

Naar meer evidence based onderwijs!

R. J. Bosker

Samenvatting

In een kritische bespreking van *Naar meer evidence based onderwijs* (Onderwijsraad, 2006) sommen Gravemeijer en Kirschner een groot aantal bezwaren op tegen dit advies van de Onderwijsraad. Die bezwaren snijden niet altijd hout. Dat het echte experiment de enige juiste onderzoeksmethode zou zijn, wordt bestreden, verwijzend naar het onderscheid tussen *context of discovery* versus *context of justification*. Dat er allerlei praktische en ethische bezwaren aan experimenten kleven, zijn geen onoverkomelijke hindernissen. Gevonden wordt op het probleem dat het helaas usance is dat er veel geëxperimenteerd wordt in de dagelijkse onderwijspraktijk en dat er menige innovatie wordt uitgerold zonder dat de werkzaamheid daarvan bewezen is.

1 Inleiding

In hun reactie op het advies van de Onderwijsraad (2006) over *evidence based* onderwijs brengen Gravemeijer en Kirschner (2007) een aantal bezwaren tegen de in dat advies geformuleerde gedachtengang over het gebruik van echte experimenten in onderwijskundig onderzoek naar voren. In deze bijdrage ga ik op deze bezwaren in, waarbij ik zal betogen dat het geen zin heeft te begrijpen *hoe* iets werkt, zoals Gravemeijer en Kirschner vooral bepleiten, als we niet weten *dat* iets werkt in het onderwijs. En over de vele praktische bezwaren die ze tegen echte experimenten naar voren brengen wordt aan aantal relativerende opmerkingen gemaakt.

2 Wat is bewijs? Het genereren versus het beproeven van kennis

Het antwoord op de vraag wat een geschikte onderzoeksmethode is, hangt af van de onder-

zoeksvraag die gesteld wordt (cf., Raudenbush, 2005). In het algemeen is daarbij het onderscheid tussen de *context of discovery* versus de *context of justification* handzaam. Bij het genereren van kennis in de *context of discovery* is elke onderzoeksmethode min of meer geoorloofd, zolang men het principe van *disciplined inquiry* recht doet: het volgen van een reeks methodologische voorschriften die garanderen dat het onderzoek transparant en het resultaat herhaalbaar is. De vraag of de uiteindelijk geconstrueerde theorie over wat er werkt in het onderwijs houdbaar is in de *context of justification*, kan alleen maar beslecht worden door onderzoek waarin op zo hard mogelijk wijze de oorzaak-gevolgrelatie wordt onderbouwd. Gesteld dat de theorie zich na vele jaren hard werken laat concretiseren in een praktische en overdraagbare interventie, en dat is toch wat we mogen verwachten van een onderwijskundige theorie, dan kan het meest harde bewijs voor het bestaan van die oorzaak-gevolgrelatie gevonden worden met behulp van een echt experiment. In een echt onderwijskundig experiment worden zowel leerlingen als leerkrachten aselect toegewezen aan de experimentele of de controlegroep. Alle verschillen tussen de experimentele en controlegroep zijn dan toevallig en verwaarloosbaar (Rubin, 1974), zodat verschillen in de uitkomsten toe te wijzen zijn aan de interventie.

In het advies "Naar meer evidence based onderwijs" (Onderwijsraad, 2006) houdt de Onderwijsraad een pleidooi om te streven naar het verzamelen van zo hard mogelijk bewijs over wat er werkt in het onderwijs, waarbij zij benadrukt dat het slechts het voorlopige sluitstuk is van een proces van zorgvuldige kennisaccumulatie. In dat proces is inzicht opgedaan hoe een aantal ingrepen mogelijk effect zouden kunnen hebben op de lerende. Dat het effect ook werkelijk optreedt wanneer die inzichten vertaald zijn in een (complexe) interventie, maakt het aannemelijk dat de inzichten over waarom die interventie werkt vermoedelijk correct zijn.

3 Over de bezwaren tegen het echte experiment

Bij de zogenaamde *cluster randomized controlled trial* worden niet alleen leerlingen aselekt toegewezen aan de verschillende condities, maar ook scholen of leerkrachten. Niet alleen vindt er randomisatie plaats, maar het gebeurt ook nog eens groepsgewijs. Vandaar de toevoeging *cluster*. Gravemeijer en Kirschner sommen een reeks bezwaren op tegen deze onderzoeksmethode.

In de eerste plaats stellen ze dat het niet uitvoerbaar is. Daarbij gaan ze ervan uit dat leerlingen in de controleconditie een placebo ontvangen, maar dat heeft niemand ooit beweet. De controleconditie is een reguliere setting, ofwel de bestaande onderwijspraktijk. Dat kan men toch niet onethisch vinden? Randomisatie van scholen achten ze vervolgens onmogelijk, maar in de praktijk hoeft dat geen probleem te zijn als men met slimme oplossingen werkt: scholen die bereid zijn met de interventie te werken, voeren die bijvoorbeeld bij een andere jaargroep in dan in de experimentele groep gebeurt (het zogenaamde *cross-over design*). Dat we altijd met zeer grote aantallen zouden moeten werken is iets te ongenueanceerd: bij de opzet van het onderzoek worden de aantallen zo gekozen dat deze in verhouding staan tot de effecten die men meent aan te kunnen treffen en de a priori gekozen risico's op statistische fouten van de eerste en tweede soort. Verwacht men grote effecten dan kunnen de aantallen neerwaarts worden bijgesteld.

Dat een echt experiment niet betaalbaar is, is een volgende tegenwerping. Gravemeijer en Kirschner verwijzen daarbij naar de grote sommen geld die met een echt experiment gemoeid zijn en de risico's dat maar in een enkel geval een effect kan worden aangetoond. Men zet echter pas een experiment op als na jaren van kennisvergaring, exploratie, vallen en opstaan, men het gerede vermoeden heeft dat er een werkzame en overdraagbare interventie mogelijk is. En dat dat experiment geld kost, valt in het niet tegen de enorme sommen geld die in het onderwijs omgaan met ongerichte en op hun effectiviteit onbewezen interventies. Als er al iets onethisch is, dan is het dat daar kinderen aan worden blootgesteld.

Bij het bezwaar dat het te generaliserend zou zijn, verwijzen Gravemeijer en Kirschner naar het probleem van de *early adopters* bij wie de aanpak wel zou kunnen werken, maar in de dagelijkse onderwijspraktijk niet. Mijn tegenwerping: toon nu eerst maar eens dat het werkt (hetzij bij de *early adopters* hetzij bij anderen), en ontwerp vervolgens een verstandig implementatietraject met intieële scholing en bij-, her- en nascholing van leerkrachten, zodat ze niet als domme en luie "boeren" (die door Gravemeijer en Kirschner als voorbeeld ten tonele worden gevoerd) met het goud van een goed onderwijsprogramma omgaan. Daarbij kan dan voor de (aanstaande) leerkrachten ruim aan bod komen waarom de interventie werkt, zodat de leerkracht professioneel met het programma aan de slag kan.

Verder wijzen Gravemeijer en Kirschner erop dat de door de Onderwijsraad voorgestane benadering vertragend en conserverend werkt. Daar staat tegenover dat er veel overhaaste en ondoordachte innovaties in het onderwijs worden geïntroduceerd die vervolgens veel te snel, namelijk voordat we weten of het echt werkt, worden verspreid over het hele onderwijs. En dat meer innovatieve doelen moeilijk toetsbaar zijn, wil niet zeggen dat ze niet toetsbaar zijn. En hierbij betekent toetsbaar niet per definitie dat het om een papier-en-potlood meerkeuzetoets dient te gaan.

4 Tot slot

Onderwijskunde is een toegepaste interdisciplinaire wetenschap met een duidelijk optimalisatieperspectief: het onderwijs beter maken. In de onderwijspraktijk van alle dag wordt er volop door leerkrachten met leerlingen geëxperimenteerd, onder meer verwijzend naar het grondwettelijke beginsel dat de overheid zich niet met de inrichting van het onderwijs zou mogen bemoeien. Enorme hoeveelheden geld worden gependend aan het oplossen van problemen (onderwijsachterstanden, voortijdig schoolverlaten, de geringe belangstelling voor bèta en techniek, uitval van beginnende leerkrachten) zonder dat er enig zicht is dat het daadwerkelijk

beter zal gaan. Er zijn leerlingen die met enorme achterstanden het basisonderwijs binnenkomen én met enorme achterstanden het funderend onderwijs verlaten. Er worden volledig dekkende en op hun effecten ononderzochte innovaties over het hele onderwijs ingevoerd, waarmee de facto met alle leerlingen wordt geëxperimenteerd. Daarvoor geldt mijns inziens het homeopathieprincipe: laat de bedenkers van die innovaties eerst maar eens aantonen dat het voor de leerlingen niet schadelijk is, en dan in het bijzonder voor die leerlingen uit risicogroepen die met name op het onderwijs zijn aangewezen om zich verder te ontwikkelen.

Niet alles laat zich door experimenten bewijzen. De vraag of een geïntegreerd onderwijssysteem beter is dan een gedifferentieerd stelsel, is daarvan een voorbeeld. In dat geval doet men er onder meer verstandig aan met wetenschappelijke nijverheid alle evidentie verzameld in internationaal vergelijkend onderzoek netjes op een rij te zetten. Niet alle theorieën laten zich door experimenten bewijzen. Maar dat prachtige theorieën in de *context of discovery* zonder experimenten tot stand zijn gekomen wil niet zeggen dat daarmee het experiment in de *context of justification* niet een nastrevenswaardig wetenschappelijk ideaal is in menige situatie.

Dit is geen pleidooi voor *teacher proof*-onderwijsprogramma's. In tegendeel. Juist in de opleiding van leerkrachten zou ruim aandacht besteed moeten worden aan een goede inleiding in de onderwijskundige kennisbasis, en bij voorkeur ook aan het ontwikkelen van een academische reflexieve houding ten aanzien van het eigen handelen. Zodat leerkrachten als echte professionals hun werk kunnen doen en met een gedegen kennis over wat er waarom voor welke leerlingen werkt in het onderwijs. We kunnen niet volhouden dat onderwijzen alleen een kunst is en geen kunde.