsopgaven). Leerkrachten kregen daarbij het minst vaak een voldoende beoordeling voor differentiatie in de instructie (Inspectie van het Onderwijs, 2014).

In de onderzoeken zijn verschillende resultaten gevonden met betrekking tot de invloed van OGW op leerresultaten. In twee studies zijn positieve effecten gevonden van een OGW-interventie op leerresultaten (Van Kuijk, 2014; Van Geel, Keuning, Fox & Visscher, submitted), ook uit onderzoeken van de inspectie volgt dat op de scholen waar hogere scores zijn gegeven op de OGW-indicatoren gemiddeld hogere leerresultaten behaald worden in vergelijking met scholen die lagere scores kregen op de OGW-indicatoren (Inspectie van het Onderwijs, 2014). In het onderzoek van Doolaard (2014) worden daarentegen geen significante positieve effecten gevonden op leerresultaten.

3. ONDERZOEKSMETHODE

Er namen 26 scholen met 77 leerkrachten deel aan dit onderzoek. Met het oog op een goede verdeling over de groepen van het primair onderwijs is gekozen voor leerkrachten uit de groepen twee, vier en zeven. In paragraaf 3.1 wordt de selectie van de scholen toegelicht, in paragraaf 3.2.1 tot en met 3.2.5 de dataverzameling en de instrumenten die daarbij zijn gebruikt evenals de training van observanten. In paragraaf 3.3 vervolgens de betrouwbaarheid van de observatiescores en de stabiliteit van de scores van leerkrachten, en tot slot in paragraaf 3.4 de gehanteerde analysemethode.

3.1. WERVING EN SELECTIE

Alle scholen die aan dit onderzoek hebben deelgenomen zijn geselecteerd uit de Focus-projecten. Van de geselecteerde scholen is vervolgens aanvullende data verzameld, en is data gebruikt die binnen de Focus-projecten reeds verzameld was. Trainers van de Focus-projecten hebben elke school beoordeeld qua OGW-kwaliteit en elke school vervolgens geclassificeerd als een sterke, gemiddelde, of zwakke OGW school. Er is naar gestreefd om uit elk van deze drie categorieën tien scholen te werven zodat er voldoende variatie tussen de scholen zou bestaan in de mate van OGW-beheersing.

Scholen zijn met een brief op de hoogte gesteld van het onderzoek en een week later is telefonisch contact met hen opgenomen. Enkele scholen zijn daarnaast bezocht om het onderzoeksproject verder toe te lichten. Als tegenprestatie voor deelname hebben scholen na afloop van het onderzoeksproject feedback ontvangen op basis van gemaakte lesobservaties.

Aanvankelijk zijn 27 scholen en 81 leerkrachten gestart met het onderzoek (Tabel 2). Een meerderheid van deze scholen had de Focus-interventie afgerond, enkele scholen bevonden zich nog in het laatste traject daarvan. Er zijn acht scholen geworven uit de *zwakke* OGW-categorie, negen uit de categorie *gemiddeld*, en uit de categorie *sterk* deden tien scholen mee. Gedurende het onderzoek viel één school en nog eens één leerkracht uit met name doordat de lesobservaties met twee camera's als te belastend werden ervaren. Uit Tabel 2 volgt dat de deelnemende scholen een gemiddeld hoog percentage gewichten leerlingen hadden, het gemiddeld percentage van Nederlandse basisscholen is namelijk 11% (CBS, 2014). Zoals verwacht waren er veel combinatiegroepen in groep twee, ongeveer een derde van de groepen vier en zeven bestond uit combinatiegroepen. Van de deelnemende leerkrachten had 78 procent een duo collega, dit percentage was iets lager bij de leerkrachten uit groep zeven.

Tabel 2: Kenmerken geselecteerde scholen

<i>Groep</i>	<i>n</i> <i>Leerlingen</i>	<i>n</i> <i>Leerkrachten</i>	% <i>Gewichten</i> <i>leerlingen</i>	% <i>Combinatie-</i> <i>groepen</i>	% <i>Duo's</i>
2	437	26	16.7%	88%	85%
4	506	25	18.8%	32%	84%
7	544	26	18.5%	35%	65%
Totaal	1487	77	18%	52%	78%

3.2. DATAVERZAMELING EN ONDERZOEKSTRUMENTEN

Voor het beantwoorden van de onderzoeksvragen zijn de competenties, behorend bij de vier componenten van OGW (Tabel 1), gemeten evenals de leerresultaten. Hiervoor zijn vier instrumenten ingezet; de leerkrachtvragenlijst die in de Focus-interventie werd afgenomen om kennis en attitudes bij leerkrachten te meten, een instrument voor de beoordeling van groepsplannen, het lesobservatie-instrument *International Comparative Analysis of Learning and Teaching* (ICALT), en tot slot de CITO leerlingvolgsteemtoetsen.

3.2.1. Leerkrachtvragenlijst

Resultaten uit de leerkrachtvragenlijst zijn gebruikt voor het meten van attitude en kennis componenten. Met 21 items werd de attitude van leerkrachten gemeten ten aanzien van OGW (algemeen) ($\alpha=0.92$), het gebruiken van een leerlingvolgsysteem (LVS) voor analyses ($\alpha=0.79$), en het stellen van prestatiedoelen ($\alpha=0.86$). Deze items zijn gemeten op een 4-puntsschaal en in de analyses zijn de gemiddelde scores gebruikt op de items. Voor het meten van kennis over het gebruik van een LVS is een 3-puntsschaal gehanteerd ($\alpha=0.68/0.74$ afhankelijk van het type leerlingvolgsysteem). In de analyses zijn de percentages goed beantwoorde vragen meegenomen. De waarden voor de interne consistentie van de items binnen de constructen zijn afkomstig uit de resultaten van Focus 1 (Staman, Visscher & Luyten, 2013). Voorbeelden van vragen uit de vragenlijst zijn:

- Het stellen van prestatiedoelen legt te veel nadruk op de leerprestaties van leerlingen (*attitude doelen*)
- Het gebruik van een leerlingvolgsysteem is een activiteit waar ik meer tijd in wil investeren (*attitude LVS*)
- Voor het bewaken en verbeteren van resultaten van mijn/ons onderwijs is OGW (*attitude OGW*)
Aan te raden 0 0 0 0 Af te raden
- Met ons leerlingvolgsysteem kunnen we zien of leerlingen wel overeenkomstig hun intelligentie presteren (ja/nee/weet ik niet) (*kennis LVS*)

De resultaten uit de leerkrachtvragenlijst van de laatste meting zijn gebruikt omdat deze het best aansluiten bij de huidige situatie. De vragenlijst is niet bij alle leerkrachten afgenomen onder andere door wisselingen in het leerkrachtenteam na de Focus-interventie. In de analyses zijn gegevens uit 20 vragenlijsten van leerkrachten uit groep twee, van 16 leerkrachten uit groep vier, en gegevens van 17 leerkrachten uit groep zeven meegenomen.

3.2.2. Beoordelingsinstrument groepsplannen

Het beoordelingsinstrument is ontwikkeld voor het meten van de vaardigheden met betrekking tot de eerste drie componenten uit Tabel 1. Het volledige instrument is terug te vinden in Bijlage B van dit rapport evenals het Focus format voor groepsplannen (Bijlage C). Het instrument is ontwikkeld binnen het Focus-project en intern meerdere malen besproken en aangepast. Het instrument is

zolang doorontwikkeld totdat een selectie groepsplannen op een consistente en overeenkomstige wijze beoordeeld werd door verschillende onderzoekers.

Met het instrument kan zowel het protocol als bijbehorend groepsplan beoordeeld worden. Het protocol is een richtlijn voor leerkrachten om de verschillende analysemogelijkheden van het LVS goed te benutten. In het groepsplan staan de instructiegroepen waarin leerlingen zijn geplaatst, welke leerdoelen de leerkracht voor leerlingen stelt en hoe de leerkracht inspeelt op vastgestelde onderwijsbehoeften (qua instructie-aanpak). Het groepsplan omvat de instructiestrategie van de leerkracht voor het onderwijs aan alle leerlingen in zijn/haar groep en welke doelen men wil behalen.

In de Focus-training maakten leerkrachten naar aanleiding van de Cito-toetsen twee keer per schooljaar een groepsplan en vulden ze een protocol in. De aan dit onderzoek deelnemende leerkrachten is dan ook gevraagd naar twee plannen en een ingevuld protocol. Op veel scholen werden de protocollen echter niet meer gebruikt en in totaal werden er 29 protocollen aangeleverd (*missing* 81%). Groepsplannen daarentegen werden nog veel gebruikt, er zijn 39 groepsplannen van groep twee (*missing* 21%), 46 groepsplannen van groep vier (*missing* 8%) en 43 plannen van groep zeven (*missing* 17%) verzameld. Voor het opstellen van het groepsplan gebruikten de leerkrachten in de training een vast format (Bijlage C), nog 26 leerkrachten gebruikten datzelfde format. Veel scholen hadden inmiddels een eigen format in gebruik, of een format vanuit een LVS.

Alle aangeleverde protocollen en groepsplannen zijn door twee observatoren gescoord aan de hand van het beoordelingsinstrument. Het instrument bestond uit de volgende vijf onderdelen:

1. Kwaliteit analyses en interpretatie: de kwaliteit van de uitgevoerde analyses met behulp van het LVS en de mate waarin de analyse uitkomsten juist zijn geïnterpreteerd ($\alpha=0.68$).
2. Kwaliteit instructiebeslissingen: de mate waarin er een verbinding was tussen de analyse uitkomsten uit het LVS met de instructiebeslissingen, en de kwaliteit van de instructiebeslissingen ($\alpha=0.59$).
3. Kwaliteit doelen: de mate waarin doelen specifiek en gedifferentieerd waren en aansloten bij meest recente leerresultaten ($\alpha=0.66$).
4. Evaluatie studievoortgang: de mate waarin systematisch geëvalueerd was en de kwaliteit en kwantiteit van toegepaste evaluatie-instrumenten ($\alpha=0.50$).
5. Subgroep specifieke items: de mate waarin er een verbinding was tussen de analyse uitkomsten uit het LVS en het opstellen van de subgroepen, en de kwaliteit van de instructiebeslissingen voor de subgroepen (leerlingen die ingedeeld zijn in subgroepen krijgen verlengde instructie gericht op een specifieke leerstofcategorie) ($\alpha=0.90$).

Elk onderdeel bestond uit items en sub-items die met een 0 of 1 gescoord zijn. Voor elk onderdeel is de som score bepaald en vervolgens is deze score geaggregeerd. De totale som score van het instrument is opgenomen in de analyse, met uitzondering van de som score voor het onderdeel 'kwaliteit analyses en interpretatie' vanwege het hoge percentage *missings*.

Uit het groepsplan zijn tevens gegevens gehaald over de indeling van leerlingen in instructiegroepen. De indeling in groepen is afkomstig uit het groepsplan van de periode januari/februari 2013/2014 tot en met het einde van het schooljaar. In het primair onderwijs is het gebruikelijk om te werken met drie verschillende instructiegroepen. Er zijn verschillende benaming mogelijk voor deze instructiegroepen, in dit rapport houden we de volgende aan:

- instructieafhankelijk,
- instructiegevoelig,
- en instructieonafhankelijk.

In de instructieafhankelijke groep zitten leerlingen die in vergelijking met de andere leerlingen in de *eigen* groep beneden gemiddeld presteren, dit zijn de leerlingen die in principe in aanmerking komen voor een verlengde instructie. Leerlingen die in de instructiegevoelige groep zitten volgen doorgaans de basisinstructie en presteren gemiddelde in de *eigen* groep. Als laatste is er een groep leerlingen die instructieonafhankelijk genoemd wordt. Zij presteren in vergelijking met de andere leerlingen van de *eigen* groep bovengemiddeld, en komen in principe in aanmerking voor het volgen van een verkorte instructie.

3.2.3. Lesobservatie instrument ICALT

OGW-vaardigheden in termen van het uitvoeren van de geplande instructie strategieën (Tabel 1: strategie uitvoeren) zijn gemeten met lesobservaties. De betrouwbaarheid van observaties verhoogt aanzienlijk wanneer meerdere lessen van dezelfde leerkracht geobserveerd worden door verschillende observanten (Hill, Charalambous & Kraft, 2012). Daarom werd ernaar gestreefd elke leerkracht drie keer te observeren gedurende het schooljaar, en deze observaties door drie verschillende getrainde observanten te laten beoordelen. In totaal zouden er dan 231 lessen opgenomen zijn die elk drie keer beoordeeld zouden worden. Vooraf werden met leerkrachten afspraken gemaakt over de observatie, zodat gefilmde lessen bestonden uit een instructie, verwerking en lesafsluiting. Gefilmde lessen duurden over het algemeen tussen de 45 en 60 minuten voor groep vier en zeven, de opnames in groep twee waren vaak korter omdat de gegeven instructies en opdrachten korter duurden (ongeveer 15 tot 30 minuten). Opnames werden gemaakt met het systeem IRIS-Connect waarmee zowel de leerkracht als de leerlingen gelijktijdig in beeld worden gebracht. Er zijn alleen lessen gefilmd waarin rekenen-wiskunde gegeven werd. Door wisseling en/of veranderingen binnen scholen is niet elke leerkracht driemaal gefilmd (*missing* 8.2% op totaal van 231).

Voor het beoordelen van de lessen is het ICALT instrument gebruikt, dit instrument bestaat uit 35 items die gescoord worden op een 4-punt Likertschaal (van 'overwegend zwak' tot 'overwegend sterk') (Van de Grift, 2007). In de analyse is de totaalscore van het gehele instrument opgenomen en het construct 'afstemmen van instructie en verwerking' ($\alpha=0.81$) apart.

Het construct 'afstemmen van instructie en verwerking' bestond uit de volgende drie items:

- de leerkracht biedt zwakke leerlingen extra leer- en instructietijd,
- de leerkracht stemt instructie af op relevante verschillen tussen leerlingen,
- en de leerkracht stemt de verwerking van de leerstof af op relevante verschillen tussen leerlingen.

De vermelde Cronbach's Alpha voor het weergeven van de interne consistentie tussen de items binnen constructen zijn afkomstig uit de huidige studie. Uit deze analyse bleek dat item 23 (de leerkracht gaat na of de lesdoelen werden bereikt) de betrouwbaarheid van het bijbehorende construct verlaagden, scores op dit item zijn daarom niet meegenomen in verdere analyses.

3.2.4. Leerresultaten

Scholen hebben de CITO LVS Rekenen-Wiskunde resultaten van groep vier en zeven, en de CITO-LVS Rekenen-voor-Kleuters resultaten van groep twee aangeleverd. In de analyses zijn de resultaten op de eindtoetsen van het schooljaar 2012/2013 als controle variabele (voortests) opgenomen, en de resultaten op de eindtoetsen van het schooljaar 2013/2014 als afhankelijke variabele. Van 437 leerlingen uit groep twee zijn de leerresultaten opgenomen in de analyses (daarvan missing voortest 15.8%, en 10.8% eindtoets), uit groep vier van 506 leerlingen (daarvan missing voortest 8.1%, en eindtoets 8.9%) en van 544 leerlingen uit groep zeven (daarvan missing voortests 13.4%, en eindtoets 16.2%).

3.2.5. Training observanten

Twee onderzoekers zijn getraind in het toepassen van het instrument voor de beoordeling van de protocollen/groepsplannen, en drie onderzoekers voor het beoordelen van de lesobservaties. Aan de trainingsbijeenkomsten vooraf had elke onderzoeker dezelfde observaties en groepsplannen/protocollen beoordeeld. Vervolgens werden de uitkomsten daarvan geanalyseerd om zo de verschillen in beoordelingen tussen beoordelaars vast te stellen. De gevonden verschillen stonden centraal in de bijeenkomsten, onderzocht werd waardoor (grote) verschillen ontstonden, en vervolgens werden beoordelingsafspraken gemaakt om grote verschillen te voorkomen. Er vonden diverse discussieronden plaats, net zolang totdat er in grote mate overeenstemming was tussen de beoordelingen. Naar aanleiding van de discussieronden werd het instrument voor de beoordeling van de protocollen/groepsplannen steeds aangepast en verbeterd.

3.3. BETROUWBAARHEID OBSERVANTEN EN LEERKRACHT STABILITEIT

De variantie in leerkrachtscores op basis van lesbeoordelingen zal doorgaans niet alleen bepaald worden door verschillen tussen leerkrachten, maar ook door verschillen tussen observanten en tijdsmomenten. De variantie veroorzaakt door verschillen tussen observanten en tijdsmomenten zou in verhouding tot de totale variantie klein moeten zijn, in dat geval beoordelen observatoren leerkrachten gelijkwaardig en is de kwaliteit van leerkrachten stabiel over de tijdsmomenten heen. Een gemiddelde score van de tijdsmomenten en de observanten is dan een betrouwbare score; het overgrote deel van de variantie kan worden toegeschreven aan verschillen tussen leerkrachten.

De *intraclass correlatie coëfficiënt* (ICC) is een maat voor het bepalen van de betrouwbaarheid van de oordelen van observanten en de stabiliteit van leerkrachtbeoordelingen. De ICC wordt bepaald door middel van een vergelijking van de variantie veroorzaakt door verschillen tussen observatoren, of tijdsmomenten, met de totale variantie (Snijders & Bosker, 1999) De ICC kan als volgt geïnterpreteerd worden:

- 0.00-0.20 = zeer geringe overeenstemming
- 0.21-0.40 = geringe overeenstemming,
- 0.41-0.60 = matige overeenstemming,
- 0.61-0.80 = goede overeenstemming,
- 0.81-1.00 = zeer goede overeenstemming (Landis & Koch, 1977).

Voor het bepalen van de ICC is uitgegaan van een *two-way mixed* model en is de consistentie in scores geanalyseerd (Tabel 3).

Tabel 3: ICC observatoren en leerkracht stabiliteit

<i>Lesobservaties</i>	<i>ICC observatoren (95% bhi) n=212</i>	<i>ICC stabiliteit leerkrachten (95% bhi) n=174</i>
Afstemmen van instructie en verwerking	0.73 (0.68-0.76)	0.63 (0.57-0.68)
<hr/>		
<i>Groepsplan</i>	<i>ICC observanten (95% bhi) n=128</i>	<i>ICC leerkracht stabiliteit (95% bhi) n=53</i>
Kwaliteit instructiebeslissingen	0.68 (0.64-0.71)	0.73 (0.70-0.75)
Kwaliteit doelen	0.8 (0.76-0.83)	0.72 (0.66-0.77)
Evaluatie studievoortgang	0.69 (0.64-0.74)	0.83 (0.8-0.86)
Subgroep specifieke items	0.47 (0.41-0.53)	0.63 (0.64-0.71)

Uit Tabel 3 volgt dat sprake is van een goede stabiliteit van de leerkrachtkwaliteit over de drie tijdstippen heen, ook de overeenstemming tussen de lesobservatiescores van de observanten is goed. De overeenstemming tussen beoordelaars is iets hoger bij de groepsplannen. De betrouwbaarheid van het onderdeel subgroep specifieke items is alleen matig. Bij de andere onderdelen is sprake van een goede overeenstemming tussen observanten en zijn de scores van leerkrachten over de twee tijdstippen heen stabiel.

3.4. ANALYSES

Voor het beantwoorden van de derde onderzoeksvraag is een multilevel analyse uitgevoerd. In dit model is uitgegaan van twee niveaus; het niveau van de leerkracht en het leerling niveau. De volgende variabelen zijn opgenomen:

- Leerresultaten eind schooljaar 2013/2014 (afhankelijke)
- Leerresultaten eind schooljaar 2012/2013 (controle)
- Geslacht leerling (controle)
- Leerlingengewicht als maat voor sociaal economische status (controle)
- Combinatiegroep (controle)
- Afstemmen van de instructie (voorspeller)
- Totaal score op ICALT (voorspeller)
- Totaal score op het groepsplan (voorspeller)
- Verdeling in instructiegroepen (interactie)

In groep twee wordt een andere toets gebruikt dan in groep vier en zeven daarom zijn de analyses voor groep twee apart uitgevoerd.

4. RESULTATEN

Paragraaf 4.1 bevat de beschrijvende resultaten van de leerkrachtvragenlijst, de lesobservaties, evenals de resultaten van het instrument waarmee de groepsplannen beoordeeld zijn. In de laatste paragraaf van dit hoofdstuk worden de resultaten gepresenteerd van de multilevel analyse.

4.1. BESCHRIJVENDE RESULTATEN

Leerkrachtvragenlijst

In Tabel 4 zijn de scores uit de leerkrachtvragenlijst vermeld, hieruit volgt dat de leerkrachten in het onderzoek een positieve attitude hebben ten opzicht van OGW in het algemeen, het gebruiken van een LVS, en het gebruiken van prestatiedoelen. Op de kennisvragen over het LVS wordt gemiddeld 70 procent van de vragen goed beantwoord door leerkrachten in groep twee, 81.4 procent door leerkrachten in groep vier en 81.6 procent goed beantwoorde vragen door leerkrachten in groep zeven. In groep zeven was de spreiding in leerkrachtscores, met betrekking tot kennis over het LVS, aanzienlijk lager dan in groep twee en vier.

Tabel 4: Scores leerkrachtvragenlijst

<i>Onderdeel</i>	<i>Groep 2</i> $\bar{X}(SD)$	<i>Groep 4</i> $\bar{X}(SD)$	<i>Groep 7</i> $\bar{X}(SD)$
Attitude OGW*	3.50 (0.54)	3.43 (0.55)	3.42 (0.36)
Attitude LVS*	3.05 (0.49)	2.94 (0.57)	3.01 (0.44)
Attitude prestatiedoelen*	2.91 (0.54)	2.63 (0.46)	2.78 (0.36)
Kennis LVS**	70% (25.4)	81.4% (20.9)	81.6% (9.9)

*De attitude items zijn beoordeeld op een schaal van 1 tot en met 4.

**Bij het onderdeel Kennis LVS is het gemiddelde percentage goed beantwoorde vragen vermeld.

Lesobservaties en groepsplannen

De scores op de lesobservaties zijn geaggregeerd tot één score per construct per leerkracht (Smolkowski & Gunn, 2012). In Tabel 5 staan de resultaten.

Uit de resultaten volgt dat de lesobservatiescores op een meerderheid van de constructen rond de drie zijn gescoord (= meer sterk dan zwak). Daarnaast valt op dat leerkrachten uit groep twee in vergelijking met de leerkrachten uit groep vier en zeven lager scoren op het construct 'afstemmen van instructie en verwerking'. Gemiddeld scoren alle leerkrachten respectievelijk laag op het construct 'leerstrategieën aanleren'. De opbouw van het instrument ICALT, van de eenvoudige naar de meer complexere vaardigheden, is duidelijk terug te zien in de scores.

De scores gegeven voor het groepsplan zijn lastiger te interpreteren doordat de maximale somscores per groepsplan onderdeel verschillen. Om die reden zijn percentages berekend per onderdeel. Voor de meeste onderdelen geldt dat gemiddeld niet meer dan de helft van het maximale aantal punten verkregen is. Leerkrachten uit groep twee scoren over het algemeen iets lager dan de leerkrachten uit groep vier en zeven, met name voor het onderdeel 'doelen' is dit verschil groot. Op het onderdeel 'doelen' scoren de leerkrachten uit groep vier en zeven in vergelijking met de andere

onderdelen relatief hoog. Voor het onderdeel ‘subgroep specifieke items’ zijn de scores lager in vergelijking met de andere onderdelen van het groepsplan.

Tabel 5: Scores lesobservaties en groepsplannen

<i>Construct ICALT</i>	<i>Groep 2</i> $\bar{X}(SD)$	<i>Groep 4</i> $\bar{X}(SD)$	<i>Groep 7</i> $\bar{X}(SD)$
Veilig en stimulerend leerklimaat	3.34 (0.24)	3.26 (0.20)	3.33 (0.26)
Efficiënte lesorganisatie	3.04 (0.31)	3.12 (0.22)	3.15 (0.26)
Duidelijke en gestructureerde instructie	2.96 (0.27)	3.01 (0.17)	3.09 (0.20)
Intensieve en activerende les	2.76 (0.27)	2.8 (0.16)	2.92 (0.24)
Afstemmen van instructie en verwerking	1.68 (0.45)	2.54 (0.32)	2.67 (0.36)
Leerstrategieën aanleren	2.08 (0.32)	1.96 (0.21)	2.29 (0.30)
Betrokkenheid van leerlingen	2.86 (0.32)	2.83 (0.17)	2.87 (0.28)
Totaal score ICALT*	18.66 (1.71)	19.46 (1.08)	20.31 (1.51)

<i>Onderdeel groepsplan</i>	<i>Groep 2</i> $\bar{X}(SD)$	<i>Groep 4</i> $\bar{X}(SD)$	<i>Groep 7</i> $\bar{X}(SD)$
Kwaliteit instructiebeslissingen	6.31 (1.8) 44%	6.73 (1.8) 47%	6.55 (2.02) 45%
Kwaliteit doelen	1.72 (1.03) 42%	2.87 (0.97) 73%	2.60 (1.22) 63%
Evaluatie studievoortgang	2.16 (1.08) 43%	2.52 (1.02) 51%	2.7 (1.05) 53%
Subgroep specifieke items	7.14 (5.05) 36%	6.34 (4.74) 32%	5.34 (5.22) 26%

* Bij een maximale score op elk construct worden 28 punten behaald.

4.2.EFFECTEN OP LEERRESULTATEN

De resultaten van het effect op leerresultaten van groep twee zijn terug te vinden in Tabel 6 en de resultaten van groep vier en zeven in Tabel 7. De resultaten van groep twee worden eerst besproken.

Tabel 6: Resultaten groep twee

<i>Variabelen</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>
<i>Leerlingniveau</i>					
Intercept	94.38** (0.86)	45.09** (4.03)	45.94** (4.71)	58.07** (4.90)	57.53** (4.95)
Cito 2013		0.62** (0.04)	0.63** (0.05)	0.37** (0.06)	0.37** (0.06)
Geslacht (meisje)		1.62 (0.99)	1.95* (1.07)	2.29* (1.00)	2.16* (1.00)
Leerlingengewicht (geen)		4.09** (1.38)	3.99** (1.51)	3.53** (1.41)	3.52** (1.42)
<i>Groepsniveau</i>					
Combinatiegroep (ja)		-3.21 (2.39)	-4.74 (3.29)	-3.99 (3.04)	-3.83 (3.12)
Afstemmen			-1.03 (1.25)	-0.89 (1.17)	-0.02 (1.58)
Totaal score ICALT			0.15 (1.29)	0.69 (1.20)	0.63 (1.23)
Groepsplan			-0.16 (1.19)	-0.15 (1.1)	-0.24 (1.13)
<i>Interactievariabele</i>					
Instructiegroep (gevoelig)				9.03** (1.41)	9.36** (1.46)
Instructiegroep (onafh.)				12.11** (1.83)	12.23** (1.88)
Gevoelig x Afstemmen					-1.82 (1.47)
Onafh. x Afstemmen					0.06 (1.62)
<i>Variantie leerlingniveau</i>	140.46** (10.44)	78.39** (6.39)	83.05** (7.09)	71.86** (6.19)	71.49** (6.17)
<i>Variantie groepsniveau</i>	8.97 (5.74)	13.21* (6.34)	13.14* (7.49)	11.03* (6.25)	11.95* (6.53)

* $p < 0.05$, ** $p < 0.01$

De analyse is stapsgewijs uitgevoerd. In het eerste model is alleen de afhankelijk variabele opgenomen, in het tweede model zijn de controle variabelen toegevoegd en vervolgens in het derde, vierde en vijfde model de onafhankelijke variabelen en de interactie variabelen.

Uit Tabel 6 volgt dat de Cito toets uit 2013 een goede voorspeller is voor de resultaten op de Cito toets in 2014. Ook blijken meisjes in groep twee gemiddeld hoger te scoren op rekenen dan jongens, in de laatste drie modellen is dit effect ook significant. Leerlingen zonder een gewicht behalen hogere scores dan de leerlingen met een gewicht (1.2 of 0.3), het effect van leerlingengewicht is in alle modellen significant. In de laatste controle variabele worden combinatiegroepen vergeleken met 'normale' groepen, de richting van dit effect is negatief en duidt op gemiddeld lagere resultaten in combinatiegroepen. Dit effect is echter in geen van de modellen significant.

In groep twee zijn geen significante effecten gevonden van de onafhankelijke variabelen. Uit Tabel 6 volgt dat de richting van het construct afstemmen van de instructie en verwerking negatief is, evenals de variabele waarmee de compleetheid van het groepsplan gemeten is. Het effect van de totale score op ICALT is daarentegen positief. Voor de indeling in instructiegroepen zijn wel significante resultaten gevonden. In vergelijking met de instructieafhankelijke en onafhankelijke groep behalen leerlingen ingedeeld in de instructiegevoelige groep hogere resultaten. Hetzelfde geldt voor de leerlingen die ingedeeld zijn in de instructieonafhankelijk groep, zoals verwacht is het effect van deze groep het grootst dit zijn immers de leerlingen die op basis van hun leerbehoeften zijn ingedeeld in een verkorte instructiegroep. Er zijn geen significantie interactie-effecten gevonden. De richting van de effecten impliceert wel dat instructie gevoelige leerlingen in vergelijking met de

andere leerlingen minder profiteren van een leerkracht die hoge scores heeft op het afstemmen van de instructie, en dat dit niet geldt voor leerlingen in de instructieonafhankelijke groep.

Tabel 7: Resultaten groep 4 en groep 7

Variabelen	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Leerlingniveau</i>					
Intercept	85.45**(2.95)	32.24**(2.44)	32.25**(2.43)	53.14**(2.85)	53.14**(2.86)
Cito 2013		0.80**(0.02)	0.80**(0.02)	0.49**(0.03)	0.49**(0.03)
Geslacht (meisje)		-2.01**(0.62)	-2.03**(0.62)	-2.16**(0.58)	-2.17**(0.59)
Leerlingengewicht (geen)		1.08 (0.87)	1.22 (0.87)	1.30 (0.81)	1.30 (0.81)
<i>Groepsniveau</i>					
Combinatiegroep (ja)		-0.75 (1.22)	-1.53 (1.26)	-0.88 (1.25)	-0.88 (1.25)
Groep 4		-4.66**(1.57)	-4.86**(1.57)	-18.62**(1.88)	-18.62**(1.88)
Afstemmen			-1.00 (0.66)	-0.84 (0.68)	-0.78 (0.93)
Totaal score ICALT			0.91 (0.62)	0.77 (0.68)	0.78 (0.68)
Groepsplan			-0.86*(0.51)	-0.83 (0.52)	-0.82 (0.53)
<i>Interactievariabele</i>					
Instructiegroep (gevoelig)				5.64**(0.86)	5.64**(0.87)
Instructiegroep (onafh.)				13.46**(1.12)	13.48**(1.13)
Gevoelig x Afstemmen					-0.13 (0.86)
Onafh. x Afstemmen					-0.04 (0.87)
<i>Variantie leerlingniveau</i>	184.67**(8.72)	78.02**(3.87)	78.07**(3.88)	64.63**(3.34)	64.79**(3.36)
<i>Variantie groepsniveau</i>	430.77**(88.35)	10.29**(3.14)	9.17**(3.04)	9.77**(3.07)	9.81**(3.09)

* $p < 0.05$, ** $p < 0.01$

In Tabel 7 staan de resultaten voor groep vier en zeven. De Cito toets eind 2013 is weer een goede voorspeller voor de resultaten eind 2014. Net als in groep twee is het effect van deze controle variabele in alle modellen significant. In plaats van meisjes behalen juist jongens in groep vier en zeven gemiddeld hogere resultaten voor rekenen-wiskunde, het effect is in alle modellen weer significant. De invloed van het leerlingengewicht is in groep vier en zeven minder groot dan in groep twee, bovendien is het effect niet significant. Hetzelfde geldt voor combinatiegroepen, de richting van dit effect is negatief maar is niet significant. Als laatst worden de leerlingen uit groep vier vergeleken met de leerlingen uit groep zeven, zoals verwacht behalen leerlingen uit groep vier lagere resultaten.

Voor de onafhankelijke variabelen geldt een vergelijkbaar resultaat als in groep 2, alleen is het effect van het groepsplan significant in het derde model. De richting van de variabelen zijn eveneens vergelijkbaar. In de analyse zijn negatieve coëfficiënten gevonden voor afstemmen in de groep en voor het groepsplan, en een positieve coëfficiënt voor de totaal score op ICALT. Ook behalen leerlingen ingedeeld in de instructie gevoelige groep gemiddeld hogere scores in vergelijking met de andere leerlingen, en hetzelfde geldt voor de leerlingen ingedeeld in de

instructieonafhankelijk groep. Voor de interactie tussen de indeling in instructiegroepen en afstemmen in de groep is geen significant resultaat gevonden. In tegenstelling tot groep twee is in groep vier en zeven de richting van de interactie tussen de instructie onafhankelijke groep en afstemmen negatief, wat inhoudt dat ook deze groep minder profiteert van een leerkracht die afstemt in vergelijking met de leerlingen in de andere instructiegroepen. Dit effect is echter niet significant.

5. CONCLUSIES EN DISCUSSIE

In het laatste hoofdstuk worden de onderzoeksvragen beantwoord en conclusies getrokken op grond van de gevonden resultaten. Daarna wordt de gehanteerde onderzoeksmethode bediscussieerd.

Het onderzoek heeft waardevolle inzichten opgeleverd over OGW in het primair onderwijs en de zaken waarmee leerkrachten ten aanzien van OGW nog problemen ondervinden. Deze inzichten zijn verkregen door het in kaart brengen van:

- de OGW-attitude van leerkrachten,
- de kennis van leerkrachten over de analysemogelijkheden die een LVS biedt,
- de kwaliteit van 128 groepsplannen,
- en de kwaliteit van 212 lessen.

De eerste onderzoeksvraag luidde in hoeverre leerkrachten de voor elk van de componenten van OGW vereiste attitudes, kennis en vaardigheden beheersen (Tabel 1). Op basis van de resultaten van de leerkrachtvragenlijst kon deze vraag beantwoord worden. Deelnemende leerkrachten blijken overwegend een positieve attitude te hebben ten opzichte van OGW. Leerkrachten oordelen iets positiever over OGW in het algemeen dan over de meer specifieke onderdelen van OGW; het gebruiken van een leerlingvolgsysteem (LVS) en prestatiedoelen, maar ook over deze onderdelen zijn leerkrachten overwegend positief. Tussen de drie leeftijdsgroepen bestaan geen noemenswaardige verschillen. Wel is deze vragenlijst alleen afgenomen onder leerkrachten van basisscholen die ook hebben deelgenomen aan een OGW-interventie en daardoor bij voorbaat al positiever zouden kunnen zijn over OGW in vergelijking met leerkrachten op andere basisscholen.

Voor een goede implementatie van OGW is kennis over de analysemogelijkheden die een LVS biedt noodzakelijk. Uit de resultaten van de leerkrachtvragenlijst blijkt dat leerkrachten na afloop van de Focus-interventie voldoende scores op de kennisvragen over welke analyses zij kunnen maken met behulp van een LVS. De spreiding in scores is het kleinst binnen de groep leerkrachten uit groep zeven; een meerderheid van deze leerkrachten behaalt een score van 71% of meer juist beantwoorde vragen. De spreiding binnen leerkrachten uit groep twee en vier is groter, een meerderheid van de leerkrachten uit groep twee behaalt een score hoger dan 42%, en in groep vier beantwoordt een meerderheid van de leerkrachten meer dan 60% vragen juist.

Deelnemende leerkrachten hadden dus overwegend een positieve OGW-attitude en voldoende kennis van de analysemogelijkheden binnen een LVS, ook in andere onderzoeken wordt gevonden dat leerkrachten door middel van een training uitgebreidere en betere analyses maken van leerresultaten (Doolaard, 2013; Staman, Visscher & Luyten, 2014).

De tweede onderzoeksvraag betreft de mate waarin de diverse OGW-componenten uitgevoerd worden in de onderwijspraktijk. Bevindingen naar aanleiding van de lesobservaties en de analyse van groepsplannen geven een indicatie van de beheersing van de eerste drie componenten uit Tabel 1. Op het lesobservatie-onderdeel 'afstemmen van instructie en verwerking' scoren de leerkrachten uit groep vier en zeven vergelijkbaar. In groep vier ligt ongeveer 68% van de gegeven scores tussen de 2.22 en de 2.86, terwijl in groep zeven ongeveer 68% van de gegeven scores tussen 2.31 en de 3.03 ligt. De gegeven scores voor leerkrachten uit groep twee liggen iets lager, ongeveer 68% van de

scores ligt in deze groep tussen de 1.23 en de 2.13. De scores liggen op een schaal van één (overwegend zwak) tot en met vier (overwegend sterk) en de gegeven scores van de leerkrachten uit groep vier en zeven liggen overwegend tussen de 'meer zwak dan sterk' (2) en 'meer sterk dan zwak' (3), terwijl de scores voor de leerkrachten uit groep twee iets lager liggen en voornamelijk als 'meer zwak dan sterk' (2) zijn beoordeeld.

De resultaten van de beoordeling van de groepsplannen indiceren dat veel leerkrachten moeite hebben met het bepalen van een instructiestrategie op papier. Gemiddeld wordt niet meer dan de helft van het maximale aantal punten behaald. Dit houdt in dat leerkrachten op papier minimaal hun instructiebeslissingen afstemmen op de uitkomsten van de analyses van de leerresultaten, en dat er minimaal gedifferentieerd wordt in de instructie, verwerkingsopdrachten, en in de in te zetten materialen. Ook ontbreekt in ongeveer de helft van de plannen een beschrijving van de afspraken omtrent de evaluatie van de effecten van de in het groepsplan voorgenomen instructiestrategie, en de voor dergelijke evaluaties te gebruiken instrumenten.

Opvallend is het beperkte gebruik van een categorieënanalyse binnen het LVS. Belangrijk daarbij is te weten dat niet elk LVS de mogelijkheid voor een categorieënanalyse biedt. Leerkrachten uit groep vier en groep zeven behaalden gemiddeld voldoende punten op het onderdeel 'kwaliteit van de doelen'. Een omschrijving van de resultaten op het onderdeel 'kwaliteit analyses en interpretatie' wordt hier buiten beschouwing gelaten vanwege het hoge percentage *missings*, dit zou wel kunnen impliceren dat leerkrachten dit onderdeel vaak achterwege laten bij het opstellen van een groepsplan. Hoewel leerkrachten op basis van de lesobservaties overwegend een voldoende score toebedeeld kregen voor het 'afstemmen van de instructie en verwerking' blijkt uit de groepsplanscores dat de OGW-component 'bepalen van de instructiestrategie' in het groepsplan weinig aandacht krijgt (zie ook Inspectie van het Onderwijs, 2014).

In de multilevel analyses is geen positief verband gevonden tussen de lesobservatiescores voor het afstemmen van de instructie en verwerking en de groepsplanscores met leerresultaten. Voor de groepsplanscores is een negatief effect gevonden in het derde model in de groepen vier en zeven, in de daarop volgende modellen was dit effect overigens niet meer significant. Een mogelijke verklaring voor deze resultaten zou kunnen zijn dat met de gebruikte instrumenten hoofdzakelijk de frequentie waarmee een aantal differentiatieactiviteiten worden uitgevoerd is gemeten en niet de *kwaliteit* daarvan. Een hoge frequentie van differentiatieactiviteiten hoeft niet altijd samen te gaan met een goede differentiatiekwaliteit.

Observanten gaven aan dat in veel groepen de instructiegroep voor de instructieafhankelijke leerlingen wel duidelijk zichtbaar was binnen de klas, maar dat goed afgestemd onderwijs vaak nog ontbrak. Leerlingen in deze groep maakten vaak apart de sommen, of samen met de leerkracht, zelden kregen zij een herhaling van de instructie, een aangepaste instructie, of werden opdrachten met behulp van extra materialen gemaakt. Dit sluit aan bij de bevindingen van het beoordelen van de groepsplannen waar ook uit bleek dat weinig gedifferentieerd wordt in de instructiebeslissingen. In een effectieve OGW-aanpak is een instructieaanpak die aansluit bij de specifieke leerbehoeften van de leerlingen in de instructieafhankelijke groep wel noodzakelijk.

De onderzoeksresultaten lijken de opvatting dat groepsplannen ook een onbedoeld neveneffect kunnen veroorzaken te bevestigen. Scholen denken door het indelen van leerlingen in de drie instructiegroepen alleen, al een voldoende aanpak te hanteren maar in een effectieve OGW-aanpak kan een nadere probleemanalyse, aanvullend diagnostisch onderzoek en het op de individuele leerling afgestemd onderwijs niet achterwege blijven (Inspectie van het Onderwijs, 2014).

Een opvallend onderzoeksresultaat is de richting van de interactie tussen de instructiegevoelige groep leerlingen met het construct 'afstemmen van instructie en verwerking' uit de lesobservaties. Deze bevinding lijkt erop te wijzen dat leerlingen in de instructiegevoelige groep het minste profijt hebben van een leerkracht met een hoge score op 'afstemmen van de instructie en verwerking'. Een vergelijkbaar resultaat werd gevonden voor de leerlingen in de instructieonafhankelijke groep (leerlingen die normaliter een verkorte instructie volgen). Ook observatoren gaven aan dat het onderwijs nog maar beperkt afgestemd wordt op de behoeften van deze specifieke groep leerlingen. Soms geven leerkrachten aan dat deze leerlingen eerder mogen beginnen met opgaven, maar hier blijft het vaak bij. Enkele leerkrachten spreken tijdens de instructie wel andere verwachtingen uit voor deze groep, bijvoorbeeld door deze leerlingen te vragen of zij andere oplossingsmethoden kennen, of door bij de begeleide oefening extra moeilijke opgaven op het bord te zetten voor deze leerlingen. Dit zijn voorbeelden van effectievere OGW-aanpakken. Daarnaast zijn er enkele leerkrachten die tijd nemen om de extra opgaven voor deze groep leerlingen kort met hen door te nemen, maar zelden wordt er ook echt een passende instructie aangeboden voor deze groep. Ook onderwijsinspecteurs constateren dat sterkere leerlingen vaak geen instructie aangeboden krijgen die passend is bij hun onderwijsbehoeften (Inspectie van het Onderwijs, 2014).

Het differentiëren met behulp van de drie instructiegroepen zou niet moeten leiden tot een te rigide lesorganisatie. Leerlingen die over het algemeen sterk zijn in rekenen kunnen soms ook baat hebben bij extra instructie over een specifiek leerstofonderdeel, en daarvoor zou ruimte moeten zijn binnen de lesplanning. Een gedeelte van de leerkrachten speelt hier wel op in door de leerlingen zelf te vragen wie nog extra uitleg wil over een bepaald onderdeel, deze leerlingen kunnen dan ook aanschuiven bij de instructietafel.

Samenvattend blijkt dat leerkrachten overwegend een positieve OGW-attitude hebben en dat zij, na de Focus-interventie, de benodigde kennis over het analyseren van leerresultaten met behulp van een LVS bezitten. De kwaliteit van de geplande instructie in het groepsplan is echter overwegend onvoldoende (gedifferentieerd), en de uitkomsten van leerresultatenanalyses met een LVS kunnen daarbij meer benut worden.

De uitvoering van OGW in de klas lijkt vaak nog te kort te schieten. Differentiëren is een complexe professionele vaardigheid (Inspectie van het Onderwijs, 2014), en er is onduidelijkheid over hoe 'goed differentiëren' eruit ziet en wanneer dit tot hogere leerresultaten leidt. De bevindingen uit dit onderzoek laten zien dat het maken van instructiegroepen alleen onvoldoende is, van belang lijkt een goede invulling van de inhoud van afgestemd onderwijs en daarbij het voorkomen van een te rigide aanpak in de lesorganisatie. Kennis over dergelijke zaken zijn belangrijk om ook een goede invulling te kunnen geven aan het trainen en ondersteunen van leerkrachten bij het differentiëren, zodat dergelijke investeringen ook daadwerkelijk gaan leiden tot verbeteringen voor leerlingen.

5.1.DISCUSSIE

In deze discussie worden beperkingen van het onderzoek besproken die bij het interpreteren en gebruiken van de resultaten in acht dienen te worden genomen.

Een eerste punt betreft de validiteit van het observatie-instrument voor het meten van differentiëren. Differentiëren is hier gemeten met drie items uit het observatie-instrument ICALT, het

is de vraag of hiermee voldoende de kwaliteit gemeten is van differentiëren. Daarnaast is het de vraag of lesopnames voldoende informatie geven voor het kunnen beoordelen van de kwaliteit van differentiëren. Aan de hand van lesobservaties kan goed beoordeeld worden of leerlingen in verschillende instructiegroepen geplaatst zijn, de kwaliteit van de instructie en de mate van differentiatie in verwerkingsopgaven tussen die instructiegroepen is daarentegen een stuk lastiger te beoordelen. Omdat het wellicht nodig is dat de observant daadwerkelijk in de klas is en daardoor werk van leerlingen kan bekijken, vragen kan stellen aan leerkrachten en/of leerlingen, zodat bepaald kan worden waarop instructiekeuzes zijn gebaseerd, en of leerlingen in de verschillende instructiegroepen ook echt verschillende vormen van instructie aangeboden krijgen. Hetzelfde geldt voor het ontwikkelde instrument waarmee de groepsplannen beoordeeld werden. Het instrument betrof een checklist waarmee beoordeeld werd of een item wel of niet opgenomen was in het groepsplan, en niet zozeer de kwaliteit daarvan ten aanzien van differentiatie.

Een tweede punt dat van belang is betreft weer het ontwikkelde instrument waarmee de groepsplannen zijn beoordeeld. Het door leerkrachten analyseren van resultaten uit de Cito LVS-toetsen en het toepassen van analyseresultaten vormde daarin een belangrijk onderdeel. Tijdens het beoordelen van groepsplannen bleek echter dat een aantal leerkrachten juist prioriteit legt bij resultaten uit methodegebonden toetsen en zij houden vaak een andere cyclus aan bij het evalueren en opstellen van groepsplannen. Dit kan nadelig zijn geweest voor deze groep bij het toekennen van scores. Gegevens uit de Cito LVS-toetsen geven toegang tot waardevolle prestatiefeedback mede doordat de scores van verschillende toets perioden op dezelfde onderliggende vaardigheidsschaal geplaatst worden, Cito biedt echter slechts tweemaal per jaar een ijkpunt en data uit methodetoetsen en andere instrumenten zijn van belang om tussen de Cito-toetsen het onderwijs af te kunnen stemmen op meest actuele leerbehoeften. Voor OGW is dus de feedback uit beide toetsvormen van belang, beide zijn daarom ook meegenomen in het instrument maar de focus lag daarbij meer op de Cito LVS-toetsen.

Tot slot zijn de gevonden resultaten verzameld binnen een groep scholen die allemaal een interventie gevolgd hebben die op OGW gericht was. Deze scholen waren daardoor bij voorbaat misschien al meer gemotiveerd voor OGW. Bovendien is het gemiddelde percentage leerlingengewichten op de scholen relatief hoog. Hoewel de resultaten hierdoor niet met zekerheid te generaliseren zijn, hebben we wel de indruk dat ze een goede indicatie geven van OGW binnen het primair onderwijs.

LITERATUURLIJST

- Blok, H., Ledoux, G., & Roeleveld, J. (2013). *Opbrengstgericht werken in het primair onderwijs: theorie en praktijk*. Amsterdam: Kohnstamm Instituut.
- Bosker, R., & Snijders, T. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, Thousand Oaks, New Delhi: SAGE Publications.
- Centraal Bureau voor Statistiek. (2014). *Onderwijs Cijfers*. Geraadpleegd op 9 februari 2015, van <http://www.cbs.nl/nl-NL/menu/themas/onderwijs/cijfers/default.htm>
- Doolaard, S. (2013). *Effecten van het trainings- en begeleidingstraject 'Streef: Gebruik maken van opbrengsten'*. Groningen: GION/RUG.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When Rater Reliability Is Not Enough: Teacher Observation Systems and a Case for the Generalizability Study. *Educational Researcher*, 41(2), 56-64.
- Inspectie van het Onderwijs. (2010). *Opbrengstgericht werken in het basisonderwijs; een onderzoek naar opbrengstgericht werken bij rekenen-wiskunde in het basisonderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2011). *De Staat van het Onderwijs. Onderwijsverslag 2009/2010*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2012). *De Staat van het Onderwijs. Onderwijsverslag 2010/2011*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2013). *De Staat van het Onderwijs. Onderwijsverslag 2011/2012*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2014). *De Staat van het Onderwijs. Onderwijsverslag 2013/2014*. Utrecht: Inspectie van het Onderwijs.
- Keuning, T., & Van Geel, M. J. M. (2012). *Focus projects II and III. The effects of a training in 'achievement oriented work' for primary school teams*. Poster presented at the International ICO fall school, Girona, Spain.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.
- Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. *American Psychologist*, 57(9), 705-717.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2007). *Scholen voor morgen: Samen op weg naar duurzame kwaliteit in het primair onderwijs*. Den Haag: Ministerie van OCW.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2011). *Actieplan 'Basis voor presteren'*. Den Haag: Ministerie van OCW.
- Smolkowski, K., & Gunn, B. (2012). Reliability and Validity of the Classroom Observations of Student-Teacher Interactions (COSTI) for Kindergarten Reading Instruction. *Early Childhood Research Quarterly*, 27(2), 316-328.
- Staman, L., Visscher, A.J., & Luyten, H. (2013). The effects of professional development on the attitudes, knowledge and skills for data-driven decision making. *Studies in Educational Evaluation* 42, 79-90.

- Tomlinson, C., Brighton, C., Hertberg, H., Callahan, C., Moon, T., Brimijoin, K., Conover, L., & Reynolds, T. (2004). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27(2/3), 119-145.
- Van de Grift, W. (2007). Quality of teaching in four European countries: a review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127-152.
- Van Geel, M., Keuning, T., Fox, J.P. & Visscher, A.J. *Assessing the Effects of a School Wide Data-Based Decision Making Intervention on Student Achievement Growth in Primary Schools* (submitted).
- Van Kuijk, M.F. (2014). *Raising the bar for reading comprehension: the effects of a teacher professional development program targeting goals, data use, and instruction* (Proefschrift). Opgevraagd van <http://dissertations.ub.rug.nl/faculties/gmw/2014/m.f.van.kuijk/>
- Visscher, A., & Ehren, M. (2011). *De eenvoud en complexiteit van Opbrengstgericht Werken*. Opgevraagd van <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2011/07/13/de-eenvoud-en-complexiteit-van-opbrengstgericht-werken.html>

BIJLAGE A: FOCUS-PROJECT

In de Focus-projecten worden scholen twee jaar lang getraind in OGW en het benutten van leerlingvolgsystemen, de training richt zich op alle leerjaren binnen de basisschool. In de eerste twee projecten was de training gericht op het vak rekenen-wiskunde, binnen het derde project konden scholen ervoor kiezen om zich te richten op een ander vakgebied. In het eerste jaar vonden zeven bijeenkomsten plaats (twee dagen en vijf dagdelen), in het tweede jaar vier en daarnaast klassenconsultaties en schoolbezoeken. De bijeenkomsten hadden een actief karakter. Nadat de leerkrachten, schoolleiders en intern begeleiders kennis hadden verworven gingen zij eigen data of leerresultaten gebruiken voor het maken van opdrachten, voeren van discussies en geven van presentaties.

In de bijeenkomsten van het eerste jaar stonden de volgende thema's centraal;

- mogelijkheden binnen leerlingvolgsystemen,
- rekendidactiek en doorgaande leerlijnen,
- goalsetting theorie en het maken van een groepsplan,
- aanscherpen en uitvoeren van het groepsplan,
- tussenevaluatie, analyse van Cito leerresultaten,
- aanscherpen en uitvoeren van het groepsplan (2),
- eindevaluatie en overdracht van gegevens naar de nieuw groep.

In de bijeenkomsten van het tweede jaar werden de volgende thema's behandeld;

- rekendidactiek,
- evaluatie Cito midden toetsen en verdieping rekenen,
- tussenevaluatie en reflectie,
- evaluatie Cito eind toetsen en afsluiting en borging.

BIJLAGE B: SCOREMODEL PROTOCOL EN GROEPSPLAN

	Omschrijving	Score	
		Nee	Ja
I	Kwaliteit analyses en interpretatie		
1	Zijn de vaardigheidsscores en vaardigheidsgroei aanwezig én wordt er iets over geschreven? (Het gaat hier niet over kwaliteit, máár om de aanwezigheid van de analyse) (N.v.t. voor groep 1; voor <u>groep 3</u> pas mogelijk na E-afname)		
a	De vaardigheidsscores en de vaardigheidsgroei zijn geanalyseerd op <u>groepsniveau</u> (het gaat hier om de hele groep in zijn totaliteit).	0	1
b	De vaardigheidsscores en de vaardigheidsgroei zijn geanalyseerd op <u>leerlingniveau</u> .	0	1
2	Zijn de vaardigheidsscores en vaardigheidsgroei op groepsniveau juist geanalyseerd én geïnterpreteerd? (Het gaat er hier om dat de analyses volledig én juist zijn.) (Voor <u>groep 1 en 3</u> pas mogelijk na E-afname)		
a	Er zijn juiste conclusies getrokken m.b.t. de gemiddelde vaardigheidsgroei van <u>de groep</u> (het gaat hier om de <u>hele</u> groep) in vergelijking met de gemiddelde vaardigheidsgroei van een groep met hetzelfde niveau. (Is de juiste tabel uit het "FOCUS handboek normgegevens" (versie FOCUS I) gebruikt?)	0	1
3	Zijn de vaardigheidsscores en vaardigheidsgroei op leerlingniveau juist geanalyseerd én geïnterpreteerd? (Het gaat er hier om dat de analyses volledig én juist zijn.) (Voor <u>groep 1 en 3</u> pas mogelijk na E-afname)		
a	Er zijn juiste conclusies getrokken m.b.t. de gemiddelde vaardigheidsgroei van <u>leerlingen</u> (Individueel en/of niveaugroepen) in vergelijking met de gemiddelde vaardigheidsgroei van leerlingen met hetzelfde niveau. (Is de juiste tabel uit het "FOCUS handboek normgegevens" (versie FOCUS I) gebruikt?)	0	1
4	Is de categorieënanalyse correct geïnterpreteerd? (Voor <u>groep 1 en 3</u> pas mogelijk na de M-afname)		
a	De getrokken conclusies m.b.t. de geconstateerde hiaten op <u>leerlingniveau</u> zijn juist (neem hier ook stap 7 en 8 uit het protocol mee). (Scoort een leerling bijvoorbeeld opvallend sterk/zwak op 1 leerstofonderdeel. Uit hoeveel opgaven bestaat de categorie? Scoort de leerling relatief sterk op de methodetoetsen en niet op de LVS-toetsen?, enz.)	0	1
b	De getrokken conclusies m.b.t. de geconstateerde hiaten op <u>groepsniveau</u> zijn juist. (Wat zijn de zwakke leerstofonderdelen voor deze groep: valt de meerderheid bijvoorbeeld uit op 1 onderdeel/categorie. Het is hier ook goed om na te gaan of het betreffende leerstofonderdeel aan de orde is geweest in de methode.)	0	1
c	Er is rekening gehouden met het niveau van de leerlingen. (Het niveau (A t/m E of I t/m V) van de leerling bepaald of een score op een bepaalde categorie opvallend is of niet. 50% goed van een bepaalde categorie heeft voor een A-II een andere waarde dan voor een D-II. Is hier rekening mee gehouden?)	0	1
d	Er is een diagnostisch gesprek gehouden om te achtergrond van de geconstateerde hiaten boven tafel is gehaald. (Stap 8 uit het protocol)	0	1
	TOTAAL		

II	Kwaliteit instructiebeslissingen;		
5	Sluit de groepsindeling bij de convergente differentiatie aan bij de beschikbare analyseresultaten?		
a	De leerlingen zijn ingedeeld in een instructiegroep passend bij hun niveau (In principe A lln instructie-onafhankelijk, B en C lln instructiegevoelig en D en E lln instructie-afhankelijk) en onderwijsbehoefte. (In een sterke groep kunnen A-lIn ook instructiegevoelig zijn. Hier zullen waarschijnlijk alleen de A+ lIn in de instructie onafhankelijke groep zitten. In een zwakke groep kunnen D-lIn ook instructiegevoelig zijn. Van leerlingen die in een "afwijkende" instructiegroep zijn geplaatst, is de onderwijsbehoefte beschreven.)	0	1
b	De instructiegevoelige groep is de grootste groep. (In principe is dit de groep waar je de basisinstructie op afstemt. Doel van de convergente differentiatie is om de groep zoveel mogelijk bij elkaar te houden. Wanneer er geen echte middengroep is kan hiervan afgeweken worden.)	0	1
6	Wordt bij de onderwijsbehoefte het leerkrachtgedrag in het groepsplan beschreven om alle leerlingen optimaal te laten functioneren?		
a	Voor enkele van de genoemde leerlingen wordt de pedagogische en/of didactische aanpak van de leerkracht beschreven.	0	1
b	Per leerling/groepje leerlingen staat de pedagogische en/of de didactische aanpak van de leerkracht beschreven.	0	1
c	Er wordt een korte beschrijving gegeven van het probleem(gedrag) van de leerling/het groepje leerlingen.	0	1
7	Wordt de instructie aangepast aan het niveau (op basis van de analyseresultaten) van <u>alle</u> leerlingen?		
a	De gedifferentieerde instructietijd wordt per instructiegroep beschreven. (Benoemd wordt in ieder geval dat er sprake is van verkorte, basis en verlengde instructie.)	0	1
b	Per instructiegroep wordt de organisatie beschreven.	0	1
c	Het didactisch handelen van de leerkracht staat per instructiegroep beschreven.	0	1
8	Wordt de verwerking aangepast aan het niveau van <u>alle</u> leerlingen?		
a	Er wordt gedifferentieerd in de omvang <u>of</u> aard van de opdrachten. (Bijv. de instructie afhankelijke groep krijgt een kleinere taak omdat zij verlengde instructie krijgen; compacting voor A+ lIn., pluswerkboek voor de sterke lIn.; bijwerkboek voor de zwakke lIn., enz.)	0	1
b	Er staat beschreven welke leerlingen in welke niveaugroep zitten. (De benaming is afhankelijk van de gebruikte rekenmethode. De Wereld in Getallen en Pluspunt gebruiken een sternotatie: * voor de zwakkere lIn; ** voor de gemiddelde lIn; *** voor de sterkere lIn.) (Dit is voor kleuters niet van toepassing.)	0	1
c	Per instructiegroep wordt gedifferentieerd in het gebruik van materialen die de reguliere rekenles ondersteunen. (Denk aan een rekenrek, een breukendoos, een tafelkaart enz.)	0	1
9	Worden de in te zetten materialen en hulpmiddelen specifiek beschreven?		
a	De in te zetten materialen, hulpmiddelen en methoden (naast de klassikale methode) worden benoemd.	0	1
b	Van de in te zetten materialen, hulpmiddelen én methoden (naast de klassikale methode) wordt beschreven <u>op welke manier</u> ze zullen worden ingezet.	0	1
c	Van de in te zetten materialen, hulpmiddelen én methoden (naast de klassikale methode) wordt specifiek beschreven <u>voor welke leerlingen</u> ze zullen worden ingezet.	0	1
TOTAAL			

III		Kwaliteit doelen; passen de gestelde doelen bij de geformuleerde onderwijsbehoeften	
10	Normgerichte doelen <u>methodetoetsen</u> : staat per instructiegroep beschreven welk percentage er minimaal goed gemaakt moet worden? (Alleen voor groep 3 t/m 8)		
a	Per instructiegroep staat beschreven welk percentage <u>minimaal</u> goed gescoord moet worden.	0	1
11	Zijn voor elk van de drie instructiegroepen, per niveau, normgerichte doelen (uitgedrukt in termen van vaardigheidsgroei) geformuleerd die aansluiten bij de analyseresultaten? (Dit is niet mogelijk voor groep 8)		
a	De doelen zijn in termen van vaardigheidsgroei beschreven. (Wanneer het niet realistisch is om groei te beschrijven (bijv. wanneer een II A+ scoort) dan moet er in ieder geval een minimaal te halen vaardigheidsscore genoemd worden.)	0	1
b	De doelen zijn per niveau geformuleerd. Bijv: A; GVG van 4,8	0	1
c	De gestelde normgerichte doelen zijn minimaal de berekende groei in het FOCUS handboek "normgegevens" (versie FOCUS I). Wanneer hiervan wordt afgeweken, alleen een punt toekennen als beschreven wordt waarom.	0	1
		TOTAAL	
IV		Evaluatie studievoortgang	
12	Is aangegeven <u>hoe en wanneer</u> resultaten (tussentijds) gemonitord worden?		
a	Volgende CITO-afname (Dit is niet mogelijk voor groep 8)	0	1
b	Methodetoetsen aan het eind van ieder blok (Dit is mogelijk voor groep 1 en 2)	0	1
c	Gemaakt werk, observaties, etc.	0	1
d	Er wordt expliciet benoemd <u>wanneer</u> de resultaten worden gemonitord.	0	1
13	Staat beschreven wat de consequenties zijn wanneer de gestelde doelen niet bereikt worden?		
a	Er staat beschreven wat de consequenties zijn wanneer niet aan de gestelde norm voldaan wordt.	0	1
		TOTAAL	

V	Subgroep specifieke items		
14	Sluit de groepsindeling van de subgroepen aan bij de beschikbare analyseresultaten?		
A	De groepsindeling is een logisch gevolg van de <u>LVS-toetsresultaten</u> .	0	1
B	Er wordt een link beschreven tussen de <u>LVS-toetsresultaten</u> en de resultaten op de <u>methodetoetsen/ ingezette observatie-instrumenten</u> .	0	1
C	De leerlingen in een subgroep hebben ongeveer hetzelfde niveau.	0	1
15	Sluit de probleemomschrijving/ diagnose aan bij de beschikbare analyseresultaten?		
A	Er is sprake van een duidelijke en kloppende probleemomschrijving/diagnose.	0	1
B	De in de probleemomschrijving beschreven hiaten zijn terug te lezen in de categorieënanalyse en/of sluiten aan op de in de methodetoetsen geconstateerde tekorten.	0	1
16	Sluiten de leerstofgerichte doelen aan bij de probleemomschrijving?		
A	De leerstofgerichte doelen hebben betrekking op hiaten in de leerstof zoals beschreven in de probleemomschrijving. (Het gaat hier niet om de kwaliteit van de doelen.)	0	1
B	Per subgroep staat één leerstofonderdeel centraal. (Het gaat hier om een korte beschrijving; opsomming.) Indien er meer leerstofonderdelen benoemd staan, alleen een punt toekennen als er samenhang bestaat tussen de verschillende onderdelen én deze keuze wordt verantwoord.	0	1
17	Zijn de leerstofgerichte doelen SMART geformuleerd?		
A	Specifiek; de doelstelling is eenduidig en gerelateerd aan te behandelen leerstofonderdeel.	0	1
B	Meetbaar; er staat duidelijk beschreven wanneer het doel is bereikt.	0	1
C	Tijdgebonden; er staat beschreven wanneer het doel bereikt moet zijn.	0	1
18	Zijn de normgerichte doelen SMART geformuleerd?		
A	Specifiek; de doelstelling is eenduidig en gerelateerd aan te behandelen leerstofonderdeel.	0	1
B	Meetbaar; er staat duidelijk beschreven wanneer het doel is bereikt.	0	1
C	Tijdgebonden; er staat beschreven wanneer het doel bereikt moet zijn.	0	1
19	Staat de rol van de leerkracht specifiek beschreven?		
A	De didactische aanpak van de leerkracht staat beschreven (Hoe gaat de leerkracht de leerstof aanbieden/aanleren)	0	1
20	Staan de aan te leren oplossingsstrategieën beschreven?		
A	De aan te leren oplossingsstrategieën staan beschreven.	0	1
21	Wordt inzet van materialen en hulpmiddelen <u>specifiek</u> beschreven		
A	De inzet van materialen en hulpmiddelen wordt specifiek beschreven; op welke manier en voor wie worden de benoemde materialen/hulpmiddelen ingezet.	0	1
21	Wordt de <i>ingeplande tijd</i> voor extra begeleiding (subgroepen) specifiek beschreven?		
A	De frequentie van de extra begeleiding staat benoemd.	0	1
b	De dag waarop de subgroepen zijn gepland staat beschreven.	0	1
c	Het tijdstip waarop de subgroep wordt begeleid staat beschreven.	0	1
d	De duur van de begeleiding staat beschreven.	0	1
TOTAAL			

BIJLAGE C: FORMAT GROEPSPLAN FOCUS

Groepsplan

Groep:
Leerkracht(en):
Datum:
Periode:
Normgericht GROEPSdoel: Deze groep scoort minimaal een ___ > vaardigheidsscore: ____
Percentageverdeling niveaus: ___% AB scores – ___% ABC scores
Welke nieuwe, extra maatregelen neem je n.a.v. de informatie in deze cursus over rekensterke scholen: <input type="checkbox"/> Meer rekentijd in het rooster. Zo ja, hoeveel extra tijd per week: ____ min <input type="checkbox"/> Maatregelen om 'weglektijd' te verminderen. Zo ja, welke: <input type="checkbox"/> Meer tijd voor automatiseren. Zo ja, hoeveel minuten per week: <input type="checkbox"/> Minder werken met individuele leerlijnen. Zo ja, hoe: <input type="checkbox"/> Maatregelen om vroeg risicoleerlingen in groep 1 en 2 te signaleren en helpen. Zo ja, hoe:

Evaluatie:				
Wanneer ga je de vorderingen van je leerlingen op basis van dit plan (tussentijds) evalueren?	▪			
Welke evaluatie-instrumenten ga je hiervoor gebruiken?	▪			
Wanneer ga je de resultaten van dit plan bespreken met:	IB-er:	Schoolleider:	Collega('s):	Ouders:

Convergente differentiatie

	Instructie onafhankelijke \mathbb{N} :		Instructie gevoelige \mathbb{N} :		Instructie afhankelijke \mathbb{N} :	
	[naam]	[niveau + v-score]	[naam]	[niveau + v-score]	[naam]	[niveau + v-score]
Groepssamenstelling (namen met daarachter behaalde vaardigheidsscore en niveau op vorige Cito-afname)						
Onderwijsbehoefte Opmerkingen m.b.t. de onderwijsbehoeften van (individuele) leerlingen						

Normgerichte GROEPSdoelen	Methode	___ % goed op methodetoets	___ % goed op methodetoets	___ % goed op methodetoets
	Cito	___ vaardigheidspunten op M6	___ vaardigheidspunten op M6	___ vaardigheidspunten op M6
Werkwijze:				
Aanvulling methode: <i>Wat doe je naast de methode met deze leerlingen voor het bereiken van de leerstofdoelen uit de methode?</i> <ul style="list-style-type: none"> • Welke materiaal gebruik je als ondersteuning? • Is er een speciale didactische aanpak vereist? Noot: <i>*nieuwe methodes geven veel handvatten voor gedifferentieerde inhoud en het vormgeven van instructiegroepen. Wanneer dat voor u als leerkracht niet het geval is zult u bewuste keuzes moeten maken in het vormgeven van uw convergente differentiatie.</i>		▪	▪	▪

Clusteren van leerlingen met specifieke onderwijsbehoeften

	Subgroep 1 D en E leerlingen	Subgroep 2 Meten	Subgroep 3 Tijd
Samenstelling subgroepjes (namen)			
Probleemomschrijving: Beschrijf kort en bondig het probleem of de beginsituatie.			
Aanvullende leerstofgerichte doelen voor volgende Cito-afname	▪ Aan het einde van deze periode kunnen of kennen de leerlingen:	▪ Aan het einde van deze periode kunnen of kennen de leerlingen:	▪ Aan het einde van deze periode kunnen of kennen de leerlingen:
Aanvullende normgerichte doelen voor volgende Cito-afname; bv - Formuleer een norm voor het % goed op de categorie? - Of, filter opgaven uit de Cito-toets die passen bij je leerstofgerichte doel en formuleer een norm.			

Tussenevaluatie - <i>Wanneer?</i> : Plan in het komende halfjaar een moment voor een tussenevaluatie waarin de aanvullende doelen worden geëvalueerd;				
- <i>Hoe?</i> Kies een vorm van evalueren om de mate van beheersing te toetsen.		■	■	■
- <i>Vervolg</i> : eventueel bijstellen.		■		
Aanpak	Welke leerstof bied je aan? (beschrijf de leerstof concreet)			
	Hoe bied je de leerstof aan? Welke specifieke (pedagogische of didactische) begeleiding is nodig? (bijv. welke strategie, materialen)	<i>Aanpak:</i> <i>Materiaal:</i> <i>Begeleiding:</i>	<i>Aanpak:</i> <i>Materiaal:</i> <i>Begeleiding:</i>	<i>Aanpak:</i> <i>Materiaal:</i> <i>Begeleiding:</i>
	Wanneer is er extra tijd of verlengde instructie gepland?			

* Het aantal subgroepen is afhankelijk van wat de klassenorganisatie toelaat.