

# Hangen cijferprestaties van leerlingen in het basisonderwijs samen met het gebruikte wiskundehandboek? Een *mixed methods* analyse.

Goffin, E., Van Dooren, W., Avau, I., Thomassen, I., Ameel, E. en Janssen, R.

## Samenvatting

Presteren leerlingen beter of net minder goed in functie van het handboek dat in de klas gebruikt wordt? Aan de hand van een verkennende *mixed methods* analyse tonen we voor cijferen hoe kan worden nagegaan of het een verschil maakt welk wiskundehandboek er gebruikt wordt in de Vlaamse basisschool. In het eerste, kwantitatieve luik van dit onderzoek voeren wij een twe-niveauregressie-analyse uit op de dataset van de tweede Vlaamse peiling wiskunde basisonderwijs van 2009. We tonen aan dat het handboek voor bepaalde deeldomeinen een samenhang vertoont met de prestaties van de leerlingen. In een tweede luik maken wij door middel van doelgerichte selectie op basis van de resultaten van de eerste studie een kwalitatieve gevalsstudie van twee veelgebruikte handboeken die specifiek voor cijferen sterk uiteenlopende peilingsresultaten lieten optekenen. We stellen vast dat de geselecteerde handboeken voor dit deeldomein ook inhoudelijk en structureel van elkaar verschillen. Zo besteedt het handboek dat het beste scoort bijvoorbeeld meer tijd aan cijferen, biedt het meer cijferoefeningen aan en legt het enigszins andere didactische accenten in de instructie van cijferen. Mogelijk kunnen deze verschilpunten de prestatieverschillen (mee) verklaren die we in de kwantitatieve analyse vonden.

**Kernwoorden:** handboekonderzoek, *mixed methods*, doelgerichte selectie, peilingsonderzoek, wiskundehandboek, basisonderwijs, cijferen

## 1 Inleiding

### 1.1 Handboek als potentieel geïmplementeerd curriculum

In onderzoek naar de effecten van wiskundeonderwijs op de wiskundeprestaties van leer-

lingen bestudeert men veelal de impact van het aantal en de types wiskundelessen, maar wordt slechts zelden specifiek gefocust op de gebruikte handboeken. Handboeken bekleden nochtans een belangrijke plaats in het onderwijsleerproces. In het driedelig model van bedoeld (*intended*), geïmplementeerd (*implemented*) en verworven (*attained*) curriculum situeert de TIMSS-groep handboeken op een apart niveau tussen de eerste en de tweede pijler. Dit niveau wordt omschreven als het *potentially implemented* of potentieel geïmplementeerd curriculum (Valverde, Bianchi, Wolfe, Schmidt, & Houang, 2002): in handboeken worden de intenties van degenen die het curriculum bepalen en vormgeven, vertaald naar leerkrachten en leerlingen toe. In die zin zijn zij een mediërende factor tussen onderwijsbeleid en klaspraktijk. Handboeken zijn met andere woorden een wezenlijk onderdeel van de zogenaamde *implementatietrap* (Vandenbergh, 2004). In Vlaanderen vertrekt men vanuit eindtermen die door de onderwijsoverheid geformuleerd worden, die vervolgens gevat worden in leerplannen en van daaruit een vertaling vinden in handboeken. De leerkracht gebruikt dit materiaal als leidraad bij het interpreteren en implementeren van de aangeboden leerinhouden, met als eindpunt de uiteindelijke leerlingprestaties. Törnroos (2005) stelt dat handboeken in hun hoedanigheid van potentieel geïmplementeerd curriculum een goede maatstaf zijn voor de geboden *opportunity to learn*, en dat een grotere *opportunity to learn* betere leerlingprestaties met zich meebrengt.

### 1.2 Handboekonderzoek

In een literatuurstudie over reeds gevoerd handboekonderzoek in het wiskundeonderwijs, onderscheiden Fan, Zhu en Miao (2013) vier categorieën. Studies over de *conceptuele rol* van het handboek in onderwijs en leren, zoals bijvoorbeeld het eerder vermelde werk van de TIMSS-groep, vormen een eerste cate-

gorie. Dit soort studies zijn eerder gering in aantal, al komen de concepten die erin naar voren gebracht worden uiteraard ook aan bod in verder empirisch onderzoek. Een tweede categorie omvat *handboekanalyses* en comparatief onderzoek waarin het handboek als *product* centraal staat. Deze categorie is duidelijk de omvangrijkste in het veld maar ze is ook erg divers. Men focust hier bijvoorbeeld op de inhoud en structuur van handboeken, op de leer- en oplossingsstrategieën die erin aangeboden worden (zie bijvoorbeeld van Zanten & van den Heuvel-Panhuizen, 2014), of op socioculturele aspecten en internationale verschillen (zie bijvoorbeeld Pingel, 2010). In een derde categorie vermelden de onderzoekers een aantal veelal kleinschalige studies over het *didactisch gebruik* van handboeken in de klas. Ze zijn vaak voornamelijk gericht op het handboekgebruik door leerkrachten en niet zozeer op het gebruik ervan door leerlingen of studenten zelf (Fan et al., 2013). Gericht onderzoek naar de *samenhang tussen het gebruikte handboek en wiskunde-prestaties* bevindt zich in de typologie van Fan e.a. (2013) in een vierde en kleine 'rest-categorie' van voornamelijk verkennend empirisch onderzoek. Er is tot nog toe weinig eensgezindheid over het voorkomen van zulke samenhang. De auteurs pleiten daarom expliciet voor de ontwikkeling van een meer gesofisticeerde en gestructureerde methodologie.

### 1.3 Onderzoeksvragen en onderzoeksstrategie

Voorbeelden van onderwijseffectiviteitsonderzoek waarin het handboek als variabele bestudeerd wordt, zijn schaars (Fan et al., 2013). Voor handboekonderzoek in Vlaanderen vermelden wij wel een recente studie van Van Steenbrugge, Valcke en Desoete (2013). In deze studie wordt in de eerste plaats onderzocht welke perceptie leerkrachten hebben van wiskundehandboeken, maar in de tweede plaats ook of er een samenhang gevonden kan worden tussen het gebruikte handboek en de prestaties van de leerlingen. De auteurs stellen vast dat leerkrachten aan verschillende handboeken duidelijk een verschillende doeltreffendheid en aantrekkelijkheid toedichten. Hun analyse van de wiskundepresta-

ties van de leerlingen, gebaseerd op een beperkte deelsteekproef, levert echter geen significante samenhang met het handboek op. Met voorliggend onderzoeksverslag willen wij een andere mogelijke aanpak verkennen. We rapporteren over een *mixed methods* analyse waarin we op basis van grootschalige data eerst een mogelijke samenhang onderzoeken tussen de handboeken en de leerlingprestaties, en deze samenhang vervolgens exemplarisch trachten te duiden door middel van een gerichte handboekanalyse voor één deeldomein.

In een eerste fase voeren wij een kwantitatieve analyse uit op data uit de tweede Vlaamse peiling wiskunde aan het einde van het basisonderwijs van 2009. Onze twee onderzoeksvragen daarbij luiden:

1. Is er een samenhang tussen het gebruikte handboek en het bereiken van de eindtermen wiskunde basisonderwijs?
2. Hoe groot is het effect van het gebruikte handboek op de leerlingprestaties?

We nemen verschillende domeinen en deeldomeinen in beschouwing. We kijken niet alleen naar ruwe prestatieverschillen, maar corrigeren ook voor een aantal achtergrondkenmerken van leerlingen en scholen die doorgaans in effectiviteitsonderzoek in het zogenaamde nettomodel worden opgenomen.

In een tweede fase gaan we over tot een kwalitatieve verdieping. Daarvoor maken we op basis van resultaten uit het eerste luik een doelgerichte selectie (*purposive* of *purposeful sampling*, Patton, 2015) van zowel het deeldomein als de wiskundehandboeken die we aan verder onderzoek zullen onderwerpen. Hoewel we voor meerdere deeldomeinen een duidelijke samenhang vonden tussen de leerlingprestaties en het gebruikte handboek, kiezen we ervoor om ons toe te spitsen op *Cijferen*. Dit is één van deeldomeinen waarvoor de verschillen tussen handboeken vrij groot bleken te zijn, maar het is ook een welomlijnd leerstofonderdeel waarvan we verwachten dat het zich ondubbelzinnig zal manifesteren in het te onderzoeken materiaal. We voeren een beschrijvende gevalsstudie uit van twee handboeken die frequent gebruikt worden in de Vlaamse basisscholen

en die in de peilingsdata duidelijk uiteenlopende prestaties lieten optekenen voor cijferen. We gaan na of er bepaalde handboekkenmerken zijn die mogelijk verband zouden kunnen houden met de prestatieverschillen die we in de eerste studie hebben gevonden. Onze onderzoeksvraag in deze fase luidt:

3. Verschillen de geselecteerde handboeken structureel en inhoudelijk van elkaar wat cijferen betreft?

Het opzet van ons *mixed methods* onderzoek laat uiteraard niet toe om onze conclusies te generaliseren. We belichten immers maar één deeldomein en voeren onze analyse uit op een niet-willekeurige steekproef van handboeken. Onze assumptie is echter dat contrasterende *cases* het rijkst aan informatie zijn en een vergelijking ervan de meest sailante verschilpunten aan het licht kan brengen. Deze verschilpunten brengen op hun beurt wellicht hypotheses aan die nader onderzocht kunnen worden. We sluiten af met een algemene conclusie en aanbevelingen voor verder onderzoek.

## 2 Kwantitatieve analyse van het handboekeffect in de tweede peiling wiskunde basisonderwijs (2009)

### 2.1 Inleiding

#### *Algemene situering peilingsonderzoek*

Sinds 1997 zijn er in Vlaanderen in het basisonderwijs secundair onderwijs eindtermen van kracht, die gelden als minimumdoelen die de overheid noodzakelijk en bereikbaar acht voor een bepaalde leerlingenpopulatie. In opdracht van de overheid worden sinds 2002 grootschalige peilingen uitgevoerd bij representatieve steekproeven van Vlaamse leerlingen en scholen, om op systeemniveau na te gaan of de beoogde basiscompetenties gerealiseerd worden voor bepaalde deelaspecten en leergebieden. Bovendien wordt onderzocht of er prestatieverschillen zijn tussen scholen en waarmee deze verschillen samenhangen (Gielen et al., 2010).

In een peilingsonderzoek worden de te onderzoeken eindtermen in inhoudelijke analyse gebundeld en vertaald in concrete toetsopgaven. Met behulp van modellen uit de

itemresponstheorie (IRT) worden zowel de toetsopgaven als de vaardigheidsscores van de leerlingen op één meetschaal weergegeven, om de leerlingen onderling, alsook de leerlingen en de toetsopgaven, met elkaar te kunnen vergelijken. Per (deel)toets wordt er een toetsnorm vastgelegd door een groep van onderwijsdeskundigen.

Om variantie in prestaties (vaardigheidsscores) op leerling-, klas- en schoolniveau te onderzoeken wordt gebruikgemaakt van multiniveaumodellen. Er wordt gestart met een leeg model waarin geen rekening gehouden wordt met achtergrondkenmerken. In een tweede stap wordt een vaste set van context- en inputvariabelen toegevoegd om tot een nettomodel te komen, een basismodel dat een faire vergelijking van de schoolresultaten mogelijk moet maken door rekening te houden met factoren die buiten de controle van de school liggen (Gielen et al., 2010). Vervolgens wordt er onderzocht of bepaalde bijkomende school-, klas- en leerlingvariabelen een significant effect hebben op de prestaties. Zowel het enkelvoudig effect van deze variabelen, als hun effect na controle voor variabelen uit het nettomodel worden bestudeerd.

#### *Peilingsonderzoek wiskunde basisonderwijs*

In 2002 werd het realiseren van de eindtermen wiskunde aan het einde van het Vlaamse basisonderwijs een eerste maal gepeild. Een tweede peiling wiskunde basisonderwijs vond plaats in 2009. Op beide meetmomenten waren de grootste prestatieverschillen op leerlingniveau gesitueerd (Luyten et al., 2003; Gielen et al., 2010). De variantie tussen leerlingen bedroeg voor de verschillende domeinen ongeveer 80 procent, de schoolvariantie ongeveer 20 procent. Slechts een klein resterend deel van de variantie bevond zich op klasniveau, een gangbaar gegeven in het peilingsonderzoek in het Vlaamse basisonderwijs (zie bijvoorbeeld ook Vlaamse overheid, Agentschap voor Kwaliteitszorg in Onderwijs en Vorming, 2011). Een grote meerderheid van de scholen heeft ook maar één klas in het zesde leerjaar. In de praktijk vallen school- en klasvariantie bijgevolg vaak samen. Wanneer er toch meerdere klassen zijn in een school, zijn die klassen vaak hete-

rogeen samengesteld, waardoor geen grote klasverschillen binnen een school te verwachten zijn.

Op basis van de dataset van de herhalingspeiling van 2009 gaan wij in het eerste, kwantitatieve luik van deze *mixed methods* analyse na of het gebruikte handboek samenhangt met het bereiken van de eindtermen wiskunde basisonderwijs, en zo ja, in welke mate. Onze hypothese is dat er een effect van handboek op de leerlingprestaties meetbaar zal zijn, maar dat dit eerder beperkt zal zijn. De variabele *handboek* situeert zich immers wellicht op het niveau van de school, aangezien in scholen meestal eenzelfde handboek gebruikt wordt in de verschillende klassen en over de verschillende leerjaren heen, en in de beide wiskundepeilingen werd op schoolniveau slechts een kleine hoeveelheid variantie gerapporteerd.

## 2.2 Methode

### *Deelnemers*

In de peiling van 2009 werden er toetsen wiskunde basisonderwijs afgenomen bij een representatieve steekproef van 6940 leerlingen uit 389 klassen, verdeeld over 274 vestigingsplaatsen van 247 scholen (Gielen et al., 2010). De onderzoekers legden daarnaast ook uitgebreide achtergrondvragenlijsten voor aan de deelnemende leerlingen en hun ouders, aan de leerkrachten van de deelnemende klassen, en aan de directies van de deelnemende scholen. Voor elke deelnemende klas was er een aparte leerkrachtvragenlijst. Van de 400 leerkrachtvragenlijsten die verstuurd werden, werden er 384 (of 96%) volledig of gedeeltelijk ingevuld teruggestuurd naar de onderzoekers. Tien niet-ingevulde vragenlijsten behoorden toe aan klassen die zich afmeldden kort voor de peiling (Gielen et al., 2010).

### *Instrumenten*

In totaal werden 55 eindtermen wiskunde voor het basisonderwijs in 22 toetsen gevat, verdeeld over drie domeinen: *Getallen en bewerkingen*, *Meten en meetkunde* en *Strategieën en probleemoplossende vaardigheden* (Gielen et al., 2010). Met deze toetsen werden negen verschillende toetsboekjes samen-

gesteld. De namen van de verschillende individuele toetsen staan vermeld in Tabel 1 waarin de resultaten van de multiniveauanalyses voorgesteld worden per meetschaal. In de peiling werden de verschillende bewerkingen onder *Snelrekenen* (schaal 1b) samen als één toets beschouwd. Toets 11 uit de peiling, *Zakrekenmachine gebruiken*, is niet in Tabel 1 opgenomen. Voor deze toets werd geen meetschaal opgesteld door de peilingsonderzoekers omdat het type items zich hier niet toe leende (Gielen et al., 2010).

In de achtergrondvragenlijst voor de leerkrachten werd onder andere gevraagd naar het gebruikte lesmateriaal. De leerkrachten konden één of meerdere opties aanduiden in een keuzelijst met de vijf meest gebruikte wiskundehandboeken (namelijk, *Zo gezegd, zo gerekend!* [Wolters Plantyn], *Eurobasis* [die Keure], *Pluspunt* [Van In], *Nieuwe Talrijk* [Wolters Plantyn] en *Rekensprong* [Van In]) en een open optie 'Andere'. Daarnaast werd met een ja-nee-vraag en een mogelijkheid tot toelichting ook gevraagd of men naast het handboek eigen lesmateriaal gebruikte in de lessen wiskunde.

### *Data-analyse*

Voor het analyseren van de dataset gebruikten we SAS versie 9.2. Eerst maakten we een korte descriptieve analyse om het handboekgebruik in kaart te brengen. Vervolgens voerden we een tweenniveau-regressieanalyse uit op de data om te zien of het gebruikte wiskundehandboek een effect heeft op het behalen van de eindtermen wiskunde. We clusterden leerlingen daarbij binnen scholen. In het peilingsonderzoek worden verschillende vestigingsplaatsen van één school (instelling) als aparte school beschouwd.

Voor elke meetschaal of toets uit het peilingsonderzoek, inclusief de drie overkoepelende schalen, werd in een leeg model eerst de verdeling van de variantie van prestaties op school- en leerlingniveau in kaart gebracht. Vervolgens stelden we drie modellen op. In Model 1 werd nagegaan hoeveel variantie verklaard kan worden door het gebruikte handboek als onafhankelijke variabele toe te voegen (enkelvoudig effect), en in Model 2 werd dezelfde analyse nogmaals uitgevoerd

maar deze keer gecorrigeerd voor achtergrondkenmerken van scholen en leerlingen. In Model 3, ten slotte, werd het effect van de achtergrondkenmerken op zich berekend. De achtergrondkenmerken die als controlevariabelen in de modellen zijn opgenomen, zijn op leerlingniveau het geslacht, de leeftijd, de thuistaal, het aantal boeken thuis, de SES en eventuele leerproblemen. Op schoolniveau gaat het om de stratificatievariabelen schoolgrootte, onderwijsnet en provincie en de concentratiegraad (het aandeel leerlingen met een sociaal-economische achterstand). In deze multiniveau-analyse werd het effect van de handboeken met dummyvariabelen in kaart gebracht, met *Zo gezegd, zo gerekend!* als referentiehandboek.

Het berekenen van het percentage verklaarde variantie gebeurde aan de hand van de volgende formule:

$$1 - \frac{\text{schoolvariantie (Model 1)}}{\text{schoolvariantie (Leeg model)}}$$

In Model 1 werden in enkele gevallen ook negatieve percentages gevonden. Dit is een consequentie van het feit dat we niet van alle scholen het handboek kenden. Bijgevolg werden de twee modellen niet op basis van exact dezelfde gegevens berekend.

Tot slot vergeleken we per meetschaal de effecten van de verschillende handboeken met het effect van één referentiehandboek, *Zo*

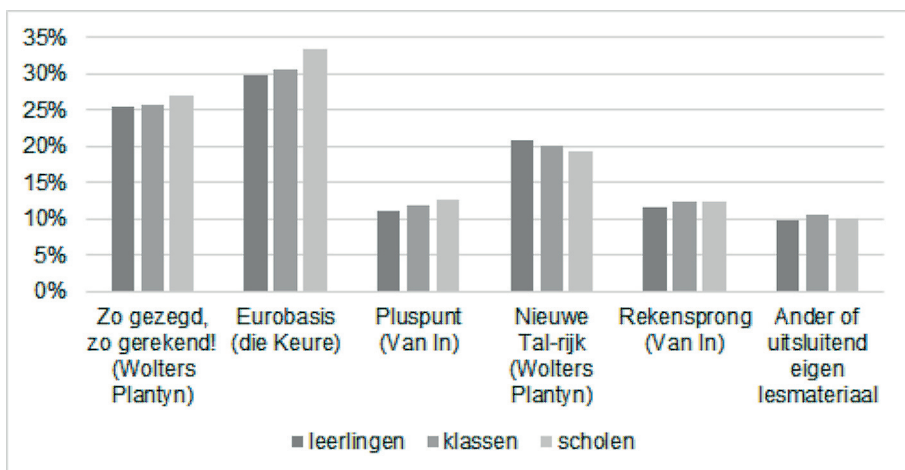
*gezegd, zo gerekend!*, door de gevonden regressiegewichten van de corresponderende dummyvariabelen met elkaar te vergelijken. De restcategorie van ‘Andere’ handboeken werd weggelaten. Deze categorie is immers weinig informatief omdat zij niet toelaat uitspraken te doen over concrete handboeken. De subschalen van schaal 1b (*Snelrekenen*) namen we niet op in de vergelijking.

### 2.3 Resultaten

#### *Beschrijvende analyse*

Uit de respons van de leerkrachten blijkt dat er meestal één wiskundehandboek gebruikt werd in de klas. In ongeveer 11 procent van de deelnemende klassen werden twee handboeken gebruikt; een combinatie van meer dan twee handboeken werd niet gerapporteerd. In de categorie ‘Andere’ vermeldden respondenten onder andere verouderde versies van de handboeken uit de keuzelijst, en een enkele respondent gaf aan uitsluitend eigen lesmateriaal te gebruiken. In totaal gaf ongeveer twee derde van de leerkrachten aan dat zij naast het handboek nog eigen lesmateriaal gebruiken, zoals bijvoorbeeld zelfgemaakte werkblaadjes.

Binnen één school werd in alle deelnemende klassen meestal hetzelfde handboek of dezelfde handboekcombinatie gebruikt. In ongeveer zeven procent van de scholen verschilde dit over de deelnemende klassen heen, of kunnen we op basis van onvolledige gegevens niet met zekerheid zeggen of het



Figuur 1. Verspreiding van de handboeken in de peilingssteekproef.

Tabel 1

Percentage (verklaarde) variantie op school- en leerlingniveau voor de verschillende meetniveaus en modellen

Schaal	Leeg model			Model 1			Model 2			Model 3		
	N <sup>a</sup>	S <sup>b</sup>	L <sup>c</sup>	N <sup>a</sup>	S <sup>d</sup>	L <sup>e</sup>	N <sup>a</sup>	S <sup>d</sup>	L <sup>e</sup>	N <sup>a</sup>	S <sup>d</sup>	L <sup>e</sup>
1: Hoofdrekenen	1521	9	91	1504	0	0	1401	53	13	1417	47	13
1b: Snelrekenen +	2297	8	92	2268	-3	0	2087	30	4	2113	28	5
1b: Snelrekenen -	2297	8	92	2268	9	0	2087	39	4	2113	30	4
1b: Snelrekenen x	2297	15	85	2268	4	0	2087	24	7	2113	22	7
1b: Snelrekenen :	2297	16	84	2268	6	0	2087	23	8	2113	21	8
2: Functies en voorstellingswijzen	1502	8	92	1484	-3	0	1376	61	8	1393	63	8
3: Begrippen en symbolen m.b.t. bewerkingen	1491	10	90	1474	-1	0	1379	60	10	1396	58	10
4: Breuken begrijpen en de terminologie hanteren	1490	19	81	1473	0	0	1378	71	21	1395	67	21
5: Natuurlijke getallen, breuken, kommagetallen en procenten	1491	17	83	1474	-3	0	1379	68	17	1396	65	17
6: Veelvouden en delers vinden en de deelbaarheid toepassen	1521	14	86	1504	-1	0	1401	55	17	1417	51	17
7: Afronden, benaderen en schatten	1549	17	83	1502	7	0	1409	56	17	1452	51	17
8: Verhoudingen en schaal vaststellen, vergelijken, berekenen	1502	17	83	1484	-2	0	1376	73	20	1393	74	20
9: Cijfers met natuurlijke getallen en kommagetallen	1521	12	88	1504	15	0	1401	38	13	1417	16	13
10: Rekenen met procenten	1501	17	83	1483	-2	0	1375	60	18	1392	62	18
12gb: Problemen oplossen bij getallen en bewerkingen	1555	18	82	1506	16	0	1424	77	16	1468	68	16
12mm: Problemen oplossen bij meten en meetkunde	1560	19	81	1510	27	0	1431	72	13	1478	61	13
13: Maateenheden	1532	20	80	1514	18	0	1385	83	15	1402	80	15
14: Betekenisvolle herleidingen met grootheden, maatgetallen en maateenheden	1532	22	78	1514	10	0	1385	55	15	1402	58	15
15a: Maten gebruiken in betekenisvolle situaties	1532	15	85	1514	9	0	1385	70	10	1402	72	9
15b: Schatten met behulp van referentiepunten	1549	18	82	1502	19	0	1409	80	13	1452	72	14
16: Omtrek, oppervlakte en inhoud berekenen	1532	19	81	1511	11	0	1370	61	13	1388	59	12
17: Rekenen met geld en klokken	1532	11	89	1513	8	0	1419	44	11	1436	43	11
18: Ruimte en ruimtelijke oriëntatie	1532	14	86	1513	8	0	1419	48	13	1436	50	13
19: Herkennen, benoemen, classificeren en vorm geven van meetkundige objecten en inzicht in meetkundige relaties	1532	17	83	1511	19	0	1370	61	12	1388	58	12
Getallen en Bewerkingen	2256	19	81	2231	-1	0	2078	67	23	2103	65	23
Meten en Meetkunde	2297	23	77	2268	12	0	2087	62	19	2113	63	19
Strategieën	2330	22	78	2259	14	0	2132	73	18	2199	68	19

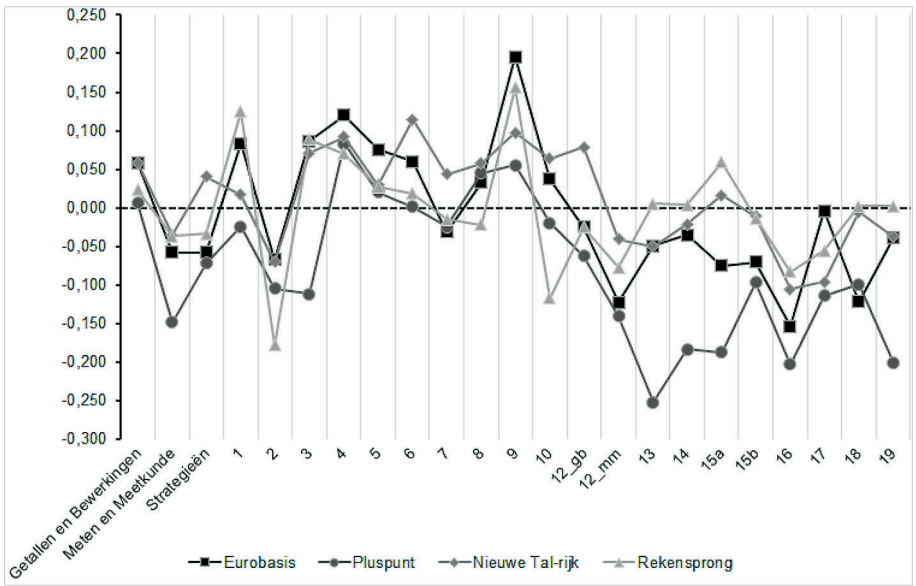
<sup>a</sup> Aantal leerlingen waarop het percentage (verklaarde) variantie berekend kon worden.

<sup>b</sup> Variantie op schoolniveau.

<sup>c</sup> Variantie op leerlingniveau.

<sup>d</sup> Verklaarde variantie op schoolniveau.

<sup>e</sup> Verklaarde variantie op leerlingniveau.



Figuur 2. Vergelijking van regressiegewichten voor het handboekeffect in Model 1 (Zo gezegd, zo gerekend! als referentiehandboek).

gebruikte handboek in de deelnemende klassen hetzelfde was. In meer dan twee derde van de scholen (64 procent) nam slechts één klas deel.

In Figuur 1 wordt de verspreiding van de verschillende handboeken weergegeven over de deelnemende leerlingen, klassen en scholen. Wanneer de weergegeven percentages per niveau opgeteld worden, is het totaal groter dan 100 procent omdat in een klas meerdere handboeken konden voorkomen. Aangezien in de leerkrachtvragenlijst niet naar de hiërarchie tussen de gebruikte handboeken werd gepeild, werd in dat laatste geval elk aangeduid handboek volwaardig meegerekend. De verspreiding die getoond wordt in Figuur 1 is vergelijkbaar over de verschillende niveaus, hetgeen betekent dat in de peilingssteekproef gemiddeld evenveel leerlingen in een klas zaten over de verschillende scholen heen, en dat ook het aantal deelnemende klassen per school grotendeels overeenkwam. De handboeken *Eurobasis*, *Zo gezegd, zo gerekend!* en *Nieuwe Tal-rijk* werden het meest gebruikt: deze drie blijken samen goed voor meer dan 75 procent van het handboekgebruik voor wiskunde in het Vlaamse basisonderwijs in 2009.

#### Multiniveau-analyse

Tabel 1 geeft de resultaten weer van de twee-niveau-regressieanalyse. We zien dat de variantie in het lege model grotendeels op het niveau van de leerlingen ligt. Globaal genomen bedraagt de schoolvariantie ongeveer 20 procent en de leerlingvariantie ongeveer 80 procent. Voor sommige schalen bedraagt de verhouding 10 procent versus 90 procent.

Wanneer we het handboek toevoegen in Model 1, wordt voor de meeste toetsen een klein percentage van de gevonden schoolvariantie verklaard. Voor zes toetsen zien we echter dat het handboek een relatief groot percentage van de gemeten prestatieverschillen tussen scholen verklaart, namelijk meer dan 15 procent: *Cijferen* (schaal 9), de toetsen rond *Probleemoplossende vaardigheden* (schaal 12gb en schaal 12mm), *Maateenheden* (schaal 13), *Schatten met behulp van referentiepunten* (schaal 15b) en *Inzicht in meetkundige relaties* (schaal 19).

We zien in Tabel 1 dat het percentage school- en leerlingvariantie dat verklaard kan worden, toeneemt wanneer de achtergrondvariabelen aan het model worden toegevoegd in Model 2 en 3. De combinatie van het gebruikte wiskundehandboek en de achtergrondkenmerken verklaart dus een groot deel van de

prestatieverschillen en het effect van handboek voor bepaalde toetsen blijft overeind na controle voor het nettomodel.

In Figuur 2 worden de regressiegewichten voor de enkelvoudige effecten (Model 1) van de vijf meest gebruikte handboeken voorgesteld voor de verschillende toetsen en overkoepelende meetschalen. Het effect van referentiehandboek *Zo gezegd, zo gerekend!* wordt als nullijn weergegeven. In de figuur merken we vooral sterke schommelingen op per schaal: de handboeken laten erg wisselende prestaties optekenen over de verschillende toetsen heen. Over het algemeen scoort *Pluspunt* het laagst.

#### 2.4 Discussie

Op basis van een tweenniveau-regressieanalyse van de dataset uit de peiling wiskunde basisonderwijs van 2009, waarbij het wiskundehandboek beschouwd werd als een variabele op schoolniveau, stelden wij vast er voor het merendeel van de schalen een effect meetbaar is van het gebruikte handboek op de wiskunde prestaties. In de meeste gevallen bleek dit zoals verwacht een beperkt effect, maar voor enkele schalen verklaart het handboek met meer dan 15 procent toch een relatief groot percentage van de prestatieverschillen tussen scholen. Wanneer we het enkelvoudig effect van de handboeken verder in kaart brengen, stellen we sterke schommelingen per schaal vast, hetgeen erop lijkt te wijzen dat elk handboek sterkere en zwakkere onderdelen heeft.

Om het handboekeffect verder te onderzoeken, kiezen we het deeldomein *Cijferen* en de handboeken *Zo gezegd, zo gerekend!* en *Eurobasis* uit voor een kwalitatieve verdieping. *Cijferen* (schaal 9) is immers één van de schalen die een relatief hoog percentage verklaarde variantie vertonen met betrekking tot het gebruikte handboek (zie Tabel 1) en betrekkelijk grote verschillen laten optekenen tussen de verschillende handboeken (zie Figuur 2). Bovendien vormt cijferen een afgebakend onderwerp, dat relatief eenduidig te onderzoeken is in handboeken. *Zo gezegd, zo gerekend!* en *Eurobasis* zijn interessant voor een vergelijkende analyse omdat bij deze handboeken de prestaties

van de leerlingen significant sterk van elkaar afwijken voor cijferen: *Eurobasis* laat beduidend betere cijferprestaties optekenen. Bovendien bleken dit ten tijde van de meting de twee meest gebruikte handboeken te zijn in Vlaamse basisscholen.

### 3 Kwalitatieve handboekanalyse voor cijferen

#### 3.1 Inleiding

##### *Cijferen*

Bij cijferen wordt er gerekend door getallen onder elkaar te schrijven (met uitzondering van de staartdeling) en er een specifiek algoritme op toe te passen. Cijferen of 'schriftelijk rekenen' leidt steeds met zekerheid tot de juiste oplossing mits de regels correct worden toegepast. In Vlaanderen wordt cijferen doorgaans halverwege het derde leerjaar van de lagere school geïntroduceerd. Het is als volgt opgenomen in de Vlaamse eindtermen wiskunde voor het einde van het basisonderwijs (Vlaamse overheid, Agentschap voor Kwaliteitszorg in Onderwijs en Vorming, 2014, *1. Wiskunde – Getallen, Procedures*):

De leerlingen (...)

1.24 kennen de cijferalgoritmen. Zij kunnen cijferend vier hoofdbewerkingen uitvoeren met natuurlijke en met kommagetallen:

- optellen met max. 5 getallen: de som  $< 10\ 000\ 000$
- aftrekken: aftrektal  $< 10\ 000\ 000$  en max. 8 cijfers
- vermenigvuldigen: vermenigvuldiger bestaat uit max. 3 cijfers; het product = max. 8 cijfers (2 cijfers na de komma)
- delen: deler bestaat uit max. 3 cijfers; quotiënt max. 2 cijfers na de komma

Cijferen bekleedt een belangrijke plaats in het curriculum van de Vlaamse lagere school. Traditioneel wordt het samen met hoofdrekennen en het informeel noteren van tussenstappen ook beschouwd als één van de belangrijkste rekenhandelingen om bewerkingen op te lossen (Selter, 2001).

De laatste decennia lijkt cijferen echter terrein te verliezen ten voordele van andere rekenstrategieën, met name de zakrekenma-



chine en schattend of flexibel rekenen (van de Craats, 2007; Koninklijke Nederlandse Akademie van Wetenschappen [KNAW], 2009). Dit gegeven past in een algemene verschuiving van eerder traditionele kennisoverdracht naar meer realistische en leerlinggerichte onderwijsvormen, die voor het eerst op de voorgrond kwam in Nederland in de jaren 70 en 80 van de vorige eeuw (KNAW, 2009). In het debat tussen beide stromingen, dat enkele jaren geleden op zijn hoogtepunt was, had het zelfs een prominente rol. Eén van de hartenkreten van het opkomende realistisch rekenonderwijs was precies de oproep om het “geestdodende cijferen” (van den Heuvel-Panhuizen & Treffers, 2010) veel minder te gaan beklemtonen.

In het eerder traditionele rekenonderwijs ligt de nadruk op progressieve complicering (Feys & Van Biervliet, 2008) en het stap voor stap laten ‘inslijpen van procedures’ (Klein, Beishuizen, & Treffers, 1998). Bij cijferen, waaraan veel belang wordt gehecht in het traditionele rekenonderwijs, leert de leerkracht het standaardalgoritme in stapjes aan. De leerlingen oefenen het individueel en op papier in, en de oefeningen en toepassingen worden geleidelijk aan moeilijker. In de traditionele cijferdidactiek wordt weinig meerwaarde gezien in schatten of het gebruik van een zakrekenmachine, en in rekenopgaven wordt weinig gebruikgemaakt van betekenisvolle contexten. Er wordt van uitgegaan dat begrip en inzicht in de geleerde kennis en vaardigheden vanzelf zullen ontstaan en toenemen als gevolg van het veelvuldig oefenen (KNAW, 2009).

Het meer realistische rekenonderwijs is daarentegen gericht op het inzichtelijk aanleren van algoritmen en het zelf laten zoeken naar een geschikte oplossingsweg (Klein et al., 1998). In de instructie en het leerproces wordt een grote rol toebedeeld aan sociale interactie (Goos, 2004). Cijferen komt wel aan bod, maar dan als één van meerdere mogelijke rekenstrategieën en minder als doel op zichzelf. Men vertrekt voor cijferen vanuit het kolomsgewijs hoofdrekenen en maakt gebruik van progressieve schematisering, waarbij bewerkingen en voorstellingswijzen geleidelijk aan vereenvoudigd en verkort worden (Treffers, 1987). Problemen

worden in de regel steeds geformuleerd vanuit een realistische context en bestudeerd aan de hand van materiaal zoals een abacus of telraam. In deze benadering wordt ook een belangrijkere plaats toebedeeld aan schatten en het gebruik van een zakrekenmachine.

#### *Handboekanalyse*

*Zo gezegd, zo gerekend!* (ZGZG) en *Eurobasis* (EB) werden doelgericht geselecteerd voor een kwalitatieve vergelijking op basis van hun contrasterende peilingsresultaten: leerlingen die EB gebruikten in de klas, bleken het beter te doen voor cijferen dan leerlingen die werkten met ZGZR. We trachten beschrijvend te analyseren of er voor cijferen ook structurele en inhoudelijke verschillen zijn tussen deze handboeken. Gelet op de duidelijk uiteenlopende prestaties in de peilingsdata is onze hypothese dat er minstens enkele verschillen gevonden zullen worden.

Als leidraad voor onze analyse nemen wij een comparatieve handboekanalyse van Son (2005). Daarin vergeleek de auteur schoolboeken uit de Verenigde Staten en Korea om na te gaan hoe het curriculumonderdeel ‘vermenigvuldigen en delen van breuken’ erin aangebracht wordt. Zij voerde daarvoor een kwalitatieve inhoudsanalyse uit van de schoolboeken en bijbehorende handleidingen, om na te gaan hoe vorm gegeven wordt aan de boogde leerdoelen, en combineerde dat met een gedetailleerde analyse van de aangeboden opgaven (oefeningen) om zo tot een rijkere vergelijking te kunnen komen. Geïnspireerd door dit onderzoek, stellen wij onze eigen instrumenten voor analyse op.

Omdat de tegenstelling tussen traditioneel en realistisch rekenonderwijs met name voor cijferen een belangrijke rol lijkt te spelen, willen we in de interpretatie van onze resultaten ook nagaan of we in de handboeken een bepaalde aansluiting kunnen vinden bij deze benaderingen. Hoewel zowel EB als ZGZG volgens hun uitgeverij geworteld zijn in het realistisch rekenonderwijs, vermoeden we dat we in beide handboeken eerder een combinatie van handboekkenmerken zullen aantreffen dan een perfecte reflectie van één bepaalde stroming. *Freedom of design* (van Zanten & van den Heuvel-Panhuizen, 2014) geeft handboek-

ontwikkelaars tot op zekere hoogte de ruimte om zelf een interpretatie en een invulling te geven aan leerinhouden en didactische principes, en zelfs binnen één strekking of methode kan dat verschillen in aanpak opleveren. Voor het hier onderzochte cijferen is dat niet anders: zoals beschreven door van den Heuvel-Panhuisen en Treffers (2010) is er ondanks een gedeelde theoretische grondslag ook binnen het realistisch methodenbestand geen vastomlijnde aanpak wat dit deeldomein betreft. In de praktijk blijkt overigens over het algemeen dat realistisch gezinde methodes nog steeds moeilijk loskomen van, of neigen terug te keren naar een hang naar afgelijnde taken of *task propensity* (Gravemeijer, Bruin-Muurling, Kraemer, & van Stiphout, 2016) die eerder geassocieerd wordt met het traditionele rekenonderwijs. We verwachten bijgevolg dat we zeker geen scherpe tegenstelling tussen beide benaderingen zullen vinden in de onderzochte handboeken, noch dat er op basis van onze handboekanalyse concrete uitspraken mogelijk zullen zijn over de superioriteit van één van beide methodes. We verwachten wel bepaalde kenmerken te kunnen identificeren die bevestigen dat er een verschil is tussen beide handboeken, zoals de resultaten van onze eerste studie suggereren, en die eventueel interessant kunnen zijn voor verder onderzoek.

### 3.2 Methode

#### *Materiaal*

Van ZGZG en EB<sup>1</sup> verzamelden wij al het materiaal dat in 2010 beschikbaar was voor het derde tot en met het zesde leerjaar. Voor beide handboeken waren dat handleidingen voor de leerkracht, kopieerbladen, toetsen, remediërings- en differentiatieboeken of -bladen en werkboeken voor de leerlingen. Wanneer wij in de rapportage over de kwalitatieve analyse over 'het handboek' of een bepaald handboek spreken zonder verdere specificering, wordt het geheel van dit verzamelde materiaal bedoeld. Enkel eventuele uitbreidingen die gericht zijn op het leerplan van een specifiek onderwijsnet werden buiten beschouwing gelaten. De manier waarop het materiaal door individuele leerkrachten gebruikt wordt, werd niet bestudeerd, aangezien een onderzoek van de klaspraktijk een

volledige andere methodologie veronderstelt en de reikwijdte van dit onderzoek overstijgt.

#### *Analyses*

Op het verzamelde handboekmateriaal voerden wij in navolging van Son (2005) een inhoudsanalyse en een opgavenanalyse uit. Op basis van verschillende elementen uit de literatuur (Klein et al., 1998; Lucangeli, Tressoldi, Bendotti, Bonanomi, & Siegel, 2003; Rezat, 2006; Son, 2005; Törnroos, 2005; Valverde et al., 2002; Vlaams Verbond van het Katholiek Basisonderwijs, 2002) stelden we daarvoor twee onderzoeksinstrumenten op. In de eerste plaats stelden wij een analyseschema samen om de macrostructuur van elk handboek te kunnen beschrijven, met bijzondere aandacht voor de positie van cijferen in de wiskunde-instructie en het leerproces. In de tweede plaats maakten wij een codeerschema op om de aangeboden oefeningen in kaart te brengen. In het codeerschema was een lijst van overte kenmerken opgenomen, verdeeld over zes grote onderdelen: werkvormen, schatting, zakrekenmachine, presentatie, soort oefeningen, en controleren en corrigeren van oefeningen.

In het eerste gedeelte van de kwalitatieve inhoudsanalyse gebruikten wij het analyseschema om onder andere na te gaan of cijferen verweven is met andere thema's en aspecten van het wiskundeonderwijs, of dat het aan bod komt als een apart thema, en om in kaart te brengen hoeveel lestijd eraan wordt besteed. Vervolgens gingen we in een tweede gedeelte dieper in op de instructie- en oefenfase van het cijferonderwijs in de beide handboeken. Op basis van het analyseschema en het codeerschema, gingen we in detail na hoe het cijferalgoritme wordt opgebouwd, aangeleerd, toegepast, ingeoevend en geautomatiseerd. Welke strategieën worden aangereikt aan de leerling en (hoe) wordt hen inzicht in procedures bijgebracht? Hoe verhouden kennisoverdracht en interactie zich ten opzichte van elkaar? Welke principes van differentiatie, remediëring en evaluatie worden er gehanteerd? We brachten ook in kaart welke soorten opgaven (oefeningen) aangeboden worden in de

werkboeken voor de leerlingen. In de bespreking van de resultaten van deze analyse gaan we na of we bepaalde handboeken kunnen identificeren die we eerder als ‘traditioneel’ dan wel eerder als ‘realistisch’ kunnen categoriseren.

### Procedure

Twee onderzoekers analyseerden elk één handboek. Om de interbeoordelaarsbetrouwbaarheid te bepalen, werd al het materiaal van één willekeurig bepaald leerjaar, met name het vierde leerjaar, uitgewisseld en opnieuw gecodeerd met behulp van het codeerschema. De beide coderingen werden vervolgens met elkaar vergeleken voor de zes grote onderdelen van het codeerschema. Binnen elk onderdeel werd per kenmerk een score van nul of één toegekend, naargelang de overeenkomst tussen de beide analyses. Een maximale score van 100 procent overeenkomst zou betekenen dat beide onderzoekers tot exact hetzelfde resultaat kwamen, maar voor elk kenmerk waarvoor een verschil gevonden werd, ging er één punt van deze maximale score af. Voor het merendeel van de zes grote onderdelen van het codeerschema bleek de overeenkomst 90 tot 100 procent te bedragen, maar voor het nakijken van oefeningen in EB werd slechts 70 procent overeenkomst gevonden. Op basis van deze vaststelling werd de betrokken categorie uit het codeerschema geherformuleerd om de interpretatie ervan eenduidig te maken. De variabele werd vervolgens opnieuw gecodeerd, waarna de totale overeenkomst tussen de analyses van beide onderzoekers 95 procent bedroeg.

Op basis van de verzamelde gegevens gingen we na op welke punten ZGZG en EB overeenkomen, dan wel van elkaar verschillen. In de presentatie van de resultaten benadrukken we vooral de verschillen die we geïdentificeerd hebben, omdat deze ons het meeste inzicht kunnen verschaffen in de verschillende effectiviteit van beide handboeken.

### 3.3 Resultaten

#### *Cijferen in het lessenpakket*

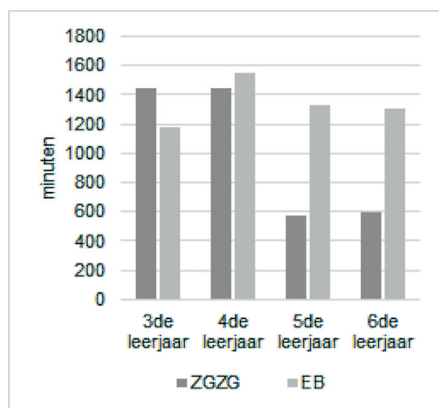
Het totale aantal uren wiskundeles over alle leerjaren heen, verschilt slechts minimaal tussen de beide handboeken. In totaal worden er 626 uur en 15 minuten wiskundeles inge-

pland bij ZGZG en 599 uur en 10 minuten bij EB. Wanneer we ervan uitgaan dat er 36 lesweken zijn per leerjaar, zouden we op basis van dit totaal dus een ruw gemiddelde kunnen berekenen van iets meer dan vier uur wiskundeles per week.

De handboeken verschillen wel in hun behandeling van de drie grote domeinen. Zo komt het domein *Toepassingen* in EB bijvoorbeeld als een apart onderdeel aan bod in elk leerjaar: tijdens toepassingslessen leren leerlingen elke week diverse problemen oplossen. Wekelijks staan hier één les of uitzonderlijk twee lessen cijferen op het programma. In ZGZG zijn toepassingen, en dus ook cijferen, geïntegreerd in de domeinen *Getallen en bewerkingen* en *Meten en meetkunde*. Hier wordt met thema's gewerkt die twee tot drie weken beslaan; per thema wordt er in één of meer lessen gecijferd.

In Figuur 3 wordt in absolute cijfers het totaal aantal lesminuten weergegeven dat de handboeken in elk leerjaar wijden aan cijferen. EB besteedt globaal genomen en zeker in de latere leerjaren veel meer onderwijstijd aan cijferen dan ZGZG. In ZGZG zien we voor cijferen een opvallende terugval in het vijfde en zesde leerjaar.

In Figuur 4 wordt weergegeven in welk schoolweek van het derde leerjaar de verschillende bewerkingen voor het eerst aan bod komen in de onderzochte handboeken. In EB wordt cijferen al iets vroeger in het schooljaar geïntroduceerd dan in ZGZG, en wordt ook elk apart algoritme vroeger aange-



Figuur 3. Aantal lesminuten gependend aan cijferen.

Schoolweek 3de leerjaar	ZGZG	EB
12		+
13		-
14		
15	+	
16		
17		
18		
19	-	x
20		
21	x	
22		
23		:
24		
25		
26		
27		
28		
29		
30	:	
31		

*Figuur 4.  
Tijdlijn: introductie van de verschillende bewerkingen.*

leerd. Hoewel de verschillende algoritmes verder in dezelfde volgorde worden aangeleerd in beide handboeken, wordt er in ZGZG meer tijd gelaten tussen het introduceren van optellen en aftrekken en tussen vermenigvuldigen en delen. Wel volgt in ZGZG de vermenigvuldiging sneller op het aftrekken dan in EB het geval is.

Over de vier onderzochte leerjaren heen stelden we verder vast dat in EB vroeger gestart wordt met grote getallen, terwijl er in ZGZG meer aandacht gaat naar kommagetalen.

### *Cijferen aanleren*

In beide handboeken begint een les cijferen doorgaans met een contextprobleem. In ZGZG zoeken de leerlingen daarvoor eerst zelfstandig of in groepjes een oplossing. Op basis daarvan volgt er een onderwijsleergesprek en van daaruit worden de algoritmen opgebouwd. De leerlingen oefenen vervolgens per twee met behulp van het werkboek. Bij EB verloopt de instructie meer leerkrachtgestuurd. De leerlingen mogen even nadenken of discussiëren over de probleemstelling, maar daarna legt de leerkracht het algoritme klassikaal uit met behulp van kopieerblaadjes en worden er afspraken gemaakt over de uitvoering van het algoritme. Deze afspraken worden opgehangen in de klas. De leerlingen oefenen de getoonde werkwijze zelfstandig in, ook met het werkboek.

Beide handboeken boden op het moment van dit onderzoek geen ondersteunende software aan, noch voor de leerkracht, noch voor de leerlingen. Hulpmiddelen zoals kurken, knikkers, kralen of MAB-materiaal (*Multi-base Arithmetic Blocks*: blokjes, staafjes, vierkanten en kubussen waarmee rekenkundige verhoudingen aanschouwelijk worden gemaakt) worden in ZGZG wel gebruikt voor de instructie maar ze worden achterwege gelaten wanneer de leerkracht van mening is dat het algoritme voldoende begrepen is. EB spoort zwakke rekenaars aan om het materiaal dat tijdens de instructie gebruikt werd, te blijven gebruiken tijdens het oefenen zolang ze dat nodig vinden. Sterke leerlingen kunnen een onthoudboekje met een samenvatting van de belangrijkste theorie gebruiken om problemen zelfstandig proberen op te lossen, voordat ze hulp inroepen van de leerkracht.

Beide handboeken adviseren de leerkracht om de leerlingen te observeren met het oog op differentiatie en remediëring. In EB wordt omschreven hoe deze observatie op een systematische wijze kan gebeuren, maar het verdere kader voor differentiatie en remediëring is wel iets minder afgelijnd dan in ZGZG. In ZGZG is differentiatie standaard ingebouwd in de les: in alle oefeningenreeksen is het voor de leerkracht duidelijk welke oefeningen bestemd zijn voor de hele leerlingengroep, en welke voor de zwakkere of de sterkere leerlingen. In dit handboek worden er vaste herhalingsweken ingelast en worden toetsen over het algemeen afgenomen na een thema. Toetsen kunnen steeds gevolgd worden door remediëringslessen waarvoor aparte werkblaadjes worden gebruikt. In EB worden herhalingslessen voor elk cijferonderdeel (bijvoorbeeld 'optellen en aftrekken tot 10 000') aangeboden, en worden die meestal gevolgd door een toets.

### *Cijferen toepassen en inoefenen*

In Tabel 2 wordt een overzicht gegeven van de verschillende soorten oefeningen die we terugvinden in de werkboeken voor het derde tot en met het zesde leerjaar. De werkboeken van EB bevatten meer cijferoefeningen dan die van ZGZG: we tellen er 2473 in EB tegenover 1963 in ZGZG. In EB zit er echter

Tabel 2  
Soorten oefeningen in de werkboeken

Soort oefening	ZGZG		EB	
	N	%	N	%
Kale oefening	1385	70.6	2085	84.3
Vraagstuk	89	4.5	240	9.7
Ontbrekende cijfers invullen	135	6.9		
Rekenverhaal schrijven	24	1.2		
Verbeteren of aanduiden	178	9.1		
Zelf oefeningen bedenken	56	2.9		
Rekentaal omzetten in formules	42	2.1		
Andere	54	2.8	148	6.0
Totaal	1963	100	2473	100

veel minder variatie in de soorten opgaven. We treffen enkel kale oefeningen en vraagstukken aan, terwijl in ZGZG ook oefeningen worden aangeboden waarbij de leerlingen bijvoorbeeld ontbrekende cijfers dienen in te vullen of fouten dienen te zoeken in reeds ingevulde oefeningen.

Kale oefeningen worden in beide handboeken op een gelijkaardige wijze gepresenteerd, evoluerend van ingevulde en voorgestructureerde oefeningen naar enkel een raster. In EB wordt het positiestelsel af en toe voorgesteld met tekeningen. De handboeken blijken te verschillen in hun aanpak van overschrijvingen, met name bij optellen en vermenigvuldigen. Wanneer er cijfers onthouden moeten worden, worden die in beide handboeken op een andere plaats genoteerd, en enkel in ZGZG wordt bij vermenigvuldiging het cijfer na verloop van tijd niet meer opgeschreven maar moeten de leerlingen het onthouden.

Terwijl in beide handboeken in de instructie aandacht gegeven wordt aan het controleren van oplossingen, wordt dat alleen in EB ook in de oefenfase nog nadrukkelijk gestimuleerd. Om een oplossing te controleren wordt in ZGZG overigens vooral gebruik gemaakt van een voorafgaande schatting en gebeurt schatten standaard op basis van afgeronde getallen. In EB daarentegen, leren de

leerlingen verschillende schatprocedures aan en zijn ze vrij om te kiezen welke zij gebruiken. Daarnaast wordt in dit handboek, met name vanaf het vierde leerjaar, vaak gebruik gemaakt van de zakrekenmachine (terwijl dat in ZGZG eerder zelden het geval is) en worden er ook andere methoden aangereikt ter controle, zoals het toepassen van de omgekeerde bewerking en het optellen van meerdere termen in een andere volgorde.

In ZGZG wordt geen informatie gegeven over huiswerk of over welke oefeningen daarvoor in aanmerking komen. EB doet dat wel: op de kopieerbladen die tijdens de instructie gebruikt worden staat ook steeds een oefeningenreeks die als huistaak bedoeld is, of eventueel – in het vijfde en zesde leerjaar – van pas kan komen tijdens hoekenwerk of contractwerk.

### 3.4 Discussie

Wij voerden een vergelijkende analyse uit van twee handboeken die in de peiling wiskunde basisonderwijs van 2009 erg verschillende prestaties lieten optekenen voor cijferen. We stellen vast dat de beide handboeken wat cijferen betreft ook structureel en inhoudelijk van elkaar verschillen op een aantal punten.

Zo valt op dat EB, dat het voor cijferen beter deed in de peiling, hier ook meer tijd

aan wijdt dan ZGZG. In EB zijn globaal genomen ook meer cijferoefeningen opgenomen. Deze vaststellingen lijken te bevestigen wat we intuïtief zouden verwachten: meer instructietijd en oefengelegenheid zorgen ervoor dat leerlingen beter presteren. Dit sluit aan bij de stelling dat een grotere *opportunity to learn* betere leerlingprestaties met zich meebrengt (Törnroos, 2005).

Het feit dat EB meer aandacht besteedt aan cijferen dan ZGZG ligt in de lijn van een eerder traditionele benadering van dit deeldomein. We vinden in het beter scorende EB nog meer ‘traditionele’ accenten voor cijferen. De instructie van cijferen kent in EB bijvoorbeeld een strikter verloop dan in ZGZG, en het is de leerkracht die het algoritme aanbrengt. Klassikale afspraken, bijvoorbeeld over notatie en formulering, worden in EB a priori en veel vroeger gemaakt dan in ZGZG. In EB ligt de nadruk iets meer op kale cijferoefeningen, terwijl ZGZG daarnaast ook een variëteit aan andere probleemstellingen aanbiedt. In EB werken de leerlingen meer op individuele basis en staan oplossingsstrategieën minder ter discussie dan in ZGZG, dat het onderwijsleergesprek in elke les promoot.

In ZGZG, dat iets minder goed scoorde voor cijferen, herkennen we over het algemeen een aantal kenmerken van realistisch rekenonderwijs die minder sterk geprononceerd zijn in EB. Leerlingen worden in ZGZG bijvoorbeeld meer gestimuleerd om eigen oplossingswijzen voor te stellen. Er ligt meer nadruk op interactie: het algoritme wordt vanuit een contextprobleem opgebouwd en benaderd door de groep. In ZGZG wordt ook iets meer nadruk gelegd op de samenhang met hoofdrekenen en andere eerder verworven kennis.

We stellen dus vast dat de onderzochte handboeken voor cijferen inderdaad in zekere mate van elkaar verschillen in hun aansluiting bij principes van het traditioneel versus het realistisch rekenonderwijs. Zoals verwacht is er echter geen sprake van een absolute tweedeling. In beide handboeken wordt bijvoorbeeld gebruikgemaakt van ondersteunend materiaal en mogen de leerlingen schatten en een zakrekenmachine gebruiken. Dat zijn kenmerken die we associëren met ‘realis-

tisch’ rekenen, en met name dat laatste kenmerk is in EB nog sterker aanwezig dan in ZGZG.

Op basis van onze analyses voor het onderzochte deeldomein kunnen we de handboeken niet volledig in één van beide strekkingen categoriseren. Bijgevolg is het ook niet mogelijk of wenselijk om één van beide benaderingen als de betere te kwalificeren. Wel zou het zinvol zijn om te trachten na te gaan welke van de gevonden kenmerken er nu precies ‘het meest toe doen’. Daarvoor is echter een systematische manipulatie van één of meerdere van deze kenmerken in een experimenteel design (met een random toewijzing aan condities) nodig. Dergelijk kleinschaliger onderzoek zou de resultaten van ons verkennend en correlatieel onderzoek verder kunnen aanvullen en verdiepen.

#### 4 Algemene conclusie en aanbevelingen voor verder onderzoek

Uit de data van de tweede peiling wiskunde in het Vlaamse basisonderwijs van 2009 blijkt dat gemeten prestatieverschillen tussen leerlingen voornamelijk verklaard kunnen worden vanuit leerlingkenmerken, maar dat het gebruikte wiskundehandboek ook een rol speelt voor sommige schalen. Het handboek-effect is met name substantieel voor *Cijferen*, *Probleemoplossende vaardigheden*, *Schatten met behulp van referentiepunten*, *Maateenheden* en *Inzicht in meetkundige relaties*. De effecten van de verschillende handboeken blijken per deeldomein wel sterk te verschillen, met een wisselend verloop van sterke en zwakke of zwakkere onderdelen. Dat betekent dat het niet mogelijk is om handboeken over de hele lijn als ‘beter’ of ‘slechter’ te bestempelen.

We kozen ervoor om een kwalitatieve handboekanalyse uit te voeren voor een deeldomein met een hoog percentage verklaarde variantie, namelijk *Cijferen*. We selecteerden twee handboeken die voor dit deeldomein qua prestaties ver uit elkaar liggen, namelijk *Zo gezegd, zo gerekend!* (ZGZG) en *Eurobasis* (EB), en gingen na op welke punten ze verschillen. De geïdentificeerde verschilpunten zouden immers mogelijk een invloed

kunnen hebben op de cijferprestaties van leerlingen.

De gevonden verschillen hebben we onder meer gesitueerd in het debat tussen traditioneel en realistisch rekenonderwijs. Het beter scorende EB blijkt voor cijferen meer kenmerken te vertonen die we associëren met traditioneel rekenonderwijs. In de traditionele benadering wordt meer belang gehecht aan het gebruik van algoritmes. Het lijkt voor de hand te liggen dat een methode die meer aandacht besteedt aan een bepaald deeldomein, en er dus een grotere opportunity to learn voor biedt, samenhangt met betere prestaties voor dit deeldomein. Toch hebben we vastgesteld dat de instructie van cijferen in EB ook realistische elementen bevat. Vanzelfsprekend hebben deze elementen ook een invloed op de prestaties van leerlingen. Dat betekent dat we op basis van onze gegevens zeker niet eenduidig één didactische methode kunnen verdedigen of bekritisieren, zelfs niet voor dit ene geselecteerde deeldomein.

In de onderzochte handboeken analyseerden we alleen hoe cijferen aangebracht wordt. Voor andere wiskundedomeinen, waarvoor we andere prestatieverschillen vaststelden in het eerste luik van ons onderzoek, ziet de aansluiting bij de verschillende rekenmethodes er mogelijk helemaal anders uit. Daarbij lijkt het aannemelijk dat leerlingen voor contrasterende, meer inzichtelijke deeldomeinen misschien wel betere scores halen wanneer die onderwerpen in de handboeken behandeld worden vanuit een meer realistische benadering. Op basis van onze beperkte gevalstudie kunnen wij deze hypothese niet staven, maar ze biedt wel stof voor verder onderzoek. Het feit dat handboeken over het algemeen een grillig prestatiepatroon vertonen over (deel) domeinen heen lijkt alleszins te suggereren dat elk handboek ook zijn eigen klemtonen legt in de uitwerking en aanpak van de verschillende (deel)domeinen.

Wij hebben getracht om aan te tonen dat *mixed methods* onderzoek een meerwaarde kan bieden voor zowel handboekanalyse en rekendidactiek als onderwijseffectiviteitsonderzoek. Het feit dat wij toetsprestaties uit een peilingsonderzoek als selectiebasis genomen hebben voor onze gevalstudie, heeft

echter enkele specifieke consequenties. Peilingen meten het bereiken van de eindtermen, in dit geval de vooropgestelde minimumdoelstellingen voor wiskunde aan het einde van het basisonderwijs. Peilingstoetsen worden bijgevolg steeds afgenomen aan het einde van het zesde leerjaar. Uit onze kwalitatieve analyse kwam naar voor dat EB in tegenstelling tot ZGZG ook in het vijfde en zesde leerjaar nog steeds erg veel aandacht besteedt aan cijferen – het is dus niet ondenkbaar dat ook dit een invloed heeft op de resultaten. Daarnaast wijzen we erop dat er vanuit het peilingonderzoek geen uitspraak gedaan kan worden over de langetermijneffecten van het wiskundeonderwijs. Zo kunnen we op basis van dit onderzoek bijvoorbeeld niet voorspellen of leerlingen die meer inzicht hebben in cijferen later beter of handiger gebruikmaken van cijferalgoritmen dan degenen die vooral goed zijn in het oplossen van klassieke opgaven.

Tot slot willen we nogmaals het verkennende karakter van ons onderzoek benadrukken. Onze gevalstudie concentreerde zich op twee handboeken die in de resultaten van het peilingsonderzoek in 2009 extreem scoorden voor cijferen ten opzichte van de andere handboeken. Er werd enkel onderzocht of in een inhoudelijke analyse van deze twee handboeken elementen gevonden werden die in de lijn liggen van het gevonden prestatieverschil. Wij kunnen ons er niet over uitspreken of het prestatieverschil tussen deze twee handboeken consistent is over de tijd en analoog is voor andere wiskundedomeinen. Om dat na te gaan en om te kunnen weten welke individuele handboekkenmerken (zoals het aantal oefeningen of het soort van oefeningen) de prestaties effectief beïnvloeden, en in welke mate, is verder onderzoek nodig. Daarin zou ook de volgende ‘trede’ van de implementatietrap betrokken kunnen worden: het gebruik van de handboeken door de leerkrachten in de klas. Hoe ‘goed’ of ‘slecht’ een handboek of lesmethode immers ook is, veel hangt af van hoe de leerkracht dit *potentially implemented curriculum* aanwendt in de praktijk. Alleszins zou het wenselijk zijn om in verder onderzoek ook langetermijneffecten te bestuderen, contrasten te (onder)zoeken voor andere wiskundeonderdelen en

patronen te identificeren in andere handboeken, om uiteindelijk te komen tot een meer verscheiden en volledig beeld van de kenmerken die de wiskunde-prestaties van leerlingen in het algemeen positief kunnen beïnvloeden. Handboekontwikkelaars zouden deze wetenschappelijke inzichten kunnen aangrijpen ter inspiratie voor een ‘ideaal’ handboek. Dezelfde oefening kan bovendien ook gemaakt worden voor andere vakken en leergebieden.

## Noot

<sup>1</sup> EB werd op dat moment niet meer uitgegeven omdat die Keure intussen in 2006 een nieuwe rekenmethode had gelanceerd: *Kompas*. Toch werd EB nog steeds door veel scholen als wiskunde-handboek gebruikt.

## Literatuur

Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: development status and directions. *ZDM*, 45(5), 633-646.

Feys, R., & Van Biervliet, P. (2008). Wiskundeoorlog en reddingsoperatie in Nederland. *Onderwijskrant*, 146, 3-15.

Gielen, S., Willem, L., De Meyst, M., Beringhs, S., Crynen, M., Luyten, B., & Janssen, R. (2010). *Peiling wiskunde in het basisonderwijs - Eindrapport*. Leuven: KU Leuven, Centrum voor Onderwijseffectiviteit en -evaluatie.

Goos, M. (2004). Learning mathematics in a classroom community of inquiry. *Journal for research in mathematics education*, 35, 258-291.

Gravemeijer, K., Bruin-Muurling, G., Kraemer, J. M., & van Stiphout, I. (2016). Shortcomings of Mathematics Education Reform in The Netherlands: A Paradigm Case?. *Mathematical Thinking and Learning*, 18(1), 25-44.

Klein, A. S., Beishuizen, M., & Treffers, A. (1998). The empty number line in Dutch second grades: Realistic versus gradual program design. *Journal for Research in Mathematics Education*, 29, 443-464.

Koninklijke Nederlandse Akademie van Wetenschappen [KNAW]. (2009). *Rekenonderwijs op de basisschool, analyse en sleutel tot verbeter-*

*ring*. Alkmaar: Bejo druk & print.

Lucangeli, D., Tressoldi, P.E., Bendotti, M., Bonanomi, M., & Siegel, L.S. (2003). Effective strategies for mental and written arithmetic calculation from the third to the fifth grade. *Educational Psychology*, 23(5), 507-520.

Luyten, B., Van Nijlen, D., Janssen, R., De Corte, E., De Boeck, P., Verschaffel, L., Daems, F., & Rymenans, R. (2003). *Eerste peiling wiskunde en lezen in het basisonderwijs - Eindrapport*. Leuven - Antwerpen: K.U.Leuven, Leuvens Instituut voor Onderwijsonderzoek - Universiteit Antwerpen, Departement Didactiek en Critiek.

Patton, M. (2015). *Qualitative research & evaluation methods*. Thousand Oaks, CA: SAGE Publications.

Pingel, F. (2010). *UNESCO guidebook on textbook research and textbook revision*. Unesco.

Rezat, S. (2006). The structures of German mathematics textbooks. *ZDM*, 38(6), 482-487.

Selter, C. (2001). Addition and subtraction of three-digit numbers: German elementary children's success, methods and strategies. *Educational Studies in Mathematics*, 47(2), 145-173.

Son, J. (2005). A comparison of how textbooks teach multiplication of fractions and division of fractions in Korea and in the U.S. In H. L. Chick & J. L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 201-208). Melbourne: International Group for the Psychology of Mathematics Education.

Törnroos, J. (2005). Mathematic textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation*, 31, 315-327.

Treffers, A. (1987). Integrated column arithmetic according to progressive schematization. *Educational Studies in Mathematics*, 18, 125-145.

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book. Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Dordrecht: Kluwer Academic Publishers.

van de Craats, J. (2007). Waarom Daan en Sanne niet kunnen rekenen. *Nieuw Archief voor Wiskunde*, 8(2), 132.

van den Heuvel-Panhuizen, M., & Treffers, A. (2010). Cijfer positieve prestaties in rekenen niet weg. *Nieuw archief voor wiskunde*. Serie 5, 11(2), 120-124.



Van Steenbrugge, H., Valcke, M., & Desoete, A. (2013). Teachers' views of mathematics textbook series in Flanders: Does it (not) matter which mathematics textbook series schools choose?. *Journal of Curriculum Studies*, 45(3), 322-353.

van Zanten, M., & van den Heuvel-Panhuizen, M. (2014). Freedom of design: The multiple faces of subtraction in Dutch primary school textbooks. In Y. Li & G. Lappan (Eds.), *Mathematics curriculum in school education* (pp. 231-259). Heidelberg/Dordrecht/London/NewYork: Springer.

Vandenbergh, R. (2004). Over stuurbaarheid van het onderwijs. Een analyse van sturend beleid, resultaten en niet-bedoelde effecten. In G. Kelchtermans (Red.), *Studia Paedagogica 37. De stuurbaarheid van onderwijs. Tussen kunnen en willen, mogen en moeten* (pp. 89-120). Leuven: Universitaire Pers.

Vlaams Verbond van het Katholiek Basisonderwijs. (2002). *Bewerkingen. Toelichtingen*. Brussel: Centrale Raad Katholiek Lager Onderwijs.

Vlaamse overheid, Agentschap voor Kwaliteitszorg in Onderwijs en Vorming. (2011). *Peiling wereldoriëntatie (tijd, ruimte, maatschappij en brongebruik) in het basisonderwijs*. Brochure. Brussel: Vlaamse overheid.

Vlaamse overheid, Agentschap voor Kwaliteitszorg in Onderwijs en Vorming. (2014). *Lager onderwijs - Wiskunde - Eindtermen. 2: Eindtermen*. Geraadpleegd op 20 november 2014 via <http://www.ond.vlaanderen.be/curriculum/basisonderwijs/lager-onderwijs/leergebieden/wiskunde/eindtermen.htm>

## Auteurs

**Rianne Janssen** en **Wim Van Dooren** zijn beiden hoofddocent aan de Faculteit Psychologie en Pedagogische Wetenschappen van de KU Leuven. **Rianne Janssen** coördineert de onderzoekseenheid Onderwijskunde en is verbonden aan het Centrum voor Onderwijseffectiviteit en -evaluatie, **Wim Van Dooren** is verbonden aan het Centrum voor Instructiepsychologie en -technologie. Onder hun begeleiding maakten **Isabel Avau** en **Ilse Thomassen** een masterproef over dit onderwerp ter verkrijging van de graad van Master in de Pedagogische Wetenschappen. **Eef Ameer** en **Evelyn Goffin** zijn respectievelijk projectcoördinator en medewerker

van het Steunpunt Toetsontwikkeling en Peilingen, onderdeel van het Centrum voor Onderwijseffectiviteit en -evaluatie.

*Correspondentieadres:* E. Goffin, Centrum voor Onderwijseffectiviteit en -evaluatie, KU Leuven, Dekenstraat 2 bus 3773, 3000 Leuven. E-mail: [evelyn.goffin@kuleuven.be](mailto:evelyn.goffin@kuleuven.be).

## Abstract

**Is there a relationship between pupils' performance of standard algorithms and the mathematics textbook used in primary school? A mixed methods analysis.**

In an exploratory mixed methods analysis we investigate pupils' performance of standard written algorithms in arithmetic in Flemish primary schools, and whether it matters which mathematics textbook is used for instruction. Our first, quantitative study consists of a multilevel regression analysis on data from the 2009 national mathematics assessment. We establish that the textbook in fact has an effect on pupils' learning outcomes for several domains. In our second, qualitative study, we compare two popular textbooks that differed in assessment results for the domain of standard written algorithms. We analyze contrasts in textbook structure and content, as these might provide an explanation for why the selected textbooks score differently. We find that the textbook with the better scores attributes more time and more exercises to standard written algorithms, and displays different didactic principles.

**Keywords:** textbook research, mixed methods, purposive sampling, national assessment, mathematics textbook, primary education, standard written algorithms