

Bruggen slaan tussen beleid, praktijk en wetenschap in curriculumontwikkeling en -onderzoek

J. van den Akker, W. Kuiper en N. Nieveen

Samenvatting

Curriculumontwikkeling is in principe gebaat bij kennis uit onderzoek, maar aan de onderzoeksmatige onderbouwing van curriculumbeleid valt nog veel te verbeteren. Sociaal-politieke en pragmatische argumenten domineren vaak de besluitvorming. In dit artikel wordt een pleidooi gevoerd voor het adequaat benutten van onderzoek binnen een integrale benadering van curriculumontwikkeling. Daartoe worden drie luiken geïntroduceerd. Het uiteindelijke ontwikkelwerk (tweede luik) wordt vooraf gegaan door monitoring en (trend)analyse van informatie uit onderwijspraktijk, wetenschap en samenleving. Bevindingen daaruit leiden via debat en overleg tot verbeter- en innovatieprojecten (tweede luik). Door bij dat overleg en debat relevante ketenpartners te betrekken wordt tevens een basis gelegd voor implementatie en opschaling van die projecten (derde luik). Tijdens het tweede luik wordt via een cyclisch proces van analyse, ontwerp en evaluatie de kwaliteit van de curriculuminnovatie geoptimaliseerd en wordt bewijs verzameld voor de kwaliteit ervan.

1 Inleiding

Non satis scire – to know is not enough. Dat motto van de AERA-conferentie in 2012 (Vancouver) is ook goed van toepassing op onze bijdrage aan dit themanummer. Kennis uit (wetenschappelijk) onderzoek leidt immers niet vanzelf tot (meer) succesvolle curriculumontwikkeling. Curriculumontwikkeling is zeker gebaat bij kennis uit onderzoek, maar de onderlinge verbindingen laten dikwijls veel te wensen over. Onderwijsbeleid in het algemeen en curriculumbeleid in het bijzonder blinken van oudsher immers (wereldwijd) niet bepaald uit door krachtige onderzoekmatige onderbouwing. Sociaal-politieke en pragmatische argumenten domineren vaak de besluitvorming. Prioriteiten voor curricu-

lumprojecten berusten slechts zelden op systematische monitoring en analyse van praktijken en uitkomsten. Anderzijds blijkt beschikbare kennis uit onderzoek veelal weinig relevant voor het nemen van ontwerpbeslissingen. Bovendien ontbreekt dikwijls empirische informatie over feitelijke implementatie en invloed van curriculuminnovaties. Kortom, zowel voorafgaand, tijdens als na afloop van ontwikkelwerk is de rol van onderzoek beperkt. Men mag dan ook concluderen dat leerplanontwikkeling – wij hanteren de termen curriculum en leerplan als uitwisselbaar – nauwelijks een *evidence-based* onderneming is. Hoewel dat in strikte vorm ook moeilijk denkbaar is gezien het waarden geladen en normatief getinte karakter van curriculumontwikkeling (Walker, 2003; vgl. ook Biesta, 2012, hoewel die zich niet expliciet uitlaat in curriculumtermen), zouden velen echter toch graag tenminste een meer *evidence-informed* of *research-informed* benadering op prijs stellen, dat wil zeggen dat de keuzes in ontwikkelwerk waar mogelijk door onderzoek onderbouwd worden. In deze bijdrage zullen wij uiteenzetten hoe een betere kruisbestuiving tussen onderwijsonderzoek en leerplanontwikkeling de informatiebasis voor (verbetering van) curriculumbeleid en schoolpraktijk kan versterken (vgl. Van den Akker, 2010). De rol van onderzoek zal geschetst worden voor de verschillende stadia van leerplanontwikkeling, met bijzondere aandacht voor: kwaliteitscriteria voor curriculumproducten; adequate onderzoeksmethoden; kennisgroei; en professionele ontwikkeling van alle deelnemers. Deze uiteenzetting wordt vooral gevoerd door ontwikkel- en onderzoekservaringen bij SLO en geïllustreerd met enkele recente voorbeelden.

SLO (Stichting Leerplanontwikkeling, opgericht in 1975) heeft als nationaal expertisecentrum leerplanontwikkeling een publieke, tweedelijns taakstelling binnen de driehoek beleid, praktijk en wetenschap. De

werkzaamheden omvatten onder meer:

- het ontwerpen en valideren van landelijke leerplankaders;
- het ontwerpen en in de praktijk beproeven van exemplarische leerplanconcretisering op school-, vak- en klasniveau;
- wetenschappelijke onderbouwing van (inclusief internationale oriëntatie op) leerplanontwikkeling.

SLO streeft naar een integrale benadering van leerplanvraagstukken teneinde relevante, consistente, bruikbare, effectieve, schaalbare en duurzame curriculumopbrengsten te realiseren (zie par. 2). Die aanpak wordt gekenmerkt door het volgende drieluik:

- Feitelijke ontwikkelwerk wordt voorafgegaan door monitoring en (trend)analyse van informatie uit wetenschap, onderwijspraktijk en samenleving. Bevindingen daaruit leiden via overleg en debat tot agendering van curriculaire verbeter- en innovatieprojecten.
- Het ontwikkelwerk wordt bij voorkeur ingericht als ontwerponderzoek in proefsituaties, met veel aandacht voor formatieve evaluatie van tussentijdse opbrengsten.
- Van meet af aan wordt geanticipeerd op implementatie en opschaling van de ontwikkelopbrengsten door middel van samenwerking met relevante partners in met name opleiding, nascholing en begeleiding.

In de volgende paragrafen worden de drie fasen nader uitgewerkt. Daaraan voorafgaande staan we kort stil bij de relatie tussen onderzoek en leerplanontwikkeling.

2 De rol van onderzoek voor de kwaliteit van leerplanontwikkeling

In algemene zin geldt dat onderzoek voor het SLO-werk vooral een ondersteunende functie vervult om de kwaliteit van leerplanontwikkeling te bevorderen; ofwel, in termen van de Commissie de Graaf (2010), er is sprake van ‘ontwikkelingsondersteunend onderzoek’ of ‘optimaliseringsonderzoek’. Binnen het genoemde drieluik variëren de relaties tussen onderzoek en leerplanontwikkeling. Dat hangt samen met een evoluerend accent op de kwaliteit van het werk. Kwaliteit is een breed be-

grip en kan bij leerplanontwikkeling worden verhelderd door de volgende differentiatie in kwaliteitscriteria:

- *Relevantie*: wordt het belang van de beoogde leerplanontwikkeling breed onderschreven door betrokken partijen in beleid en praktijk, en is dat belang ook te onderbouwen op basis van wat er bekend is aan onderzoeksbevindingen ter zake?
- *Consistentie*: is het (voorlopige) curriculumontwerp adequaat uitgewerkt in samenhangende keuzes die (naar het oordeel van experts) ook stroken met beschikbaar wetenschappelijk onderzoek?
- *Praktische bruikbaarheid*: is het curriculumontwerp tijdens formatieve evaluatie in proefsituaties ook praktisch uitvoerbaar (voor met name leraren)?
- *Effectiviteit*: levert het curriculumontwerp tijdens empirische beproeving de gewenste opbrengsten op?
- *Schaalbaarheid*: in hoeverre is het curriculumontwerp ook bruikbaar en effectief in situaties buiten de kring van scholen en leraren die bij het ontwerp en de formatieve evaluatie betrokken waren?
- *Duurzaamheid*: in welke mate zijn de beoogde curriculumverbeteringen of -vernieuwingen ook op langere termijn bruikbaar en effectief?

Het zal duidelijk zijn dat deze zes kwaliteitscriteria op elkaar voortbouwen. Positieve antwoorden op eerdere vragen zijn voorwaardelijk voor de zinvolheid van de latere vragen. De Relevantie-vraag staat voorop in het eerste luik: Monitoring & Analyse. De aspecten Consistentie en Praktische Bruikbaarheid staan centraal tijdens het hoofdruik van Ontwikkeling. De laatste drie aspecten (Effectiviteit, Schaalbaarheid en Duurzaamheid) behoren vooral tot het derde luik (Implementatie en Opschaling). Elk van deze fasen en criteria vergen andersoortige vormen van onderzoek. In de volgende paragrafen wordt een en ander toegelicht.

3 De rol van onderzoek in het eerste luik: Monitoring & Analyse

Zoals eerder is aangegeven, berusten prioriteiten voor curriculumprojecten zelden op

een systematische monitoring en analyse van praktijken en uitkomsten. De hier bepleite integrale benadering van curriculumontwikkeling beoogt in deze tekortkoming te voorzien door het feitelijke ontwikkelwerk (tweede luik) vooraf te laten gaan door monitoring en (trend)analyse van informatie uit onderwijspraktijk, wetenschap en samenleving.

Kenmerkend is de brede focus, gegeven het uiteindelijke oogmerk: het maken van beargumenteerde keuzes omtrent ter hand te nemen verbeter- en innovatieprojecten. Daar waar het trends in de onderwijspraktijk betreft gaat het niet om alleen een analyse van onderwijsopbrengsten (vergelijk de accenten in PPO van Cito, in Inspectie-analyses en in veel effectiviteitsonderzoek) of om alleen een analyse van de onderwijspraktijk (het accent in veel vakdidactische analyses). Wat voorop staat, is inzicht in en duiding van de relatie c.q. spanning tussen opbrengsten, praktijken en intenties, ook wel aangeduid als het bereikte, uitgevoerde resp. beoogde curriculum (Goodlad *et al.*, 1979; zie ook Thijs & Van den Akker, 2009). Kennis van ervaringen, percepties, motieven, knelpunten en verbeterwensen van de kant van leraren, schoolleiders, leerlingen en allerlei andere systeempartners is daarbij onmisbaar (Fullan, 2007). En vaak ook helpt het zaken in internationaal perspectief te bezien en vanuit enig historisch besef te duiden. Richtinggevend bij het zicht krijgen op trends in wetenschap en samenleving zijn vragen als welke maatschappelijke en wetenschappelijke ontwikkelingen aanleiding geven tot herziening van leerdoelen en -inhouden en hoe actoren prioriteiten voor leren en onderwijzen wegen.

Bij de brede focus hoort ook het onderscheid in *generieke* en *vakspecifieke* monitoring en analyse. Het eerste richt zich op algemene leerplankundige thema's (bv. inhoudelijk maatwerk bieden aan leerlingen of curriculaire competenties van leraren) en niet op één vak, vakkencluster of leergebied. Bij vakspecifieke monitoring en analyse is dat laatste juist wel het geval. SLO is eind 2011 begonnen met het per vakkencluster (bijv. natuur & techniek) systematisch op orde brengen, ontsluiten en duiden van de grote hoeveelheid informatie over beleid, schoolpraktijk en leeropbrengsten. Dat gebeurt in

de vorm van 'digitale vakportalen'. Die zijn niet alleen bedoeld als een dienst aan leraren en lerarenopleidingen, maar ook – en dat is het punt waar het hier om draait – als input voor vakspecifieke trendanalyse (met mogelijke implicaties voor agendabepaling ten aanzien van initiatieven voor vakvernieuwing).

Kenmerkend is tevens de fasering in activiteiten. De blik richt zich:

- eerst op de *signalering* van trends, ontwikkelingen en/of problemen in praktijk, beleid, wetenschap en maatschappij in combinatie met het maken van een inschatting van de mogelijke betekenisvolheid van die signalen voor leerplanontwikkeling, in generieke of vakspecifieke zin;
- dan (soms) op een *probleemanalyse* waartoe wordt overgegaan als een meer indringende analyse en duiding van redenen, kansen en bedreigingen voor leerplanvernieuwing, generiek dan wel vakspecifiek, gewenst is;
- en uiteindelijk op een *agendabepaling* waarin beargumenteerde keuzes worden gemaakt omtrent ter hand te nemen verbeter- en innovatieprojecten.

Binnen deze fases bestaan monitoring- en analyseactiviteiten uit een mix van:

- het kennisnemen van relevante (inter)nationale bronnen, publicaties en websites alsmede het bezoeken van relevante bijeenkomsten en onderzoekcongressen (*signalering* en *probleemanalyse*).
- het systematisch en op basis van een variëteit aan methoden analyseren en synthetiseren – en soms ook eerst zelf verzamelen – van gegevens over opbrengsten, praktijken en/of intenties (*probleemanalyse*).
- het organiseren en reguleren van 'conversatie en debat' (Cras, 2009), zowel publiek als professioneel, over de resultaten van een uitgevoerde probleemanalyse en de implicaties daarvan voor te maken keuzes (*agendabepaling*).
- het participeren in strategisch relevante relatienetwerken (*signalering*, *probleemanalyse* en *agendabepaling*).

Een laatste kenmerk van monitoring en trendanalyse is het proportionele karakter

ervan. Systematische monitoring en analyse van alle mogelijk denkbare generieke en vakspecifieke thema's en aspecten is – alleen al uit kostenoverwegingen – onhaalbaar. Ook op dit punt is prioritering noodzakelijk en dienen activiteiten op opportuniteit te worden beoordeeld.

Het systematisch analyseren en synthetiseren van gegevens over opbrengsten, praktijken en/of intenties illustreren we aan de hand van de leerplankundige analyse van PISA-trends die SLO in opdracht van het ministerie van OCW heeft uitgevoerd (Kuiper *et al.*, 2010). Aanleiding tot deze analyse was de dalende lijn – zowel absoluut (gemiddelde scores) als relatief (positie op de internationale ranglijst) – in prestaties van Nederlandse leerlingen voor wiskunde en deels ook voor leesvaardigheid en natuurwetenschappen bij PISA in de periode 2003-2009. Deze volgens OCW zorgwekkende trend staat haaks op beleidsambities die mikken op een top-5 positie en vormde daarmee reden voor nadere analyse. Centraal stond de vraag hoe de dalende trend in prestaties van Nederlandse leerlingen bij de PISA-metingen in de periode 2003-2009 voor wiskunde, natuurwetenschappen en leesvaardigheid leerplankundig zou kunnen worden geduid. Bij de analyse is als kapstok de indeling in leerplankundige verschijningsvormen gehanteerd: beoogd, uitgevoerd en bereikt curriculum. Het ging daarbij om een duiding van mogelijke discrepanties tussen de prestaties van 15-jarigen in PISA 2003, 2006 en 2009 (bereikt), de feitelijke lesuitvoering voor de drie domeinen (uitgevoerd) en het huidige ambitieniveau zoals verwoord in 58 kerndoelen voor het primair onderwijs en 58 kerndoelen voor de onderbouw VO (beoogd). Met name ging het daarbij om de vraag hoe dat wat gevraagd en getoetst is bij PISA zich verhoudt tot de kerndoelen, de meest gebruikte methoden en de lespraktijk op basis van die methoden. De leerplankundige duiding op het niveau van het uitgevoerde en beoogde curriculum is uitgevoerd aan de hand van het curriculaire spinnenweb (Van den Akker, 2003; Thijs & Van den Akker, 2009). De kern en negen draden van dit spinnenweb verwijzen naar tien onderdelen van het curriculum: visie, leerdoelen, leerinhoud, leeractiviteiten,

docentrollen, bronnen en materialen, groepeeringsvormen, leeromgeving, tijd en toetsing. Idealiter zijn de onderdelen met elkaar verbonden en is er sprake van consistentie en samenhang. Teneinde de duiding van de dalende trend in prestaties van Nederlandse 15-jarigen betekenisvol te laten zijn, is gekeken naar de gehele fase van het funderend onderwijs.

De analyse liet onder andere zien dat er, kijkend naar wat bij PISA belangrijk wordt geacht, in Nederland bij wiskunde te weinig aandacht is voor mathematiseren, redeneren, reflectie en authentiek probleemoplossen, bij natuurwetenschappen voor het geven van natuurwetenschappelijke verklaringen en bij lezen voor volledig tekstbegrip. Opmerkelijk was ook dat de daling in gemiddelde wiskundescores vooral is toe te schrijven aan lagere prestaties van meisjes in alle opleidingstypen en dat er bij wiskunde ruimte voor prestatieverbetering is voor zowel de betere leerlingen (er zijn minder hoge scores) als de zwakkere. Ook bleek dat de ontwikkeling van de vakinhoudelijke en vakdidactische kwaliteit van leraren krachtige impulsen behoeft met het oog op realisering van meer diepgang. Een hoofdconclusie uit de analyse is dat meer ambitie in de vorm van een beter uitgelijnde, specifiekere uitgewerkte en breed gedragen inhoudelijke koers van het funderend onderwijs nodig is, daarbij het PISA-raamwerk als een nuttige spiegel voor doelen en inhoudskeuzes beschouwend.

Dat laatste – en dat ter illustrering van genoemde agendavorming – is voor het ministerie van OCW (mede) aanleiding geweest SLO in het verlengde van het Actieplan 'Beter presteren' (OCW, 2011) een vooronderzoek te laten uitvoeren naar de mogelijke functies, doelgroepen, vorm en inhoud van zogenaamde 'kennisbases', onder andere voor het natuurwetenschappelijke domein (*science*), voor de onderbouw van het voortgezet onderwijs (Van der Hoeven *et al.*, 2012). Een kennisbasis is een richtinggevend kader voor wat leerlingen aan het eind van de onderbouw moeten kennen aan vakinhouden en moeten beheersen aan (vakspecifieke) vaardigheden, werkwijzen en denkwijzen. Primair oogpunt van de kennisbases is niet een vernieuwing van het onderwijsaanbod,

maar een bijdrage aan (de politieke en professionele discussie over) een verbetering van de leeropbrengst in de onderbouw. Bij dat laatste gaat het om een verbetering in internationaal perspectief (bv. PISA) en met het oog op het onderwijs in de bovenbouw vmbo en havo/vwo. De veronderstelling is dat een zekere mate van concretisering en aanvulling/aanscherping van de huidige kern-doelen scholen en systeempartners kan ondersteunen bij het realiseren van die ambitie.

4 De rol van onderzoek in het tweede luik: Curriculair Ontwerp-onderzoek

De werkzaamheden van het eerste luik leiden tot agendering van relevante curriculaire verbeter- en innovatieprojecten die de kern vormen van het tweede luik. In een vaak langdurig en iteratief proces worden wensen en idealen in een cyclisch proces van analyse, ontwerp, ontwikkeling, implementatie en evaluatie uitgewerkt tot tastbare resultaten in de praktijk Zo'n cyclische manier van werken draagt bij voorkeur de trekken van *ontwikkelingsonderzoek* of (curriculair) *ontwerponderzoek* (Van den Akker, 1999; Van den Akker *et al.*, 2006; Plomp & Nieveen, 2010; McKenney & Reeves, 2012). Bij ontwerponderzoek worden onderzoeksdraden ingeweven in het ontwikkelwerk teneinde allerlei onvermijdelijke onzekerheden in ontwerpkeuzes te verminderen en de robuustheid en kans van slagen van het leerplan-in-woording te optimaliseren. Daarnaast wordt via systematische reflectie ook kennisvermeerdering over succesvolle aanpakken nagestreefd. Het perspectief van ontwerponderzoek biedt vele mogelijkheden om de kwaliteit van leerplanontwikkeling en leerplanproducten te versterken. Dit krijgt met name gestalte door:

- analyse bij de start van het traject (vooronderzoek, beginsituatieonderzoek) waarmee het fundament wordt gelegd voor allerlei ontwerp- en implementatiekeuzes;
- frequente formatieve evaluatie tijdens het proces, gericht op het bepalen van de kwaliteit van tussenproducten en het genereren van verbeteringsvoorstellen.

We lichten beide activiteiten toe.

4.1 Analyse

Analyse, als geïntegreerd onderdeel van curriculaire ontwerponderzoek, bouwt vanzelfsprekend voort op (opbrengsten van) monitoractiviteiten uit het eerste luik en heeft een tweeledig doel:

- daar waar (nog) nodig meer specifiek inzicht krijgen in de bestaande situatie en in mogelijkheden voor verbetering en vernieuwing;
- specificeren van beoogde functies, inhoud en kenmerken van het ontwerp (ontwerpspecificaties) en de wijze waarop deze gerealiseerd kunnen worden (methodologische specificaties).

Belangrijke activiteiten daartoe zijn een analyse van de gebruikspraktijk (context- en behoefteanalyse) en een verkenning van de beschikbare kennis (literatuurstudie en expertbevraging). Het eerste is nodig om inzicht te krijgen in de bestaande praktijk en in de mogelijkheid en wenselijkheid van verbetering en vernieuwing van deze situatie. Een behoefteanalyse vraagt naar de visie van betrokkenen op de huidige situatie (wat werkt wel/niet en hoe ziet men de noodzaak voor vernieuwing) en op kenmerken van een meer wenselijke situatie. Een contextanalyse is erop gericht de probleemomgeving te verkennen en de ruimte voor vernieuwing in kaart te brengen. Veel gebruikte onderzoeksmethoden bij een behoeften- of contextanalyse zijn interviews, focusgroepen, lesobservaties en case studies.

Daarnaast is het van belang bestaande kennis te benutten voor het funderen van ontwerpbeslissingen. Centrale activiteiten daarbij zijn literatuurstudie, bevraging van deskundigen en analyse of evaluatie van bestaande projecten en producten. Vragen waar een dergelijke analyse zich op richt zijn:

- Welke (recente) inzichten uit onderwijs-onderzoek en vakwetenschap kunnen benut worden bij het ontwerpen?
- Welke reeds beschikbare (verwante, veelbelovende) leerplanproducten kunnen als inspiratiebron fungeren en welke lessen kunnen geleerd worden uit de invoering en het effect van deze producten?

Het vooronderzoek mondt uit in een gefundeerde beschrijving van ontwerprijlijnen,

in termen van genoemde ontwerpspecificaties en methodologische specificaties.

4.2 Formatieve evaluatie

Formatieve evaluatie is gericht op het bepalen en verbeteren van de kwaliteit van een tussenproduct. Het plannen van een formatieve evaluatie start met het formuleren van een duidelijke evaluatievraag. Naarmate de productontwikkeling vordert, verschuift de aandacht van de formatieve evaluatie veelal van vragen met betrekking tot de relevantie en consistentie van het tussenproduct naar vragen ten aanzien van de bruikbaarheid, effectiviteit, schaalbaarheid en duurzaamheid. De verschuiving van de vraag brengt ook andere evaluatiemethoden in beeld. We onderscheiden hier de volgende vijf evaluatiemethoden (cf. Nieveen, 2009):

- *screening*:
Het ontwikkelteam vergelijkt het ontwikkelde materiaal met een checklist van gewenste kenmerken van onderdelen van het ontwerp.
- *focusgroep*:
Respondenten (bijvoorbeeld vakexperts, docenten) geven een reactie op het ontwerp aan de hand van een richtlijn met punten waarop het ontwerp team commentaar en suggesties wenst.
- *walkthrough*:
Respondent(en) doorlopen samen met ontwikkelaar(s) het ontwerp.
- *micro-evaluatie*:
Een kleine groep gebruikers gebruikt een gedeelte van het ontwikkelde materiaal buiten de normale praktijksituatie.
- *try-out*:
Gebruikers zetten de ontwikkelde materialen binnen de beoogde praktijksituatie in.

Naarmate het product meer is uitgewerkt, verschuiven de methoden langzamerhand van screening naar try-out.

Vooraf in eerdere ontwikkelfasen is het voornaamste doel van een formatieve evaluatie het inventariseren van tekortkomingen en het genereren van ideeën voor verdere verbetering van het ontwerp. Voor dit doel zijn grote aantallen respondenten niet noodzakelijk. Triangulatie is van groter belang, dat wil zeggen het hanteren van complementaire

vormen van gegevensverzameling en betrekken van respondenten uit uiteenlopende contexten.

Bij het plannen van een evaluatie is het ook van belang stil te staan bij de rol die de ontwikkelaars zelf vervullen tijdens de evaluatie. In latere ontwikkelfasen is het wenselijk dat externe evaluatoren, en niet de oorspronkelijke ontwerpers, aan de slag gaan om de werkelijke effectiviteit van het leerplan vast te stellen. In eerdere ontwikkelfasen lijkt het om twee redenen legitiem en zelfs wenselijk om de evaluatieactiviteiten wel door de ontwikkelaars zelf te laten uitvoeren. Het vergroot de kans dat de uitkomsten van evaluatieactiviteiten daadwerkelijk leiden tot aanpassing van het leerplan. De evaluatie kan op elk door het ontwikkelteam gewenste moment worden uitgevoerd en de resultaten kunnen snel worden benut bij het reviseren van het product. Daarnaast leren ontwikkelaars zelf ook veel van de evaluatieactiviteiten. Ze zien bijvoorbeeld met eigen ogen wat er mis kan gaan tijdens het gebruik van het product in de praktijk. Echter, ontwikkelaars moeten zich er wel van bewust van zijn dat zij zozeer gehecht kunnen raken aan hun eigen ontwerp dat ze eventuele problemen met het ontwerp niet meer objectief kunnen beoordelen. In dat geval komt de keuze voor externe evaluatoren in beeld. In alle gevallen is een kritisch-constructieve distantie ten opzichte van het nut, de noodzaak en de aard van het te evalueren object geboden.

Een voorbeeld van vervaechting van analyse, ontwikkeling en formatieve evaluatie is een project over voorbereidend wetenschappelijk Engels in vwo 5 en 6 (Beeker, 2009). Aanleiding tot het project waren aansluitingsproblemen van eerstejaarsstudenten met Engels. Doel van het project was scherper te krijgen waaruit de aansluitingsproblematiek bestond, de oorzaken daarvan te achterhalen en vervolgens een mogelijke oplossingsrichting uit te werken. Voor de analyse zijn drie focusgroeptijdenkomsten georganiseerd met universitaire en hogeschooldocenten, met eerstejaars studenten en met vwo 6-leerlingen. Daarnaast is een vragenlijstonderzoek uitgevoerd onder eerstejaarsstudenten. Een van de resultaten was dat de leesvaardigheid de meeste problemen gaf, met name door de

hoeveelheid tekst en de abstracte aard van de meeste studieboeken. Teneinde de aansluiting te verbeteren is met drie pilotscholen intensief overlegd over praktische lesideeën voor voorbereidend wetenschappelijk Engels in vwo 5 en 6. Daartoe zijn als ontwerprichtlijnen gehanteerd: lesmodules zouden bij een schoolexamen gebruikt moeten kunnen worden; het materiaal zou leerlingen moeten voorbereiden op het soort Engels dat ze op de universiteit krijgen; het materiaal zou in ieder geval moeten leiden tot verbetering van de lees- en schrijfvaardigheid; en het materiaal zou makkelijk toegankelijk moeten zijn. Bij het ontwerpen van de lesmodules is uitgegaan van het idee van *scaffolding*. De taakten waaraan de leerlingen werken liggen steeds op een hoger taalniveau dan waarover ze beschikken. De docent reikt vervolgens via een vast stappenplan *scaffolds* (steigers) aan bedoeld om de leerlingen naar een hoger niveau te brengen. De lesmodules zijn via interviews met deskundigen en via focusgroepenbijeenkomsten met docenten Engels van de pilotscholen op consistentie geëvalueerd. Aangepaste versies zijn vervolgens via try-outs met observaties en interviews in de lessen beproefd op bruikbaarheid. Indicaties voor de effectiviteit zijn verkregen via een voor- en nameting bij de leerlingen.

5 De rol van onderzoek in het derde luik: Implementatie en Opschaling

Om tot relevante en kwalitatief hoogwaardig leerplanproducten te komen is het verweven van onderzoek in het eerste en tweede luik van groot belang. Vanuit zorg over het regelmatig falen van (grootschalige) vernieuwingstrajecten (cf. Fullan, 2000, 2007; ook in Nederland onderstreept door de Commissie Dijsselbloem, 2008), is de aandacht voor onderzoek naar de effectiviteit van (leerplan)vernieuwing in het onderwijs groeiende. Eerder al pleitte de Onderwijsraad (2006) voor een meer *evidence-based* benadering van onderwijsontwikkeling, waarin door middel van gerichte evaluatie gezocht wordt naar empirisch bewijs van de effectiviteit van de interventie. Geconstateerd moet echter

worden dat over de feitelijke implementatie en de effectiviteit van curriculuminnovaties in de praktijk maar weinig empirische gegevens worden verzameld. Voordat we ingaan op de wijze waarop onderzoek een duidelijke plaats kan krijgen in dit laatste luik, bespreken we eerst kort de belangrijkste redenen voor het veelal ontbreken van empirische gegevens over de effectiviteit van de implementatie en opschaling van curriculumvernieuwingen.

Een belangrijke reden is de complexiteit van curriculumimplementatie en -opshaling zelf. Iedere vernieuwing kent een gelaagdheid met geneste structuren (Van den Akker, 2003; Coburn, 2003). Om uiteindelijk de leerlingen met een vernieuwing te bereiken en effect te sorteren zullen school, locatie, team en leraren met de vernieuwing aan de slag moeten gaan (McLaughlin & Mitra, 2001). Naast deze gelaagdheid moet rekening gehouden worden met het gegeven dat implementatiecondities van school tot school, maar ook van leraar tot leraar kunnen verschillen. Dit komt deels door de ruimte die (zeker in Nederland) aan scholen wordt geboden om eigen curriculumbeleid te voeren, waardoor scholen verschillende vernieuwingsprioriteiten kunnen stellen (Kuiper *et al.*, 2007; Nieveen, Van den Akker & Resink, 2010), maar ook door diversiteit in achtergrond van leerlingen, ervaringen van leraren, ondersteuning van de kant van de schoolleiding, toegang tot materialen en ondersteuning (Penuel & Fishman, 2012). Bovendien is doorgaans een groot aantal intermediaire partners betrokken bij de uitwerking van leerplanbeleid, elk vanuit een eigen verantwoordelijkheid. De lijst van mogelijke betrokkenen is lang. Naast beleidsontwikkelaars (waaronder OCW), zijn dat toets- en examenontwikkelaars, schoolexterne belangbehartigers, professionals binnen de schoolcontext, lerarenopleidingen en nascholingsinstituten, leermiddelenontwikkelaars, onderzoekers, schoolexterne ondersteuners, de media en de onderwijsinspectie. Figuur 1 geeft de relatie weer tussen activiteiten van deze systeempartners vanuit het perspectief van leerplanbeleid. Deze intermediairs kunnen echter uiteenlopende beelden hebben van het leerplanbeleid zelf, hetgeen zijn weerslag

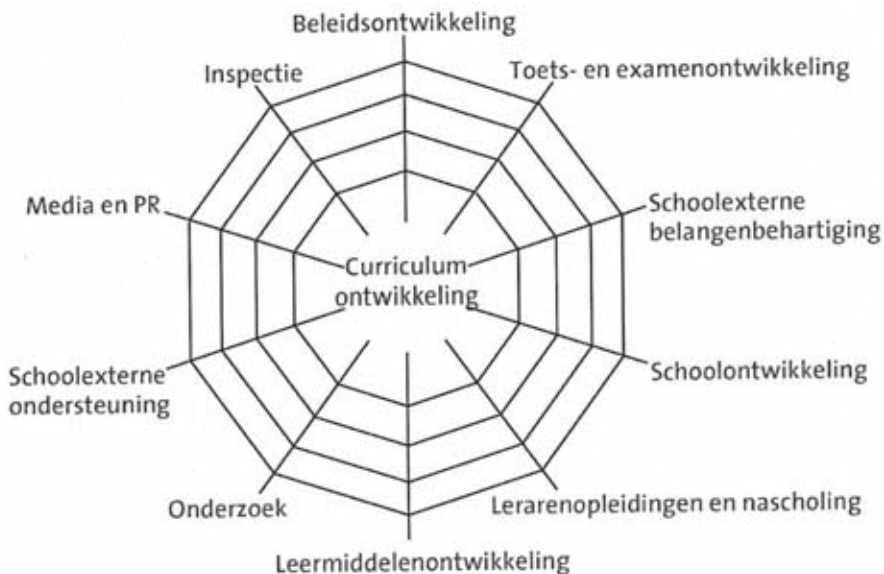
heeft op het type steun dat ze bieden naar scholen en leraren. Interactie tussen alle betrokkenen is van vitaal belang, maar is tevens een complexe aangelegenheid (Marsh, 2009; Nieveen *et al.*, 2010). Daar komt bij dat voor een echte verandering veelal meerdere jaren uitgetrokken moeten worden wil deze op duurzame wijze gestalte krijgen (Elmore, 1996; Hargreaves & Fink, 2006).

Deze kenschets van curriculuminnovaties leidt ertoe dat het op (quasi-)experimentele wijze onderzoeken van effecten van curriculuminnovaties uiterst gecompliceerd is. Vanwege de gelaagdheid en diversiteit van condities waaronder curriculumimplementatie plaatsvindt, is het zeer lastig (en kostbaar) uiteenlopende situaties te organiseren en onder controle te houden. Onderzoekers krijgen in de praktijk vaak te maken met bijvoorbeeld onvolledige implementatie, kruisbestuiving (bijv. experimentele groep beïnvloedt de controlegroep) en het verwateren van de interventie (Shadish, Cook & Campbell, 2002; Coburn, 2003). Zie voor een uitgebreidere bespreking van mogelijke bezwaren die kleven aan *randomized controlled trials* bij onderwijsonderzoek ook de discussiebijdragen van Bosker (2008a), Grave-meijer en Kirschner (2008) en Bosker (2008b).

Een mogelijkheid om uit deze impasse te

raken is het stapelen van bewijs voor kwaliteit van de onderwijsinnovatie. In haar advies 'Naar meer *evidence-based* onderwijs' (2006) stelt de Onderwijsraad een gefaseerde aanpak voor. Afhankelijk van de stand van kennisontwikkeling op een bepaald gebied, zijn verschillende gradaties van bewijsvoering mogelijk. De verschillende manieren van bewijsvoering vormen samen een oplopende schaal. Aan de linkerkant liggen weinig bouwstenen voor bewijsvoering (zacht bewijs), aan de rechterkant stapelen de bouwstenen zich op (hard bewijs). De hardste vorm bewijs met veel bouwstenen zijn gecontroleerde experimenten met aselechte toewijzing van proefpersonen aan condities.

Zoals betoogd is evaluatie van groot belang om tot een kwalitatief hoogwaardige onderwijsinnovatie te komen. In de eerste plaats (tijdens het tweede luik) via formatieve evaluatie gericht op het bepalen van de kwaliteit van tussenproducten en het genereren van verbetervoorstellen. Door daar stevig in te investeren valt er veel winst te boeken. Daarnaast, wanneer de tijd rijp is, via summatieve evaluatie aan het eind van het traject, gericht op het vaststellen van de effectiviteit van het eindproduct. De vraag is echter gerechtvaardigd of het leveren van hard bewijs middels (quasi-)experimenteel onderzoek altijd noodzakelijk is. Als er voldoende zacht bewijs is



Figuur 1 Systeemweb met ketenpartners vanuit curriculumperspectief

gestapeld in de loop van systematisch uitgevoerd ontwerponderzoek, dan is dat mogelijk-kerwijs voldoende. In dat geval kan worden volstaan met het volgen (monitoren) van de implementatie van de vernieuwing via surveys gecombineerd met case studies. Daarbij wordt nagegaan in hoeverre er met de vernieuwing vorderingen worden gemaakt ten opzichte van de oorspronkelijke situatie. Een duidelijk voorbeeld daarvan is de monitoring van de invoering van de nieuwe examenprogramma's en de daarbij behorende syllabi voor biologie, natuurkunde en scheikunde voor havo en vwo. Dit door SLO in opdracht van het ministerie van OCW uit te voeren monitoringonderzoek richt zich op de vraag in hoeverre de beoogde bètavakvernieuwing door docenten en leerlingen geïmplementeerd en gerealiseerd wordt. Het onderzoek start met een nulmeting aan het begin van het schooljaar 2012-2013 en eindigt met een meting aan het eind van het schooljaar 2017-2018. Dit onderzoek ligt in het verlengde van de evaluatie van de examenpilots natuurkunde, scheikunde en biologie en van de ontwikkeling en invoering van het profielkeuzevak 'Natuur, Leven en Technologie' in havo en vwo in de periode 2007-2010 (Kuiper *et al.*, 2011). Als toch gekozen wordt voor (quasi-)experimenteel onderzoek naar effecten van vernieuwing – bijvoorbeeld omdat er sprake is van betekenisvolle meningsverschillen over de verwachte effectiviteit van verschillende curriculumvarianten – dan dient de uitvoering daarvan bij voorkeur te liggen bij externe onderzoeksinstituten.

6 Afstemming tussen diverse vormen van onderzoek

In dit artikel hebben we een lans gebroken voor het adequaat inzetten en benutten van onderzoek binnen een integrale benadering van curriculumontwikkeling en hebben daartoe drie luiken geïntroduceerd. Bij de integrale benadering wordt het uiteindelijke ontwikkelwerk (tweede luik) vooraf gegaan door monitoring en (trend)analyse van informatie uit onderwijspraktijk, wetenschap en samenleving. Bevindingen daaruit leiden via debat en overleg tot agendering van de ver-

beter- en innovatieprojecten (tweede luik). Door bij dat overleg en debat relevante ketenpartners te betrekken wordt tevens een basis gelegd voor implementatie en opschaling van die projecten (derde luik). Tijdens het tweede luik wordt via een cyclisch proces van analyse, ontwerp en evaluatie de kwaliteit van de curriculuminnovatie geoptimaliseerd en wordt bewijs verzameld voor de kwaliteit ervan. Door deze bewijzen op systematische wijze te stapelen en er transparant over te communiceren wordt de noodzaak verkleind voor het uitvoeren van gecompliceerd (quasi-)experimenteel onderzoek in het derde luik. Aan onderzoeksinstellingen die de handschoen willen oppakken van het uitvoeren van (quasi-)experimenteel effectonderzoek naar curriculuminnovaties adviseren wij dat type effectonderzoek te laten aansluiten bij via ontwerponderzoek uitgerijpte interventies waarvan voorlopig bewijs voor kwaliteit op tafel ligt. Die suggestie gaat dan wel gepaard met een belangrijke toevoeging van Coburn (2003, p. 9): “...develop research designs that capture what is important rather than only what is easily measurable”. De validiteit van het onderzoeksinstrumentarium is immers bij curriculumonderzoek cruciaal.

De kans op implementatie en opschaling wordt vergroot door al tijdens het eerste en tweede luik te anticiperen op de complexiteit van implementatie en opschaling van curriculuminnovaties. Dat betekent ten eerste dat van meet af aan een systeembenadering gehanteerd dient te worden. Daarbij is het essentieel dat via proactieve communicatie verschillende systeempartners elkaars perspectieven, waarden en motieven (onder)kennen en hun werkzaamheden bewust op elkaar afstemmen (Fullan, 2007; Pareja Roblin *et al.*, 2012). Ten tweede dient tijdens het tweede luik (waar de ontwikkelactiviteiten plaatsvinden) geanticipeerd te worden op het feit dat leraren genoodzaakt zullen zijn de ontwikkelde materialen aan te passen aan hun lokale situatie (Coburn, 2003; Penuel & Fishman, 2012). Dat betekent de ontwikkeling van flexibele materialen en/of de ontwikkeling van meerdere uitwerkingen van een vernieuwing (Elmore, 1996). Parallel aan de vernieuwing dient er aanhoudende aandacht te zijn voor professionalisering van leraren en

andere systeembetrokkenen. Dat kan door te informeren (bijeenkomsten, nieuwsbrieven, publicaties, workshops), te scholen en te adviseren (informeel en formeel), mee te laten ontwikkelen (pilots, expertmeetings) en samen te werken in de vorm kennisnetwerken (De Vries & Pieters, 2005), professionele leergemeenschappen (o.a. Lieberman & Miller, 2008) dan wel docentontwerpteam (Handelzalts, 2010). Naast vakinhoudelijke en pedagogisch-didactische scholing is het vanwege de herontwerpplanning die leraren maken met het ontwikkelde materiaal van belang dat ook professionalisering plaatsvindt op het punt van curriculumontwerp (Nieveen *et al.*, 2010; Nieveen & Van der Hoeven, 2011). Of zoals Penuel en Fishman het verwoorden (2012, p. 295): “*professional development that prepares teachers to adapt high-quality materials*”.

Terugkomend op de openingszin van dit artikel (*Non satis scire*), kunnen we ter afsluiting concluderen dat op het vlak van leerplanontwikkeling (onderzoeksmatig verworven) kennis *sec* niet vanzelfsprekend en probleemloos leidt tot verbeteringen. Er zullen doelbewust voortdurend vele bruggen gebouwd moeten worden. Tussen beleid, praktijk en wetenschap, tussen vele soorten spelers en partijen op diverse niveaus, tussen uiteenlopende perspectieven en belangen. Dan nog zullen normatieve argumenten en contextuele bijzonderheden immer een prominente rol blijven spelen, maar onderzoek vergroot onmiskenbaar de kans op een transparante onderbouwing van keuzes in curriculaire besluitvorming en professioneel handelen.

Literatuur

Akker, J. van den (1999). Principles and methods of development research. In J. van den Akker, R. Branch, K. Gustafson, N. Nieveen, & T. Plomp (eds.), *Design approaches and tools in education and training* (pp. 1-14). Dordrecht: Kluwer Academic Publishers.

Akker, J. van den (2003). Curriculum perspectives: An introduction. In J. van den Akker, W. Kuiper, & U. Hameyer (eds.), *Curriculum landscapes and trends* (pp.1-10). Dordrecht:

Kluwer Academic Publishers.

Akker, J. van den (2010). Building bridges: How research may improve curriculum policies and classroom practices. In S. Stoney (ed.), *Beyond Lisbon 2010: Perspectives from research and development for education policy in Europe* (pp. 177- 195). [CIDREE Yearbook 2010]. Slough: NFER.

Akker, J. van den, & Kuiper, W. (2008). Research on models for instructional design. In J. Spector, M. Merrill, J. van Merriënboer, & M. Driscoll (eds.), *Handbook of research for educational communications and technology* (pp. 739-748). New York: Lawrence Erlbaum Associates.

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

Beeker, A. (2009). Mind the gap. Een website met lesideeën Engels voor vwo 6-lerlingen. *Levende Talen Magazine*, 3, 16-19.

Biesta, G. (2012). *Goed onderwijs en de cultuur van het meten*. Den Haag: Boom Lemma.

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

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

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3-12.

Commissie Dijsselbloem (2008). *Tijd voor onderwijs*. Tweede Kamer vergaderjaar 2007-2008, 31 007, No.6.

Commissie De Graaf (2010). *Nationaal Plan Onderwijs/leerwetenschappen*. Den Haag: Ministerie van OCW.

Cras, A. (2009). *Meesterschap over het leerplan* (proefschrift Universiteit Twente). Enschede: SLO.

Elmore, R. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, 66(1), 1-26.

Fullan, M. (2000). The return of large-scale reform. *Journal of Educational Change*, 1, 1-23

Fullan, M. (2007). *The new meaning of educational change* (fourth edition). New York: Teachers College Press.

Goodlad, J. I. and Associates (1979). *Curriculum inquiry: The study of curriculum practice*. New York: McGraw-Hill.

Gravemeijer, K. P. E., & Kirschner, P. A. (2008).

Discussie 'Dupliek: een te simpele voorstelling van zaken'. *Pedagogische Studiën*, 85, 195-197.

- Handelzalts, A. (2010). *Samenwerking tussen docenten bij curriculumontwikkeling in het kader van onderwijsvernieuwing* (proefschrift). Enschede: Universiteit Twente.
- Hargreaves, A. & Fink, D. (2006). *Sustainable leadership*. San Francisco, CA: Jossey-Bass.
- Hoeven, M. van der, Kuiper, W., Meijs, L., Ottevanger, W., Spek, W., & Eijkelhof, H. (2012). *Kennisbases onderbouw voortgezet onderwijs. Rapportage van een vooronderzoek*. Enschede: SLO.
- Kuiper, W. (2009). *Curriculumevaluatie en verantwoordende vernieuwing van bètaonderwijs* (oratie). Utrecht/Enschede: Universiteit Utrecht/SLO.
- Kuiper, W., Folmer, E., Ottevanger, W., & Bruning, L. (2011). *Curriculumevaluatie bètaonderwijs tweede fase. Samenvattend eindrapport*. SLO: Enschede.
- Kuiper, W., Akker, J. van den, Letschert, J., & Hooghoff, H. (2007). *Curriculum policy and practices in European comparative perspective: Finding a balance between central structure and local autonomy*. Enschede: SLO.
- Kuiper, W., Hoeven, M. van der, Folmer, E., Graft, M. van, & Akker, J. van den (2010). *Leerplankundig analyse van PISA-trends*. Enschede: SLO.
- Lieberman, A., & Miller, L. (eds.). (2008). *Teachers in professional communities: Improving teaching and learning*. New York: Teachers College Press.
- Marsh, C. (2009). *Key concepts for understanding curriculum*. London: Routledge.
- McKenney, S., & Reeves, T. (2012). *Conducting educational design research*. London: Routledge.
- McLaughlin, M., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper and going broader. *Journal of Educational Change*, 2(4), 301-323.
- Nieveen, N. (2009). Formative evaluation in educational design research. In T. Plomp & N. Nieveen (eds). *An introduction to educational design research* (pp. 89-101), Enschede: SLO.
- Nieveen, N., van den Akker, J., & Resink, F. (2010). Framing and supporting school-based curriculum development in the Netherlands.

In E. H-F. Law & N. Nieveen (eds). *Schools as curriculum agencies: Asian and European perspectives on school-based curriculum development* (pp. 273-283), Rotterdam: Sense.

- Nieveen, N., & van der Hoeven, M. (2011). Building the curricular capacity of teachers: Insights from the Netherlands. In P. Picard & L. Ria (eds.), *Beginning teachers: A challenge for educational systems – CIDREE Yearbook 2011* (pp. 49-64). Lyon, France: ENS de Lyon, Institut Français de l'Éducation.
- Onderwijsraad (2006). *Naar meer evidence-based onderwijs*. Den Haag: Onderwijsraad.
- Pareja Roblin, N., Corbalan Perez, G., McKenney, S., Nieveen, N., & van den Akker, J. (2012). *Designing for scale: How relationships shape curriculum change*. Paper presented at AERA, April 13-17, Vancouver.
- Penuel, W., & Fishman, B. (2012). Large-scale science education intervention research we can use. *Journal of Research in Science Teaching*, 49(3), 281-304.
- Pieters, J. M., & Vries, B. de (2005). *Kennisproductie en kennisdisseminatie in het Nederlandse onderwijsveld: Een voorstudie naar de rol van kennisgemeenschappen*. Enschede: Universiteit Twente.
- Plomp, T., & Nieveen, N. (eds.) (2009). *An introduction to educational design research*. Enschede: SLO.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston/New York: Houghton Mifflin Company.
- Thijs, A., & Akker, J. van den (2009). *Leerplan in ontwikkeling*. Enschede: SLO.

Manuscript aanvaard op: 19 september 2012

Auteurs

De auteurs zijn werkzaam bij SLO: **Jan van den Akker** als algemeen directeur/bestuurder, **Wilma Kuiper** als manager van de afdeling 'Onderzoek en advies' en **Nienke Nieveen** als senior-onderzoeker.

Correspondentieadres: prof. dr. J. J. H van den Akker, SLO, postbus 2041, 7500CA Enschede.
E-mail: j.vandenakker@slo.nl.

Abstract

Building bridges between policy, practice and research to guide curriculum development

It is helpful when curriculum development is research-informed, but the relationship between educational research and curriculum development leaves much to be desired. In this article, we propose adequate use of research within an integral approach to curriculum development. The approach is visualized as a triptych. Actual curriculum development (second panel) is preceded by monitoring and analysis of trends in practice, science, and society (first panel). Discourse about monitoring and analysis outcomes contributes to agenda setting about improvement and innovation projects (second panel). The deliberate involvement of relevant system partners with monitoring and analysis activities as well as with improvement and innovation projects is also meant to anticipate favorable conditions for implementation and up-scaling (third panel). Typical for the second panel is a design research approach consisting of an interactive cycle of analysis, design and formative evaluation activities, with great user involvement.