



'Verder met rekenen' als voorbeeld van evidence-based onderwijsontwikkeling

- pleidooi voor het combineren van effectevaluatie met procesanalyse -

K. Buijs, C.M. van Putten, E. Folmer & N. Nieveen
SLO, Enschede / Universiteit Leiden / SLO, Enschede

Deze bijdrage sluit aan bij de discussie over de evidence-based onderwijsinnovaties. Naar aanleiding van het rapport van de Onderwijsraad (2006), waarin gepleit wordt voor zo hard mogelijke onderzoeksmethoden naar de effectiviteit van innovaties, argumenteerden Gravemeijer en Kirschner (2007, 2008) enige tijd geleden dat het via dergelijke methoden gegenereerde bewijs voor effectiviteit betrekkelijk is. Zij betoogden dat er veeleer behoefte is aan een ander type, meer procesgericht onderzoek dat zich richt op het vastleggen van hoe een innovatie precies werkt, en op het verzamelen van bewijs voor een bijbehorende onderwijstheorie. Deze bijdrage beoogt duidelijk te maken dat het, in de context van leerplanontwikkeling, juist de combinatie van beide typen onderzoek is die tot effectieve en bruikbare leerplanproducten kan leiden. Om dit standpunt te onderbouwen, wordt een leerplanproject op het gebied van het reken-wiskundeonderwijs als voorbeeld genomen. Eerst wordt beschreven hoe binnen SLO, het nationaal expertisecentrum leerplanontwikkeling, onderzoeksmatige leerplanontwikkeling gestalte krijgt in de vorm van ontwerponderzoek (Educational Design Research; Van den Akker, 2009). Daarna worden de achtergronden en inhoud van het betreffende project 'Verder met Rekenen' (Buijs, 2009) geschetst. Op basis van een analyse van de resultaten daarvan wordt vervolgens beargumenteerd waarom het juist de genoemde combinatie van onderzoek naar effectiviteit met onderzoek naar een kern van 'essentiële werkingsmechanismen' is, die tot evidence-based leerplanproducten kan leiden. Het artikel sluit af met een beschouwing over twee van de gehanteerde kwaliteitscriteria, te weten effectiviteit en bruikbaarheid.

1 Inleiding: de discussie over evidence-based onderwijs

Enige tijd geleden ontstond een discussie over de vraag in hoeverre onderwijsinnovaties gebaseerd dienen te zijn op zo hard mogelijk bewijs voor de effectiviteit daarvan. Aanleiding vormde het advies van de Onderwijsraad (2006) waarin een pleidooi wordt gehouden voor evidence-based onderwijs waarbij innovaties vooraf worden gegaan door gedegen onderzoek naar effectiviteit. Experimenteel onderzoek waarbij gewerkt wordt met gerandomiseerde steekproeven, werd aangevoerd als de meest harde vorm van bewijs dat boven tafel gebracht kan worden. Gravemeijer en Kirschner (2007) plaatsten kritische kanttekeningen bij dit pleidooi. Zij betoogden dat er niet alleen het nodige valt af te dingen op de uitvoerbaarheid en generaliseerbaarheid van dergelijk onderzoek, maar ook dat er in de eerste plaats een ander type, meer procesgericht onderzoek, nodig is dat zich richt op het ontwikkelen van een funderende onderwijstheorie bij een innovatie. Zo'n innovatie wordt in hun visie pas bruikbaar als sprake is van een bewezen theorie over de essentiële werkingsmechanismen die verklaren hoe de innovatie werkt en die voor potentiële gebruikers onmisbare informatie bevat met het oog op aanpassing aan de eigen

onderwijssituatie. In het verlengde van dat artikel kwam een discussie tot stand waarin aan de ene kant een van de leden van de Onderwijsraad, Bosker (2008a en b), het pleidooi voor effectiviteitsonderzoek ondersteunde, terwijl Gravemeijer en Kirschner (2008) aan de andere kant hun bezwaren tegen zulk onderzoek verder uiteenzetten.

In deze bijdrage wordt de problematiek van de wenselijkheid van evidence-based onderwijsinnovaties besproken vanuit de optiek van leerplanontwikkeling. Binnen SLO, het nationaal expertisecentrum leerplanontwikkeling, wordt ernaar gestreefd om nieuwe leerplanproducten op een zodanige wijze te ontwikkelen dat tevens bewijs wordt verzameld omtrent de bruikbaarheid en effectiviteit van zulke producten. Daartoe worden leerplanprojecten zoveel mogelijk ingericht als ontwerponderzoek (Van den Akker, Gravemeijer, McKenney & Nieveen, 2006; Plomp, 2009). Dergelijk onderzoek begint gewoonlijk met een vooronderzoek waarin het betreffende leerplankundige probleem wordt geanalyseerd en waarbij mogelijke opties voor een oplossing worden overwogen. Het voorlopige ontwerp waarin dit vooronderzoek uitmondt, maakt vervolgens een zeker rijpingsproces door waarbij groepen experts en potentiële gebruikers worden geconsulteerd en waarbij de relevantie en consistentie worden onderzocht. Levert dit voldoende perspectief voor verdere ontwikkeling op, dan volgt een

onderwijspilot waarin de effectiviteit en bruikbaarheid in eerste ronde worden getest (Nieveen, 2009; Van den Akker, Nieveen & Kuiper, 2013). Op basis van de resultaten worden bijstellingen aangebracht waardoor het ontwerp aan kracht en bruikbaarheid wint, waarna een volgende pilot plaatsvindt die uiteindelijk tot een goed eindproduct moet leiden. Een centrale vraag bij deze werkwijze is de vraag hoe dergelijke onderwijspilots het beste ingericht kunnen worden. Dient het onderzoek zich, zoals Bosker bepleit, primair te richten op het achterhalen van de effectiviteit van de innovatie, bij voorkeur via zo hard mogelijke onderzoeksmethoden aan de hand van ‘echte experimenten’? Of dient de onderzoeksmethode veel meer procesgericht te zijn, zoals Gravemeijer en Kirschner betogen, en bij voorkeur te mikken op het genereren van ‘bewezen’ kennis over de essentiële werkingsmechanismen?

Aan de hand van een voorbeeld op het gebied van het reken-wiskundeonderwijs wordt hieronder betoogd dat het juist de combinatie van beide typen onderzoek is die voor doeleinden van leerplanontwikkeling perspectief biedt en die voor de onderwijspraktijk bruikbare en betrouwbare informatie oplevert. Het genereren van kennis omtrent de effectiviteit van een innovatief leerplanproduct, zo wordt betoogd, dient hand in hand te gaan met het genereren van kennis omtrent wat genoemd wordt de fundamentele werkingsmechanismen van de innovatie. Aan de hand van het project ‘Verder met Rekenen’ (vmR) wordt dit standpunt nader uitgewerkt. Eerst wordt de leerplankundige context onder de loep genomen, gevolgd door een toelichting van het project in kwestie. Beschreven wordt hoe er in het onderzoek sprake was van twee onderzoekscomponenten: de eerste gericht op onderzoek naar effectiviteit, de tweede bedoeld om het feitelijke proces in de klas te analyseren en te onderzoeken wat de essentiële werkingsmechanismen van de innovatie zijn. Nadat is beschreven wat de voornaamste resultaten van het effectiviteitsonderzoek waren, wordt de aandacht op de tweede onderzoekscomponent gericht. Er wordt een drietal relevante patronen in de procesgegevens beschreven die naar aanleiding van een retrospectieve analyse naar voren kwamen. Op basis daarvan wordt vervolgens beargumenteerd waarom juist de genoemde combinatie van onderzoek, gericht op het vaststellen van leereffecten met onderzoek en gericht op het analyseren van het proces dat zich tijdens het onderwijsexperiment afspeelt, zo vruchtbaar is.

2 Ontwerponderzoek in de context van leerplanontwikkeling

Leerplanontwikkeling vindt niet in het luchtledige plaats. Het zijn veelal reëel bestaande, in de onderwijspraktijk geconstateerde complexe problemen waarvan in brede

kring wordt onderkend dat het om maatschappelijk en onderwijskundig relevante problemen gaat, die het aangrijppingspunt vormen. Beleidsmatige ondersteuning vanuit de overheid alsmede samenwerking met relevante partners zoals groepen scholen, onderwijsbegeleidingsdiensten en universiteiten vormen belangrijke aandachtspunten. Om te bereiken dat er een zekere garantie is dat de beoogde innovatie of interventie tot passende onderwijsresultaten leidt, worden leerplanprojecten veelal ingericht als een vorm van ontwerponderzoek. Kenmerkend voor deze onderzoeksvorm is onder meer dat zij interventionistisch van aard is in de zin dat zij gericht is op het ontwerpen en beproeven van oplossingen voor problemen in de alledaagse onderwijspraktijk (Van den Akker e.a., 2006; Gravemeijer & Cobb, 2006; Plomp, 2009; Nieveen, 2009). Voorts is kenmerkend dat het onderzoeksproces iteratief van karakter is in de zin dat opeenvolgende benaderingen van ideaaltypische interventies worden ontworpen, beproefd en geëvalueerd (Kelly, 2006). Een derde belangrijk kenmerk is dat ontwerponderzoek theoriegeoriënteerd is in de zin dat het beoogde ontwerp gebaseerd is op relevante theoretische vooronderstellingen en dat het onderzoek bijdraagt aan verdere theorievorming.

In de context van leerplanontwikkeling zijn er op dit laatste punt evenwel beperkingen. Weliswaar zijn leerplanontwerpen vanzelfsprekend gebaseerd op onderliggende vakinhoudelijke en onderwijskundige theorieën. Maar anders dan in andere onderzoekssettings, is het primaire doel niet om bij te dragen aan voortgaande theorievorming. Doel is vooraleerst om de effectiviteit en bruikbaarheid van een interventie te testen teneinde potentiële gebruikers de nodige ter zake doende informatie te kunnen verschaffen. In dat verband wordt wel gesproken van ‘ontwikkelings-ondersteunend onderzoek’ (Van den Akker e.a., 2013). Uiteraard betekent dit meer dan alleen maar het verzamelen van gegevens over effectiviteit in bepaalde experimentele situaties. Het gaat er evenzeer om betrouwbare informatie te genereren over de wijze waarop de beoogde interventie in de praktijk kan werken. Zulke informatie hoeft echter niet het karakter te hebben van een volledige lokale onderwijstheorie, zoals in andere onderzoekssettings soms wordt nagestreefd (zie bijvoorbeeld: Van Nes, 2009). Volstaan kan worden met wat de ontwerpers gelet op de doelen van de interventie én op de ervaringen en uitkomsten van het onderwijsexperiment als de grondprincipes van de interventie beschouwen. Men zou dit de essentiële werkingsmechanismen van interventie kunnen noemen - de kern van mechanismen die maken dat de interventie anders (en beter) werkt dan gangbare, veelal ontoereikende benaderingen. Om de kwaliteit van leerplanproducten te waarborgen, wordt door SLO gewoonlijk een aantal kwaliteitscriteria gehanteerd. De eerste twee daarvan, relevantie en consistentie, worden veelal in de eerste onderzoeksfasen gewaarborgd. De derde en vierde, bruikbaarheid en

effectiviteit, komen in het kader van de tweede, experimentele onderzoeksfase onder de aandacht. Hoe dit in het onderhavige project is gerealiseerd, wordt hieronder verder besproken.

3 Verder met rekenen

Voorbeeld van onderzoeksmatige leerplanontwikkeling

‘Verder met Rekenen’ is een driejarig leerplanproject (2008-2010) waarin een prototype van een onderwijsleertraject voortgezet rekenen werd ontwikkeld bestemd voor klas 1 van de basisberoepsgerichte leerweg van het vmbo (Buijs & Van der Zwaard, 2011). Achtergrond van het project vormt het feit dat de betreffende groep leerlingen veelal met een gebrekkige rekenvaardigheid het voortgezet onderwijs instroomt. Wordt er niets aan de verbetering van deze rekenvaardigheid gedaan, dan is de kans groot dat deze steeds verder achteruitgaat en dat leerlingen in toenemende mate moeite hebben om zich op rekengebied staande te houden, met als gevolg dat bepaalde leerlingen in de bovenbouw van het vmbo of in het aansluitende mbo soms voortijdig de school verlaten. Door de commissie Meijerink (2008) werd dan ook geconstateerd dat er voor deze groep leerlingen onvoldoende sprake is van een doorlopende leerlijn, en werden aanbevelingen gedaan om zowel in de bovenbouw van het primair onderwijs als in de onderbouw van het voortgezet onderwijs het onderwijsaanbod beter op deze leerlingen af te stemmen.

Hierop inhakend werd door het ministerie van OCW vastgesteld dat voortgezet rekenen in de onderbouw van het voortgezet onderwijs een duidelijkere plaats dient te krijgen (OCW, 2008), waarbij vanaf 2014 via een landelijke toets zal worden gepeild hoe de leerresultaten zich ontwikkelen. Wat de problematiek complex maakt, is onder meer de grote variëteit aan beginniveaus van instromende leerlingen (Janssen, Van der Schoot & Hemker, 2005; Gille, Loijens, Noijens & Zwitser, 2010). Bovendien heeft een deel van de leerlingen door gebrek aan succeservaringen in de bovenbouw van het primair onderwijs veelal een aversie tegen rekenen ontwikkeld die het lastig maakt om ze binnen dit vakgebied verdere progressie te laten boeken.

Een typerend kenmerk van het leerplan in de onderbouw van het voortgezet onderwijs is verder dat rekenen tot voor kort een enigszins onderbelicht leerstofgebied vormde. Weliswaar bevatten de bestaande wiskundemethodes in klas 1 een of twee hoofdstukken die op rekenen betrekking hebben, maar hierin wordt er veelal vanuit gegaan dat de rekenvaardigheid bij aanvang reeds op een redelijk peil is zodat volstaan kan worden met een

beknopte samenvatting van wat op de basisschool geleerd is.

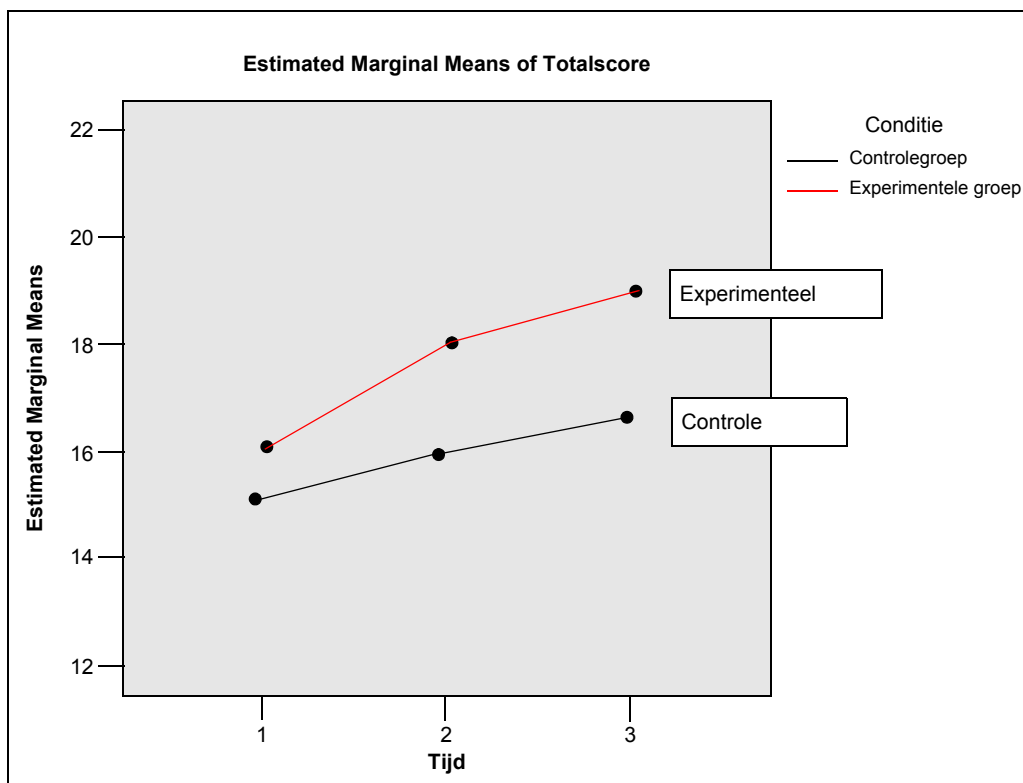
In aanvulling hierop is er de laatste jaren een aantal oefenprogramma's in boekvorm of digitaal in omloop gekomen die zich vooral richten op zelfstandig oefenen van basiskennis, standaardprocedures van het rekenen, kennis van breuken, procenten en verhoudingen, en dergelijke.

Het valt echter te betwijfelen of dergelijke programma's effectief zijn, gelet op het feit dat het onderwijs voor de betreffende leerlingen in de groepen 7 en 8 deels ook al gericht was op zelfstandig oefenen. Met het oog op deze omstandigheid werd bij vmr gekozen voor een op de ontwikkeling van inzicht gerichte, interactieve onderwijsbenadering waarbij gebruikssituaties, dat zijn voor de leerlingen relevante situaties uit het alledaags leven of uit de beroepspraktijk, en het leren hanteren van informele, modelondersteunde werkwijzen een centrale rol spelen (Buijs & Van der Zwaard, 2007).

In voorafgaande projecten werd reeds een analyse van de geschetste problematiek gemaakt waarbij voorbeeldmatig werd onderzocht op welke wijze voortgezet rekenen thans in de onderbouw van het vmbo aan de orde komt (Van der Zwaard, 2005) en waarbij een toets in klas 1 van vmbo en havo/vwo werd afgenomen (Buijs, Van Scherpenzeel, Ten Voorde & Van der Zwaard, 2008). Mede op grond hiervan werd een prototype van een onderwijsleertraject ontwikkeld rond drie rekenkundige kerninhouden, namelijk ‘geld’, ‘procenten’ en ‘meten’. In twee opeenvolgende pilots van een jaar werd dit traject op vier respectievelijk acht vmbo-scholen beproefd.

In beide pilots werd gewerkt volgens een quasi-experimentele onderzoeksofzet met niet-gerandomiseerde experimentele en controlegroepen (Shadish, Cook & Campbell, 2002), ontwikkeld in samenwerking met de Universiteit Leiden. Om de progressie in leerresultaten te meten, werden onder alle leerlingen in de experimentele en de controlegroep drie grotendeels identieke toetsen vooraf, halverwege en aan het eind van het traject afgenomen waarvan de resultaten op de Universiteit Leiden statistisch zijn geanalyseerd. Bij de aldus uitgevoerde analyses bleek zich tijdens de eerste pilot een bescheiden maar significante vooruitgang voor te doen van de experimentele groep ten opzichte van de controlegroep (De Kraker, 2009; Plomp, 2010). Tijdens de tweede pilot tekende zich een soortgelijke trend af.

In figuur 1 een overzicht van de toetsresultaten bij de eerste, tweede en derde toets op vier scholen ($n = 219$) bij deze tweede pilot. Daarnaast werden regelmatig signalen geregistreerd die erop wezen dat zich ook in andere opzichten een positieve ontwikkeling voordeed in de zin dat leerlingen minder aversie tegen het rekenen gingen vertonen en dat ze het op prijs gingen stellen op hun eigen denkvermogen aangesproken te worden (‘hier leer je



figuur 1: overzicht van de totaalscores op vier scholen tijdens de tweede pilot ($n = 219$). Zowel in de experimentele als in de controlegroep is er een verbetering van de prestaties; in de experimentele groep is deze sterker dan in de controlegroep. De getallen langs de verticale as verwijzen naar het aantal goed gemaakte opgaven (op een totaal van 33)

nadenken'; Van Os, 2010). Evenzo betoonden docenten soms hun instemming met de gekozen inzichtelijke interactieve aanpak.

Tijdens de pilots werden systematisch gegevens verzameld over het uitgevoerde onderwijs. Deze werden gebruikt om het onderwijsleertraject te verbeteren en om een systeem te ontwikkelen voor het vaststellen van de mate van implementatie. Daarnaast werden ze gebruikt om patronen op te sporen van regelmatig binnen alle scholen optredende gebeurtenissen die aanwijzingen opleverden voor de geldigheid van theoretische vooronderstellingen rond essentiële werkingsmechanismen en voor de mogelijke invloed van bepaalde uitvoeringsomstandigheden op de mate van effectiviteit. In de volgende paragraaf worden drie van dergelijke patronen beschreven.

4 Drie relevante patronen uit de onderzoeksgegevens

Behalve gegevens over de resultaten van het onderwijsleertraject werden ook gegevens verzameld over de wijze

waarop de lessen werden uitgevoerd, over de omstandigheden waaronder dit gebeurde en over het soort van redeneringen en oplossingen dat leerlingen bleken te produceren. Deze gegevens werden per school geanalyseerd om te achterhalen in hoeverre zich vergelijkbare patronen voordeden. De bedoeling was om, 'over de scholen en over de pilots heen', zo goed mogelijk zicht te krijgen op de ontwikkeling van de prestaties van de leerlingen, op de essentie van het doorlopen onderwijsleerproces en op de voornaamste factoren en omstandigheden die daarbij van invloed leken. De aldus opgezette onderzoekscyclus werd tijdens het onderzoek twee keer doorlopen - bij een eerste pilot met vier scholen, bij een tweede pilot met acht scholen. Bij beide pilots viel tussentijds een school uit, omdat men de afspraken over het uitvoeren van de interventie niet goed kon nakomen. In een afsluitende analyse van alle aldus verzamelde gegevens kwamen onder meer de volgende drie patronen naar voren.

Samenhang tussen de mate van effectiviteit en mate van implementatie

Er zijn duidelijke aanwijzingen dat de mate van effectiviteit van de interventie samenhangt met de mate van implementatie. In situaties waarin de interventie in grote lijnen werd uitgevoerd, zoals deze bedoeld is, blijken de

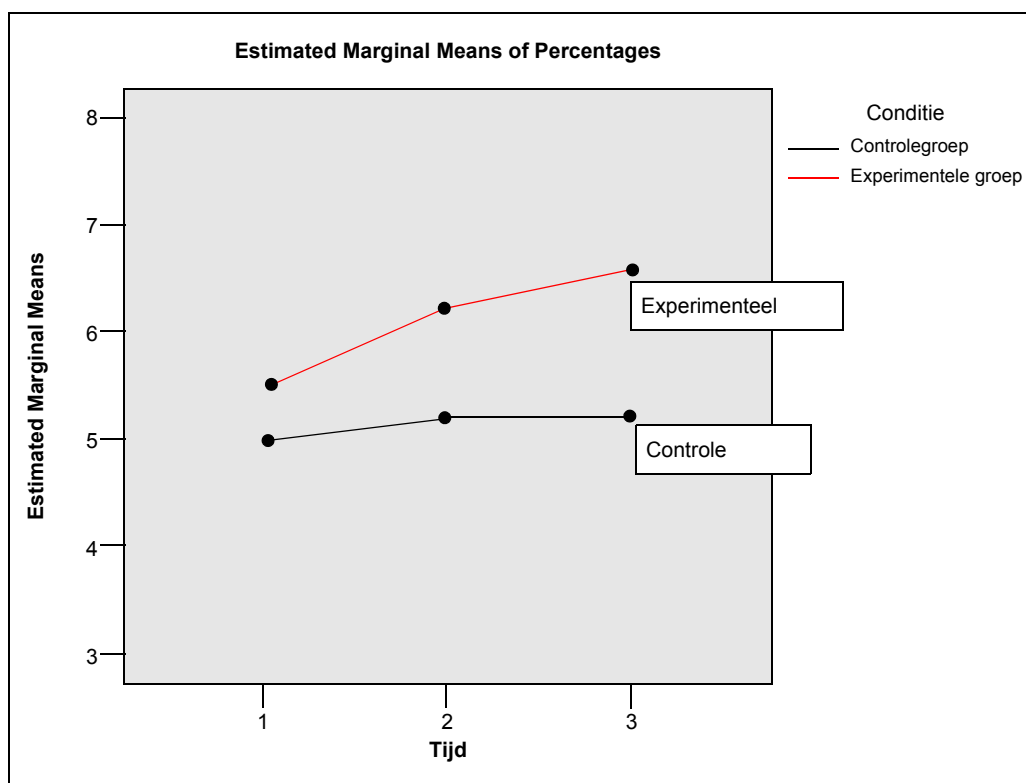
resultaten sterker vooruit te gaan dan in situaties waarin dat niet het geval was. Dat lijkt zowel te gelden voor de prestaties van de leerlingen zoals die in de toetsresultaten tot uitdrukking kwamen, als voor enkele andere onderzochte vaardigheidsaspecten zoals strategiegebruik en notatiegedrag (Stals, 2010; Plomp, 2010). Het komt erop neer dat als een docent erin slaagt om het interactieve, inzichtelijke aspect van het onderwijs redelijk goed gestalte te geven, met aandacht voor de eigen informele kennis van leerlingen, de leeropbrengsten, een sterkere stijging vertonen dan in situaties waarin een docent het onderwijs vormgeeft zonder acht te slaan op pedagogisch-didactische bedoelingen van de interventie. Tevens valt op dat de opbrengsten in de experimentele groep zich bij een redelijke mate van implementatie positief onderscheiden van de opbrengsten in de controlegroep.

Kern van de essentiële werkingsmechanismen

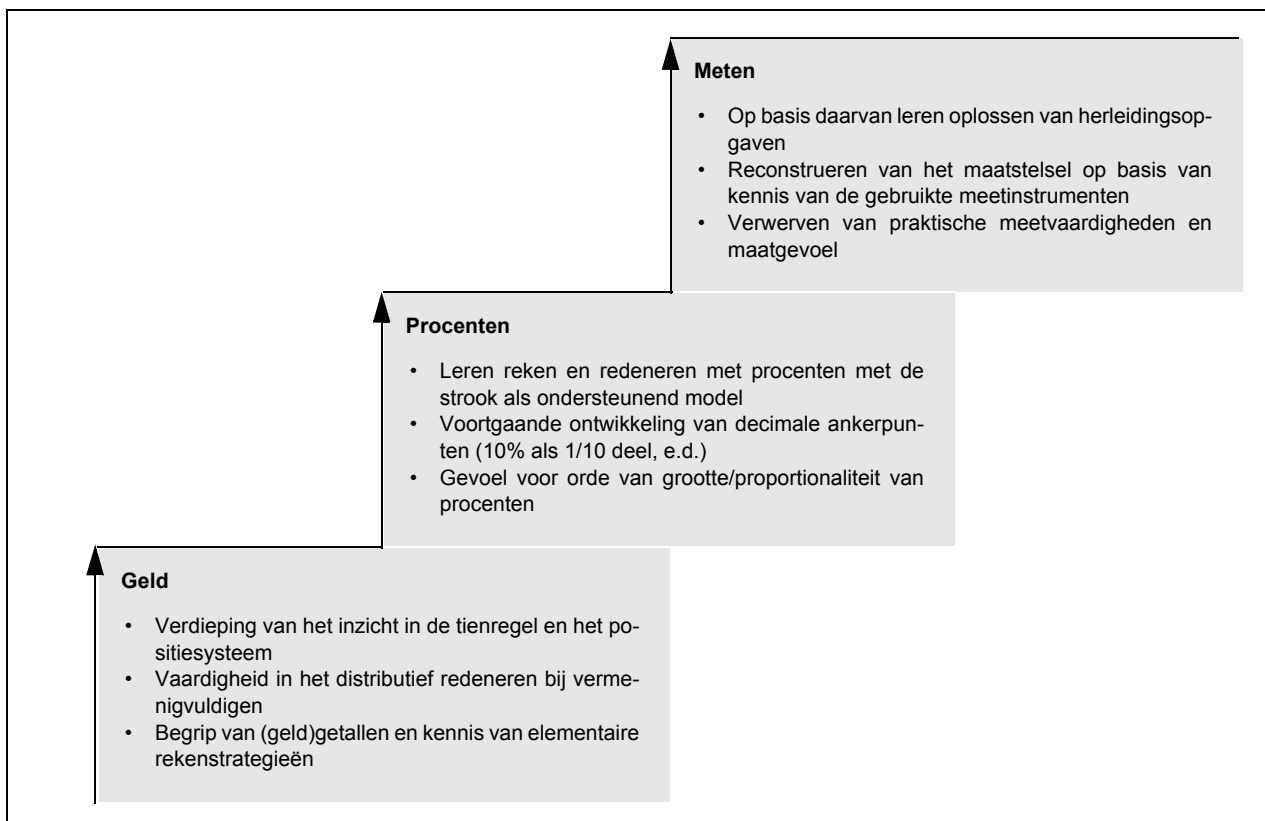
Uit de statistische analyses van de toetsgegevens kwam ook naar voren dat het vooral het domein 'procenten' was waarop zich in de experimentele groep een substantiële progressie ten opzichte van de controlegroep voordeed. Figuur 2 toont een overzicht van deze progressie. Bij de domeinen 'geld' en 'meten' was minder sprake van een duidelijke vooruitgang van de experimentele groep ten opzichte van de controlegroep. In principe is deze eenzijdigheid grotendeels te verklaren uit het feit dat de activi-

teiten rond geld er vooral op gericht waren om de leerlingen enig hernieuwd zelfvertrouwen alsmede beter notatiegedrag te laten ontwikkelen. Bij het meten daarentegen lag het accent voor een aanzienlijk deel op versterking van de praktische meetvaardigheid en het aankweken van maatgevoel, hetgeen moeilijk in een toets is te vangen. De lesobservaties en registraties van individuele begeleidingsmomenten van leerlingen wijzen erop dat de leerlingen op deze gebieden wel degelijk een zekere vooruitgang boekten. Voor het identificeren van de essentiële werkingsmechanismen werd daarom vooral naar de longitudinale opbouw van het traject gekeken, waarbij gezocht werd naar die aspecten die op de lange termijn het meest aan de progressie lijken te hebben bijgedragen. Op basis van een analyse van gegevens uit lesobservaties, leerlingmateriaal en interviews met leerlingen kon worden vastgesteld dat de kern van deze werkingsmechanismen kan worden voorgesteld als zijnde opgebouwd uit drie lagen waarvan de eerste laag betrekking heeft op kerninzichten en vaardigheden die binnen het domein 'geld' tot ontwikkeling kwamen, de tweede laag op kerninzichten en vaardigheden binnen het domein 'procenten' zelf, en de derde laag op zaken die in het verlengde daarvan liggen dan wel die specifiek betrekking hebben op het 'meten'.

Figuur 3 geeft hiervan een overzicht. Het lijkt aannemelijk dat naarmate docenten zich beter bewust werden van deze werkzame bestanddelen en naarmate ze meer bedre-



figuur 2: overzicht van de totaalscore bij de drie toetsen op de vier scholen samen bij het domein procenten ($n = 219$) tijdens de tweede pilot. De stijging in de experimentele groep is significant hoger dan die in de controlegroep. De getallen langs de verticale as verwijzen weer naar het aantal goed gemaakte opgaven (op een totaal van elf)



figuur 3: overzicht van de voornaamste werkzame bestanddelen van de interventie. Voor elk domein zijn drie vakdidactische kernpunten geformuleerd die richting kunnen geven aan het beoogde onderwijs

venheid ontwikkelden in het tot z'n recht laten komen daarvan, de leerresultaten zowel wat betreft goedscores, strategiegebruik als notatiegedrag hoger zijn uitgevallen en dat de interventie dus effectiever was.

De noodzaak van het analyseren van uitvoeringsomstandigheden

Tijdens de eerste pilot doet zich op drie van de vier scholen een enigszins vergelijkbare trend in de ontwikkeling van de prestaties van de leerlingen voor in de zin dat er een geleidelijke progressie tot stand komt die in de experimentele groep groter is dan in de controlegroep. Echter, op de vierde school is sprake van een duidelijk afwijkende trend doordat er bij de tweede toets sprake is van een tamelijk spectaculaire achteruitgang in de experimentele groep terwijl de prestaties in de controlegroep min of meer stabiel blijven.

Het blijkt aanvankelijk moeilijk om deze achteruitgang, die zich afspeelt op een school waarvan geen gegevens over de wijze van uitvoering van de lessen werden verzameld, te verklaren. Pas nadat studenten van de Universiteit Leiden via interviews meer over de gang van zaken aan de weet zijn gekomen, blijkt er een duidelijke oorzaak te zijn. De docent die de lessen aanvankelijk verzorgde, is na drie weken overspannen geraakt.

Daarna hebben de lessen drie weken stil gelegen en zijn ze uiteindelijk overgenomen door een docent die daartoe

volgens de schoolleiding eigenlijk niet capabel was. Het lijkt aannemelijk dat de teruggang in prestaties met deze ongunstige uitvoeringsomstandigheden te maken heeft. In de tweede pilot doet zich op een andere school een soortgelijk verschijnsel voor doordat er vlak voor het begin van het onderwijsexperiment onenigheid ontstaat over de vraag wie de lessen zal uitvoeren. Al met al treden er in dergelijke situaties dus leereffecten (of juist niet) op die moeilijk aan de interventie zijn toe te schrijven. Het zijn de uitvoeringsomstandigheden die in hoge mate mede bepalend zijn voor de leereffecten. Het registreren en analyseren van die uitvoeringsomstandigheden in relatie tot waargenomen leereffecten lijkt van groot belang om mogelijke verbanden tussen interventie en leereffecten eenduidig te kunnen vaststellen.

5 Aandachtspunten voor evidence-based onderzoeksbenadering

Uit het voorgaande blijkt duidelijk dat het in het soort van onderzoek zoals bij vmR werd uitgevoerd, belangrijk is om, zoals Bosker stelt, zo goed mogelijk uitsluitel te krijgen over de effectiviteit van een ontwikkelde innovatie of interventie door systematisch te meten wat de leereffecten zijn en door deze te vergelijken met de leer-

effecten bij een controlegroep waar een meer gangbare onderwijsbenadering wordt gevolgd. Op deze wijze wordt een zekere garantie verkregen dat de interventie, indien deze min of meer naar behoren wordt uitgevoerd, niet alleen een positief effect op de leerprestaties als zodanig heeft, maar dat deze zich ook positief onderscheidt van de leereffecten in een groep leerlingen die de interventie niet heeft gehad. Naast aspecten van stijging in de leerprestaties zoals die in de toetsresultaten tot uitdrukking komen, spelen daarbij ook aspecten van verbetering van strategiegebruik en notatiegedrag en aspecten van toegenomen succeservaringen een rol. Via het uitgevoerde onderzoek komt voor de toekomstige gebruikers zodoende waardevolle informatie beschikbaar over de mogelijke resultaten die met de interventie behaald kunnen worden.

Uit het hierboven gememoreerde blijkt echter ook duidelijk dat er zich bij een dergelijk experiment grote verschillen kunnen voordoen tussen wat bij leerplanontwikkeling veelal wordt aangeduid als het beoogde leerplan en het uitgevoerde leerplan (Van den Akker, 2005; Van den Heuvel-Panhuizen, 2009). De leereffecten blijken zowel in de experimentele als in de controlegroep in hoge mate bepaald te worden door het feitelijke onderwijsleerproces zoals zich dat in de klas afspeelt, de specifieke omstandigheden waaronder de interventie wordt uitgevoerd, de mate van zorgvuldigheid waarmee docenten hun lessen voorbereiden. Het is derhalve van groot belang om naast de leereffecten ook in grote lijnen te registreren hoe het proces in de klas zich afspeelt, hoe lessen worden uitgevoerd, en, algemener, in hoeverre er sprake is van een redelijke mate van implementatie van de interventie.

Uit het voorgaande blijkt verder dat dit zo goed mogelijk monitoren van het proces in de klas niet alleen van belang is om de ontwikkeling van de leerprestaties in het juiste perspectief te kunnen plaatsen, maar ook nog om andere redenen. Door te registreren wat voor cruciale onderwijsmomenten zich in de klas voordoen, welke mislukkingen en successen (Freudenthal, 1991) optreden, en door na te gaan in hoeverre hypothesen omtrent beoogde didactische werkwijzen en leerdoelen correct blijken te zijn, komt informatie beschikbaar die van essentieel belang is voor verdere optimalisering van de interventie, voor het beter uitlijnen van het experimentele programma en voor het beter uitvoerbaar maken daarvan. Zodoende kan, in overeenstemming met wat Gravemeijer en Kirschner betogen, ook een steeds scherper beeld ontstaan van wat als de vakdidactische essentie van de interventie moet worden beschouwd, van de kern van essentiële werkingsmechanismen waar een docent zicht op moet hebben om de interventie adequaat te kunnen uitvoeren en deze aan de eigen onderwijssituatie te kunnen aanpassen. Pas als het onderzoek daarover voldoende duidelijkheid heeft gecreëerd, wordt het mogelijk de interventie op een door-

dachte manier in het leerplan van de eigen school in te passen en uit te voeren. Daarbij moge duidelijk zijn dat onderzoek naar de ‘productkant’ van de interventie (cq. de leereffecten) en onderzoek naar de proceskant (cq. naar de wijze waarop de interventie in de klas gestalte krijgt) elkaar geenszins in de weg hoeven te zitten. In tegendeel, waarnemingen omtrent de leereffecten van de interventie en waarnemingen omtrent het onderwijsleerproces dat zich in de klas afspeelt, vullen elkaar op een natuurlijke manier aan. Tezamen vormen deze een solide empirische grondslag voor het vaststellen van leereffecten, de relatie tussen mate van effectiviteit en mate van implementatie van de interventie, en het identificeren van de essentiële werkingsmechanismen van de interventie.

6 Besluit

Effectiviteit en bruikbaarheid als kwaliteitscriteria voor leerplanontwikkeling

In het voorgaande gaat het om een specifieke vorm van leerplanontwikkeling waarbij het doel de ontwikkeling van een prototype van een onderwijsleertraject is met een bijbehorende karakteristiek van de essentiële werkingsmechanismen - een vorm die meestal het microniveau van leerplanontwikkeling wordt genoemd. Deze vorm is zeker niet onbelangrijk. Het gaat om het verbeteren van de doorgaande leerlijn van primair naar voortgezet onderwijs voor een bepaalde doelgroep van leerlingen, zodat deze leerlingen veel minder dan thans het geval is, een discontinuïteit in hun leerproces ervaren (Van Streun (red.), 2008). Met het oog op het voorkomen van voortijdige uitval en schoolverlaten is dit een belangrijke problematiek.

Op basis van het hierboven weergegeven betoog kan geconcludeerd worden dat het bij deze vorm van leerplanontwikkeling zeker waardevol is om bewijs voor effectiviteit te genereren. Door met experimentele en controlegroepen te werken en door de leereffecten daarvan te vergelijken, kan tot op zekere hoogte uitsluitel verkregen worden over de vraag in hoeverre het experimentele traject tot betere (en andere) leeropbrengsten leidt. Zoals uit het voorgaande blijkt, lijkt het daarbij aan te bevelen om het begrip effectiviteit breed op te vatten. Niet alleen de toetsprestaties in de vorm van goedscores dienen in beschouwing genomen te worden, maar ook andere cruciale kennisaspecten, in het bijzonder strategiegebruik en notatiegedrag.

Voor leerplanontwikkeling is het verder van belang om bewijs te genereren voor de bruikbaarheid van een product. Dit betreft in de eerste plaats de praktische bruikbaarheid in de zin dat het ontwikkelde lesmateriaal in de

klas op een betrekkelijk eenvoudige manier te gebruiken is zonder dat de docent zich zeer uitgebreid hoeft voor te bereiden, allerlei hulpmiddelen en materialen nodig heeft, enzovoorts. In de voorbereidende fase van het onderzoek kan over deze zaken al een zekere mate van uitsluitel verkregen worden. Maar, zoals in het bovenstaande naar voren is gebracht, is er nog een ander aspect van bruikbaarheid.

Zoals Gravemeijer en Kirschner betogen, is het van groot belang dat een docent gelegenheid heeft om inzicht te verwerven in de kern van essentiële werkingsmechanismen die tot op zekere hoogte verklaren waarom de interventie werkt zoals zij werkt. Terecht stellen zij dat elke docent te maken heeft met een eigen onderwijssituatie en dat deze bovendien in de loop van de tijd veelal een eigen ‘onderwijsstiel’ heeft ontwikkeld op basis waarvan hij of zij instructie geeft, leerlingen begeleidt en hulp biedt. Pas als de docent zicht heeft op de genoemde werkzame bestanddelen, is hij of zij in staat om deze adequaat te gebruiken in de eigen onderwijssituatie en te integreren in de eigen wijze van lesgeven.

Om dit mogelijk te maken, dient onderzoeksmatige leerplanontwikkeling zich erop toe te leggen om deze werkingsmechanismen bloot te leggen en bewijs voor de geldigheid ervan te genereren. Omdat van tevoren over deze bestanddelen veelal slechts vermoedens bestaan, ligt hier een belangrijke onderzoekstaak in de zin dat systematisch aanwijzingen worden verzameld voor de geldigheid daarvan. Het gaat hier om een van de essenties van ontwerponderzoek die de praktische bruikbaarheid van onderzoeksresultaten in hoge mate ten goede komt.

Literatuur

Akker, J. van den (2005). Curriculum Development Re-invented: evolving Challenges for SLO. In: J. Letschert, (ed.). *Curriculum Development Re-invented*. Enschede: SLO, 16-32.

Akker, J. van den, K. Gravemeijer, S. McKenney & N. Nieveen (2006). Introducing educational design research. In: J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (eds.). *Educational Design Research*. London: Routledge.

Akker, J. van den (2009). Curriculum Design Research. In: T. Plomp & N. Nieveen (eds.), *An Introduction to Educational Design Research*. Enschede: SLO, 37-53.

Akker, J. van den, N. Nieveen & W. Kuiper (2013). Bruggen slaan tussen beleid, praktijk en wetenschap in curriculumontwikkeling en -onderzoek. *Pedagogische Studiën*, 89, 399-410.

Buijs, K. & P. van der Zwaard (2007). *Aandachtsgebieden voor een doorgaande lijn rekenen-wiskunde van PO naar VMBO*. Enschede: SLO.
http://www.slo.nl/downloads/archief/Aandachtsgebieden_20doorgaande_20lijn_20rekenen-wiskunde_20van_20po_20naar_20vmbo.pdf/download

Buijs, K., P. van Scherpenzeel, M. ten Voorde & P. van der Zwaard (2008). *Werken aan de doorlopende leerlijn rekenen-wiskunde van PO naar VO*. Enschede: SLO.

Buijs, K. (2009). Werken aan rekenvaardigheid in het vmbo. *Euclides*, 84(8), 281-285.

Buijs, K. & P. van der Zwaard (2011). *Lesmap 'Verder met Rekenen'*. Enschede: SLO.

Bosker, R.J. (2008a). Naar meer evidence-based onderwijs! *Pedagogische Studiën*, 85, 49-51.

Bosker, R.J. (2008b). Tripliek: een onnodig gecompliceerde voorstelling van zaken. *Pedagogische Studiën*, 85, 305-308.

Commissie Meijerink (2008). *Over de drempels met taal en rekenen* (drie rapporten). Enschede: SLO.

Freudenthal, H. (1991). *Revisiting Mathematics Education. China Lectures*. Dordrecht: Kluwer.

Gille, E., C. Loijens, J. Noijens & R. Zwitser (2010). *Resultaten PISA-2009. Praktische kennis en vaardigheden van 15-jarigen*. Arnhem: Cito.

Gravemeijer, K. & P. Cobb (2006). Design research from a learning design perspective. In: J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (eds.). *Educational Design Research*. London: Routledge.

Gravemeijer, K.P.E. & P.A. Kirschner (2007). Naar meer evidence-based onderwijs? *Pedagogische Studiën*, 84, 463-471.

Gravemeijer, K.P.E. & P.A. Kirschner (2008). Dupliek. Een te simpele voorstelling van zaken. *Pedagogische Studiën*, 85, 195-197.

Kelly, A.E. (2006). Quality criteria for design research: evidence and commitments. In: J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (eds.). *Educational Design Research*. London: Routledge.

Heuvel-Panhuizen, M. van den (2009). *Hoe rekent Nederland?* Utrecht: Freudenthal Instituut (oratie).

Janssen, J., F. van der Schoot & B. Hemker (2005). *Balans van het reken-wiskundeonderwijs aan het einde van de basisschool 4*. Arnhem: Cito.

Nes, F. van (2009). *Young Children's Spatial Structuring Ability and Emerging Number Sense*. Utrecht: Freudenthal Instituut (proefschrift).

Kraker, L.J. de (2009). *A Formative Evaluation of a Mathematics Program for Students with Special Needs in Pre-vocational Secondary Education: a Multilevel Perspective*. Leiden: Universiteit Leiden.

Nieveen, N. (2009). Formative Evaluation in Educational Design Research. In: T. Plomp & N. Nieveen (eds.). *An Introduction to Educational Design Research*. Enschede: SLO, 89-103.

OCW (2008). *Kwaliteitsagenda Voortgezet Onderwijs 2008-2011*. Den Haag: Ministerie van OCW.

Onderwijsraad (2006). *Naar meer evidence-based onderwijs*. Den Haag: Onderwijsraad.

Os, S. van (2010). *Verslag Case Study Verder met Rekenen in de praktijk*. Enschede: SLO (interne SLO-publicatie).

Plomp, T. (2009). Educational Design Research: an Introduction. In: T. Plomp & N. Nieveen (eds.). *An Introduction to Educational Design Research*. Enschede: SLO.

Plomp, W. (2010). *Evaluatie van het VMBO-programma Verder met Rekenen*. Leiden: Universiteit Leiden.

Shadish, W.R., T.D. Cook & D.T. Campbell (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston/New York: Houghton Mifflin Company.

Stals, R. (2010). *Evaluatie van het onderwijsprogramma 'Verder met Rekenen'*. Leiden: Universiteit Leiden.
Streun, A (red.) (2008). *Over de drempels met rekenen*. SLO:

Enschede.
Zwaart, P. van der (2005). *Kommagetallen in diverse BAVO-methoden*. (interne SLO-publicatie).

Recently, in the Netherlands there has been a discussion about the effectiveness of evidence based educational innovations. This discussion resulted from a report by the Onderwijsraad (National Educational Council) advocating the use of solid quantitative research methods based on randomized trials, in order to prove the effectiveness of innovative educational approaches and products. In response to this appeal, Gravemeijer and Kirschner objected to the use of such methods arguing that what is really needed, is a more process oriented research method aiming to identify the crucial educational mechanisms of an innovation, as well as a supporting educational theory. In this contribution it is argued that it is the very combination of research on effectiveness with process oriented research that can lead to effective and useful innovations. To underpin this point of view, a case of curriculum development is described that aims at developing a learning trajectory relating to arithmetical proficiency of vmbo-students (prevocational secondary education).