

Sardes

Speciale Editie

Ideeën over het Bewijs voor Onderwijs

Een debat over de (on)mogelijkheden van
evidence based onderwijs in de praktijk

NUMMER 7 | APRIL 2009





Voorwoord

Sabine Peterink, Anne Luc van der Vegt en Sandra Beekhoven | Sardes

Het streven naar evidence based onderwijs is in opmars. Vorig jaar wees de commissie Dijsselbloem erop dat de effectiviteit van veel onderwijskundige vernieuwingen onvoldoende wetenschappelijk is bewezen (dus niet evidence based is). In mei vorig jaar werd TIER opgericht: het Top Institute for Evidence Based Education Research, een samenwerkingsverband van de universiteiten van Maastricht, Groningen en Amsterdam. Via het Actieprogramma Onderwijs Bewijs stelt de rijksoverheid geld beschikbaar om experimenteel onderzoek uit te voeren in het primair en voortgezet onderwijs. Doel hiervan is kennis te verkrijgen over wat wel en niet werkt in het onderwijs. Scholen en onderzoeksinstellingen konden hiervoor samen een aanvraag indienen.

Voor Sardes is de toenemende populariteit van evidence based onderwijs een aanleiding om een Speciale Editie aan dit onderwerp te wijden. Elke Speciale Editie van Sardes is gewijd aan één bijzonder thema, dat altijd wordt beschouwd vanuit de driehoek wetenschap - beleid - praktijk. En dat is niet voor niets; het is de werkwijze van Sardes om inzichten uit onderzoek bruikbaar te maken voor beleid en praktijk.

Voor deze Speciale Editie hebben we een aantal voorlopers en spelers uit het 'evidence based veld' gevraagd om hun opvattingen en ervaringen met u te delen. Eerst laten we de wetenschap aan het woord: de positieve invalshoek en de kritische noot. Wim Groot en Henriëtte Maassen van den Brink bezetten gezamenlijk de leerstoel evidence based onderwijs bij TIER. Zij beschrijven een aantal hoopgevende ontwikkelingen in het empirisch onderzoek naar de effectiviteit van onderwijsinterventies. Paul Kirschner, hoogleraar Psychologie en Leven-Lang-Leren bij de Open Universiteit Nederland, maakt een aantal kritische kanttekeningen bij de mogelijkheden om in het onderwijs daadwerkelijk evidence based te werken. Jo Nelissen van het Freudenthal Instituut beschouwt het streven naar werken met effectief bewezen programma's in het onderwijs in historisch perspectief en staat stil bij de vraag wat 'evidence' nu eigenlijk inhoudt. De brug naar de praktijk wordt gelegd door Alfons ten Brummelhuis, die een kijkje geeft in de keuken van Kennisnet, waar men werkt aan kennisontwikkeling om effectiever gebruik te maken van ICT in de onderwijspraktijk. Marian de Graaf, Hans Meij en Ton van Yperen van het Nederlands Jeugdinstituut geven een beeld van de werkwijze en de criteria die worden gebruikt bij de databank Effectieve Jeugdinterventies. Zij pleiten voor inspraak van de beroepsverenigingen van professionals om de ontwikkeling van een effectief aanbod te bevorderen. En wat vindt het onderwijsveld zelf? Kars Veling, directeur van de Johan de Witt scholengroep, pleit op basis van zijn ervaringen met het 'Talentencentrum' binnen zijn school voor het verbinden van onderwijspraktijk en wetenschappelijk onderzoek in een gezamenlijk gedragen cyclus van onderwijs, ontwerp en onderzoek.

In basisschool Het Baken in Veenendaal is onderzocht volgens welke methode leerstof het beste aan leerlingen kan worden aangeboden. Monique de Haan testte twee onderwijsconcepten bij twee vergelijkbare groepen leerlingen en beschrijft wat er allemaal in de school moest gebeuren om dit te kunnen doen.

Onderwijsonderzoekers IJsbrand Jepma en Anne Luc van der Vegt gaan een stapje verder door te kijken naar het onderwijsbeleid: kan van beleid worden aangetoond of het effectief is?

Jo Kloprogge, adviseur van Sardes, stelt ten slotte een aantal kritische vragen over de trendgevoeligheid van evidence based onderwijs.

Wij wensen u veel leesplezier en nieuwe inzichten toe.

Ideeën over het bewijs voor onderwijs

Een debat over de (on)mogelijkheden van evidence based onderwijs in de praktijk

Inhoudsopgave

- 1 Onderwijs werkt niet?
Wim Groot en Henriëtte Maassen van den Brink 5
- 2 Evidence based onderwijs: waar gaat het nou over?
Paul Kirschner 9
- 3 Kwalitatief onderzoek en objectiviteit
Jo Nelissen 13
- 4 Praktijkgerichte kennisontwikkeling
Alfons ten Brummelhuis 17
- 5 Allemaal aan de RCT?
Marian de Graaf, Hans Meij en Ton van Yperen 21
- 6 Naar structurele coalities van scholen en onderzoekers
Kars Veling 25
- 7 De noodzaak van de onderzoekende leerkracht
Monique de Haan 31
- 8 Onderwijsbeleid op wetenschappelijke leest geschoeid
IJsbrand Jepma en Anne Luc van der Vegt 35
- 9 Vier vragen over evidence based werken in het onderwijs
Jo Kloprogge 41

Onderwijs werkt niet?

Wim Groot en Henriëtte Maassen van den Brink | [Topinstituut Evidence Based Education Research \(TIER\)](#), [Universiteit Maastricht/Universiteit van Amsterdam](#)

Nederland loopt in de ontwikkeling van evidence based onderwijs achter bij Groot-Brittannië en de Verenigde Staten. Onderwijsonderzoek is in Nederland onvoldoende richtinggevend voor de praktijk in de scholen en onderzoekers zijn weinig zichtbaar in de beleidsdiscussies over de toekomst van het onderwijs. Toch zijn er hoopgevende ontwikkelingen. Nieuwe vormen van empirisch onderzoek naar de effectiviteit van onderwijsbeleid blijken zeer waardevol. De auteurs geven hiervan twee voorbeelden.

Het onderwijs kampt met een aantal problemen. Zo bestaat er in brede kring ontevredenheid over de kwaliteit van het onderwijs en stevent het onderwijs af op een tekort aan leraren. Ruim vijftigduizend oudere leraren zullen de komende tien jaar het voortgezet onderwijs verlaten en de instroom van jonge leraren is bij lange na niet voldoende om dit gat te dichten: vorig jaar telde de eerstegraads lerarenopleiding in het HBO welgeteld 1.372 studenten en de universitaire lerarenopleiding 1.415 studenten. Het lerarentekort is een bedreiging voor de kwaliteit van het onderwijs. In het voortgezet onderwijs wordt 11% van de lessen door onbevoegde docenten gegeven en 16% door onderbevoegde docenten. Naarmate het lerarentekort nijpender wordt, zullen deze percentages verder toenemen.

Wat werkt?

Om bij een toenemend lerarentekort de kwaliteit van het onderwijs te verhogen, zal er slimmer moeten worden gewerkt. Dat wil zeggen dat leraren niet 'zomaar wat doen', maar zich in hun professionele handelen moeten laten leiden door 'wat werkt'. Als zij zich daardoor laten leiden in hun werk, wordt hun effectiviteit groter. Om erachter te komen wat werkt, welke onderwijsinterventies effectief zijn, is het nodig dat onderwijsinterventies wetenschappelijk worden geëvalueerd. Evidence based onderwijs is de verzameling van onderwijsinterventies met een bewezen positieve (kosten) effectiviteit. Het is onderwijs dat zich baseert op de beste evidentie over wat werkt.

Om evidence based onderwijs te realiseren, moeten specifieke onderwijsinterventies, strategieën en beleid wetenschappelijk worden geëvalueerd voordat deze worden aanbevolen of op brede schaal worden ingevoerd. Is die evaluatie nog niet uitgevoerd, dan dienen de interventies op experimentele basis op zo'n manier te worden ingevoerd, dat de effecten ervan wetenschappelijk geëvalueerd kunnen worden. Evidence based onderwijs verwijst enerzijds naar de verzameling van interventies die positief geëvalueerd zijn, anderzijds naar een attitude van docenten en beleidsmakers, namelijk: werken op basis van evidence based onderwijsinterventies. Leraren en scholen in het voortgezet onderwijs kunnen zich onderscheiden door evidence based te werken.

Nederland loopt achter

Nederland loopt in de ontwikkeling van evidence based onderwijs achter bij Groot-Brittannië en de Verenigde Staten. Doordat systematisch onderzoek naar kosten-effectiviteit van interventies in het onderwijs ontbreekt, wordt de invoering en verspreiding van doelmatige innovaties belemmerd. Het onderwijsonderzoek in Nederland kampt met een aantal problemen. Zo is de onmiddellijke bruikbaarheid van dit onderzoek voor de onderwijspraktijk vaak moeilijk aantoonbaar. Vaak klinkt het verwijt dat de resultaten van onderwijskundig onderzoek onvoldoende relevant zijn voor de onderwijspraktijk. Verder zijn onderwijskundige onderzoekers weinig zichtbaar in de beleidsdiscussies over de toekomst van het onderwijs en is er een gebrek aan kennis over deugdelijk evaluatieonderzoek.

Het beschikbare onderzoek naar de kosteneffectiviteit van onderwijsinterventies heeft vooral betrekking op interventies buiten Nederland. De resultaten hiervan zijn niet altijd toepasbaar op de Nederlandse situatie. Er is een gebrek aan een heldere en eenduidige systematiek in het onderzoek, vooral waar het gaat om het bepalen van de kosten en effecten van interventies. Hierdoor zijn de kwaliteit en de relevantie van het beschikbare onderzoek vaak gering en zijn studies onderling onvergelijkbaar. Dit maakt een prioritering op basis van kosteneffectiviteitsratio's onmogelijk. Het zou helpen als er richtlijnen worden geformuleerd waaraan deugdelijk kosteneffectiviteitsonderzoek in het onderwijs moet voldoen.

De kennispiramide, met bovenin de hardste vormen van bewijs over wat werkt - het experiment met randomisatie - houdt in Nederland meestal op bij quasi-experimenteel onderzoek, waarbij natuurlijke implementatievarianten van nationaal beleid op hun effecten worden vergeleken. Echte experimenten veronderstellen dat overheden en besturen bereid zijn om tijdelijk geld toe te wijzen voor uitbreiding van de formatie op scholen die deelnemen aan de evaluatie van onderwijskundige interventies. Het ethische principe dat leerlingen niet mogen lijden onder experimenten, en zeker niet als ze in de controlegroep zitten, impliceert dat ook de controlegroep iets extra's geboden moet worden. Het argument dat leerlingen die betrokken zijn bij het experiment 'proefkonijnen' zouden zijn, is overigens niet valide: we onthouden immers vele kinderen succesvol onderwijs. En bovendien experimenteren scholen, ongecontroleerd en zonder deugdelijk evaluatieonderzoek, volop met leerlingen.

Evidence based onderwijs legt de nadruk op het investeringskarakter van onderwijsuitgaven: onderwijs is een investering die een positief (en hoog) rendement kan opleveren voor zowel de betrokkenen als voor de samenleving als geheel. Op grond van een investeringsbenadering heeft de maatschappij er profijt van als middelen worden besteed aan interventies met een gunstige kosten-batenverhouding. Dit geldt voor elke onderwijsinvestering, maar in het bijzonder voor investeringen in jongeren uit

achterstandsgroepen. Of, zoals de economen en onderwijskundigen Levin, Belfield, Muennig en Rouse¹ opmerken: “The investment criterion is a simple one: public investments are worth making if the benefits exceed the costs. Even if education is expensive, poor and inadequate education for substantial numbers of our young may have consequences that are even more costly. Such an analysis goes beyond the more basic question of social justice. If life chances depend heavily on education, it is important that inequalities in education associated with race, gender, immigrant status, language, and handicap be redressed as a basis for equalizing opportunities in a democratic society. But, even beyond the issue of injustice is the question of whether a poor quality education has consequences for the larger society. Social science research shows that poor education imposes social burdens via lower incomes and economic growth, lower tax revenues, and higher costs of such public services as health, criminal justice, and public assistance. In this respect, it is possible to view efforts to improve educational outcomes for at-risk populations as a public investment that may have benefits for the entire society in excess of investment costs.”

Nieuwe mogelijkheden

Uit de internationale literatuur blijkt dat er in het afgelopen decennium een belangrijke vordering is gemaakt in de ontwikkeling van methoden om effecten van onderwijsinterventies op overtuigende wijze te meten op macro-, meso- en microniveau. Nieuwe vormen van empirisch onderzoek naar de effectiviteit van verschillende vormen van onderwijsbeleid blijken zeer waardevol. Het gaat hierbij vooral om het uitvoeren van experimenten en het gebruikmaken van vaak bij toeval ontstane omstandigheden die sterk op experimenten lijken. De effecten van onderwijsinterventies worden vastgesteld door de uitkomsten van een experimentele groep, die is onderworpen aan bepaald beleid, te vergelijken met de uitkomsten van een door toeval gevormde controlegroep, die niet aan dit beleid is onderworpen. Deze vergelijking levert veelal harder bewijs op over de effecten van onderwijsinterventies.

Enkele toepassingen in Nederland

In de afgelopen jaren zijn deze methoden ook door Nederlandse onderwijseconomen en onderwijskundigen toegepast op (onderdelen van) onderwijsinterventies in Nederland. Hieronder volgt een korte weergave van enkele van deze studies.

Leuven, Oosterbeek en Van der Klaauw² hebben een veldexperiment uitgevoerd waarin eerstejaars economiestudenten aan de Universiteit van Amsterdam een financiële beloning in het vooruitzicht werd gesteld als zij hun propedeuse voor het begin van het volgende studiejaar zouden behalen. De resultaten laten zien dat goede studenten (met hoge wiskundecijfers in het voortgezet onderwijs) door deze financiële prikkel meer studiepunten haalden, terwijl slechte studenten (met lage wiskundecijfers) juist minder studiepunten haalden. Bovendien bleek dat - hoewel de prikkel alleen in het eerste stu-

diejaar gold - na drie jaar het initiële effect fors was uitvergroot. Dit kleinschalige veldexperiment levert waardevolle inzichten op over het effect van financiële prikkels op de studievoortgang en over het belang van 'op schema' studeren. Dergelijke inzichten kunnen worden gebruikt bij de vormgeving van een stelsel van studiefinanciering.

Doolaard en Bosker³ onderzochten de effecten van vier verschillende manieren waarop scholen hun formatie inzetten in de onderbouw van het basisonderwijs: kleine groepen, gemiddeld grote groepen, groepen met (naast de leerkracht) een klassenassistent en grote groepen. Door de taal- en rekenontwikkeling van leerlingen gedurende een aantal jaren te volgen, is vastgesteld dat er geen effecten traceerbaar zijn in groep 2, maar wel in groep 3. Leerlingen in kleine en gemiddeld grote groepen ontwikkelen zich gunstiger dan andere leerlingen en deze effecten beklijven tot en met het einde van groep 4. Een aanvullende analyse van de kosteneffectiviteit leverde het verrassende inzicht op, dat uiteindelijk de middelgrote groep (met ca. 20 leerlingen) het meest kosteneffectief was: hier deed zich de gunstigste verhouding voor tussen effecten en kosten.

De genoemde voorbeelden zijn creatief opgezette studies. Geen van deze afzonderlijke studies levert echter het definitieve bewijs dat een interventie wel of niet werkt. Daarmee leveren deze studies ook geen direct bruikbare informatie op voor beleid of praktijk. Het zijn stukjes van een puzzel die pas door gedegen meta-analyses inzicht bieden in wat er werkt⁴.

Bovenstaand overzicht illustreert de mogelijkheden van op moderne leest geschoeide evaluatiestudies in Nederland. Dat het aantal van dergelijke studies desondanks beperkt is, komt niet door een gebrek aan beleid dat om een gedegen evaluatie vraagt, maar door:

- (1) een tot nu toe onvoldoende besef van het nut van evidence based policy
- (2) een (te) beperkte onderzoekscapaciteit
- (3) een matige data-infrastructuur
- (4) het ontbreken van middelen waarmee scholen de experimentele interventie respectievelijk een controle-interventie getrouw kunnen implementeren
- (5) onwetendheid, onbegrip en aversie tegen onderzoek naar effectiviteit.

- 1 Levin, H.M., Belfield, C.R., Muennig, P. & Rouse, C.E. (2006). The Costs and Benefits of an Excellent Education for America's Children. Working Paper, Teachers College, Columbia University
- 2 Leuven, E., Lindahl, M., Oosterbeek, H. & Webbink, D. (2008). The effect of extra funding for disadvantaged students on achievement. In: Review of Economics and Statistics (forthcoming)
- 3 Doolaard, S. & Bosker, R.J. (2006). Effecten van formatie-inzet in de onderbouw van het basisonderwijs. Groningen: GION
- 4 Groot, W. & Maassen van den Brink, H. (2000). Overeducation in the labor market: a meta-analysis. In: Economics of Education Review, 19, p. 149-158; Furnee, C., Groot, W. & Maassen van den Brink, H. (2008). The health effects of education: a meta-analysis. In: European Journal of Public Health, 18, p. 417-421; Huang, J., Maassen van den Brink, H. & Groot, W. (2008). A meta-analysis of the effect of education on social capital. In: Economics of Education Review (forthcoming)

Evidence based onderwijs: waar gaat het nou over?

Paul A. Kirschner, hoogleraar Psychologie en Leven-Lang-Leren | Open Universiteit Nederland

Is het wel mogelijk om in het onderwijs evidence based te werken? Pleitbezorgers van evidence based onderwijs zien een aantal problemen over het hoofd, zo betoogt dit artikel, dat een bewerking is van een eerder gepubliceerde column op SURF Edusite uit 2006¹. Er zijn vier problemen, die het pleidooi voor evidence based onderwijs ontkrachten. Deze problemen verwijzen onder andere naar de onderzoeksmethodologie en naar de aard van het onderwijs.

Het verleden heeft ons geleerd dat wat er in de Verenigde Staten gebeurt, na enige tijd overwaait naar Nederland. Dat geldt voor mode en muziek, maar ook voor het overheidsbeleid, zoals de deregulering en het marktdenken van Ronald Reagan en op dit moment, de economische crisis. Ook evidence based onderwijs werd voor het eerst in de VS geïmplementeerd.

No-brainer

Aanleiding voor mijn eerste column over evidence based onderwijs was een uitspraak van professor Fons van Wieringen, voorzitter van de Onderwijsraad, in *Trouw*, waarin hij stelde dat het onderwijs (en ook degenen die bezig zijn met onderwijsonderzoek) "...een voorbeeld kan nemen aan de gezondheidszorg. Daar komen medicijnen pas op de markt als de werkzaamheid wetenschappelijk is aangetoond. Het onderwijs kan daarvan leren..."² Van Wieringen deed deze uitspraak bij het verschijnen van het advies van de Onderwijsraad *Naar meer evidence based onderwijs*.

De gedachte dat een onderwijsverandering alleen mag worden ingevoerd als deze bewezen effectief is, vormt de basis van President Bush's *No child left behind act*. Deze Amerikaanse wet - volgens velen een bedekte poging van het Witte Huis om meer greep te krijgen op onderwijsonderzoek en -vernieuwing - houdt in dat een onderwijsprogramma pas geïmplementeerd mag worden als onomstotelijk kan worden bewezen dat er "...evidence [is], based on rigorous research and evaluation, that [it] can effectively help children to succeed in school." Zoiets klinkt heel goed; of zoals de Amerikanen het zeggen: "A no-brainer".

Laten wij wél onze hersenen gebruiken en goed kijken naar wat evidence based onderwijs met zich meebrengt. Ik ben er wel degelijk van overtuigd dat het niet slim is om onderwijsvormen in te voeren zonder dat we weten of deze enig nut hebben, maar naar mijn mening is evidence based onderwijs *niet* de oplossing. De evidence based benadering brengt ten minste vier problemen met zich mee. Het eerste probleem is methodologisch van aard, het tweede financieel, het derde is de tijdsfactor en het vierde probleem ligt in de aard van het onderwijs zelf.

Methodologische problemen

De evidence based benadering komt voort uit de gezondheidszorg en vooral de farmaceutische industrie. Voordat een geneesmiddel op de markt komt, wordt onderzocht of het voldoende werkzaam en veilig is. We moeten er echter voor oppassen om een methodologie uit het ene domein (farmacie) zomaar toe te passen in een ander domein (onderwijs).

In de farmaceutische industrie heeft het onderzoek de volgende kenmerken:

- Er wordt gebruikgemaakt van controlegroepen, die in plaats van het testgeneesmiddel een placebo of een bestaand geneesmiddel krijgen. Deelnemers en behandelend artsen weten niet wie in welke groep zit (zogenoemde double-blind experimenten).
- Randomisatie: deelnemers worden willekeurig ingedeeld in de experimentele groep of de controlegroep.
- Er is sprake van onderlinge vergelijkbaarheid van de groepen, door patiënten te kiezen met dezelfde eigenschappen, bijvoorbeeld alleen mannen tussen de 18 en 30 jaar.
- Er doen zeer veel mensen mee aan het onderzoek.
- Men streeft naar statistisch significante resultaten.

Robert Slavin³, pleitbezorger van de benadering, stelt deze eisen ook aan evidence based onderwijs. Maar kan dit - zowel praktisch als ethisch - wel allemaal in een onderwijssetting? Kunnen we een controlegroep leerlingen onderwijs geven waarvan we weten dat het geen of zelfs een negatief effect heeft? Kunnen wij docenten twee compleet verschillende didactieken laten toepassen in hun klas? En hoe kan dit double-blind gebeuren? Mogen of, beter gezegd, kunnen leerlingen willekeurig worden verdeeld in verschillende onderzoeksgroepen? Kunnen we vergelijkbaarheid controleren door bijvoorbeeld alleen leerlingen met een bepaald IQ of van een bepaalde leeftijd te laten participeren? En wat betekent het als iets statistisch significant (niet gebaseerd op toeval) is? Ter illustratie, de gemiddelde effectgrootte van levensreddende interventies met geneesmiddelen (zoals geneesmiddelen tegen hoge bloeddruk ter voorkoming van beroertes) zit tussen de 0,02 en 0,08, terwijl de gemiddelde grootte van het effect van onderwijskundige ingrepen 0,452 is! Dit is tussen de zeven en twintig keer zo hoog! Een effect kan significant zijn, maar kan nog steeds heel klein zijn. Zo blijkt bijvoorbeeld dat leerlingen bij een bepaalde lesmethode een duizendste kans hebben op een hoger onderwijsniveau: significant, maar niet van betekenis.

Een controlegroep moet een 'reguliere setting' zijn, ofwel de bestaande onderwijspraktijk. Maar hoe definiëren we de reguliere setting? Is dat onderwijs in een school in Haren of op Kanaleneiland Utrecht? Is dat Montessorionderwijs of de Vrije School? Is het klassikaal frontaal onderwijs of leerwijzen?

In een repliek op een artikel van Koeno Gravemeijer en mij⁴ stelde Roel Bosker⁵ dat randomisatie geen probleem zou zijn als men 'slim is'. Volgens Bosker is het pro-

bleem opgelost als er scholen te vinden zijn die bereid zijn met de interventie te werken in een *cross-over design*, waarbij verwisseling van interventie plaatsvindt. Maar zo eenvoudig (slim?) is dit niet! Ten eerste zal het niet gemakkelijk zijn docenten te vinden die bereid zijn om in twee verschillende klassen twee verschillende methoden/didactische aanpakken toe te passen. Daarnaast moeten we dan nog hopen dat er geen overdrachteffecten optreden. Volgens Clark⁶ moeten veel verschillen die bij onderwijsinterventies worden gevonden, eerder aan de docent dan aan de interventies worden toegeschreven. Met andere woorden: een *cross-over design* kan alleen werken in situaties waar de scholen bereid zijn deze organisatorische zaken te regelen en de docent geen persoonlijk stempel op de interventie drukt.

De financiën

Dan het financiële aspect. In 2005 kostte het een groot farmaceutisch bedrijf 359 miljoen dollar (300 miljoen euro) om één nieuw geneesmiddel op de markt te brengen. Slechts vijf uit vijftienduizend 'kandidaat' geneesmiddelen die pre-klinisch getest worden, worden ooit op mensen getest. Eén van die vijf middelen wordt vervolgens goedgekeurd. Men schat dat er jaarlijks 38 miljard dollar (32 miljard euro) wordt besteed aan onderzoek⁷. De uitgaven voor onderzoek en ontwikkeling bedragen circa een kwart van de inkomsten van zo'n bedrijf⁸. Het is de vraag of de politiek of de minister bereid is om een vergelijkbaar bedrag - een kwart van het onderwijsbudget - vrij te maken voor gedegen onderwijsonderzoek mét het risico dat slechts 0,02% van de onderzochte onderwijsveranderingen 'evidence based' zal zijn.

Tijd en onderwijs

Vervolgens noemden we het tijdsprobleem. Gemiddeld duurt het 12 tot 21,5 jaar om een geneesmiddel goed te keuren op basis van evidence based research⁹. Gek genoeg is dit ook ongeveer de tijd die het kost voordat resultaten van onderzoek op grote schaal zijn ingevoerd in de onderwijspraktijk. Het is helaas de ervaring dat bestuurders en politici vaak denken in termen van benoemings- en regeerperioden, dus hooguit over een periode van een jaar of vier en vaak zelfs korter. Zijn zij bereid een langere periode af te wachten, zodat een onderwijsinnovatie zorgvuldig kan worden uitgedacht en ingevoerd?

Het laatste probleem heeft met het onderwijs zelf te maken. Als een farmaceutisch bedrijf in evidence based onderzoek een medicijn of placebo toedient, maakt het niet uit of de zon schijnt, of het regent, of het in de keuken of in de badkamer gebeurt. De dosering is constant en het effect van het medicijn is niet afhankelijk van de persoon die het toedient. Onderwijs daarentegen is contextgebonden. Wat voor lerares A goed werkt in het eerste lesuur, hoeft voor haar niet per definitie ook in het tweede uur te werken. En wat werkt voor leraar B omdat hij zeer aimabel is, hoeft niet te werken voor lerares C die wat minder geliefd is.

De aanhangers van evidence based onderwijs houden zichzelf voor de gek met het idee dat het mogelijk is om harde bewijzen voor oorzaak-gevolgrelaties te vinden. In het onderwijs hebben wij te maken met docenten en leerlingen die, afzonderlijk én tezamen, invloed uitoefenen op de werking van een interventie. Daardoor kunnen er in een bepaalde situatie onder bepaalde voorwaarden verschillende dingen gebeuren. Zo kan een 'geplande' discussie in een werkgroep door de inbreng van één van de groepsleden ontaarden in iets heel anders of kan een goede samenwerking worden verpest door de houding van één groepslid. De evidence based beweging heeft te weinig oog voor dit gegeven. Je hebt er weinig aan om te weten dat een bepaalde aanpak 'gemiddeld beter werkt', als je niet weet hoe je ervoor kunt zorgen dat deze aanpak in jouw eigen situatie zal werken.

Op zijn Amerikaans

Een eenvoudig voorbeeld: stel dat keer op keer, in verschillende scholen, een positief leereffect op lezen wordt gemeten als de leerlingen tussen de reken- en de taallessen een half uurtje springoefeningen doen. We weten dan dus dat het werkt en volgens de evidence based benadering zouden we 'een half uur springen tussen de lessen door' invoeren. Maar dan moeten we wel zeker weten dat het effect verwijst naar het springen. Misschien is het effect wel te verklaren uit de afwisseling en zou een half uurtje zingen ook helpen, of maar 5 minuten, of.... Met andere woorden: waarom werkt iets?

Misschien ben ik wel een voorstander van evidence based onderwijs, maar dan echt op zijn Amerikaans. Dat wil zeggen dat we eerst een groot aantal - honderd tot tweehonderd - 'grassrootsprojecten' financieren voor een periode van twee jaar. Vervolgens worden de meest veelbelovende innovaties (10-20%) verder uitgetest en verfijnd. Daarna komen de beste projecten in aanmerking voor een degelijk en uitvoerig veldonderzoek (vergelijkbaar met klinische testing). En pas wanneer ook daaruit blijkt dat de innovatie goede resultaten oplevert, wordt de innovatie grootschalig in het onderwijs ingevoerd. Voor een dergelijk scenario zou ik wel tekenen; en mijns inziens de Commissie Dijsselbloem ook.

- 1 Deze bijdrage is een update van een Edusite column uit 2006 die toen heette "Evidence based onderzoek: De nieuwe wonderpil?!"
- 2 Trouw, 20 januari 2006, Onderwijsraad: Het nieuwe leren kan nadelig zijn <http://www.trouw.nl/deverdieping/dossiers/article134145.ece/>
- 3 Slavin, R. E. (1987). Ability grouping: A best-evidence synthesis. In: Review of Educational Research, 57, 293-336
- 4 Gravemeijer, K.P.E. & Kirschner, P.A. (2007). Naar meer evidence based onderwijs? [Towards more evidence based education?]. In: Pedagogische Studiën, 84, 463-472
- 5 Bosker, R.J. (2008). Naar meer evidence based onderwijs [Towards more evidence based education!]. In: Pedagogische Studiën, 85, 49-51
- 6 Clark, R. (2001). Learning from media: Arguments, analysis, and evidence. Greenwich, CT: Information Age Publishing
- 7 Drug development fell in slump in 2005. In: International Herald Tribune, 11 januari 2006; <http://www.iht.com/articles/2006/01/11/business/research.php>
- 8 http://www.netsci.org/cgi-bin/Courseware/projector.pl?Course_num=course1&Filename=slide09.html
- 9 Wisconsin Technology Network; <http://wistechnology.com/article.php?id=377>

3 Kwalitatief onderzoek en objectiviteit

Jo Nelissen | Freudenthal Instituut voor didactiek van wiskunde en natuurwetenschappen, Universiteit Utrecht

Voordat we in het onderwijs een bepaalde vernieuwing invoeren, moeten we over argumenten beschikken die de invoering van de innovatie rechtvaardigen. Volgens sommige onderzoekers levert vooral ‘evidence based onderzoek’ het harde bewijs en argumenten. Wat wordt beschouwd als koersbepalende ‘evidentie’? Bestaat er in de wetenschap overeenstemming over de inhoud en definitie van dit begrip? En hoe belangrijk is het om, vóór het kiezen van een onderzoeksmethode, eerst te definiëren wat we eigenlijk te weten willen komen?

Men moet over goede argumenten beschikken om (grote) vernieuwingen in het onderwijs te kunnen invoeren. Die argumenten moeten worden ondersteund door voldoende en controleerbare feiten die, als overtuigende ‘evidence’, in onderzoek zijn gegeneerd. Dat klinkt heel redelijk en dat *is* het ook. Ondanks die redelijkheid echter, ontkomen we niet aan de vragen wat onderzoekers als ‘evidence’ beschouwen en of daarover voldoende overeenstemming bestaat.

Wat is evidence?

De Amerikaanse psycholoog Kagan¹ stelt vast dat wetenschappers van verschillende disciplines “...rely on different forms of evidence when justifying their assertions”. De evidence waarvan wordt uitgegaan, is vaak evidence die een specifieke groep onderzoekers beschouwt als “evidence it regards as defining its territory”. Als voorbeeld bespreekt Kagan een onderzoek naar de vraag wanneer de mensheid kleren ging dragen. Uit DNA-onderzoek werd geconcludeerd dat 72.000 jaar geleden de mensheid voor het eerst kleren ging dragen. Dit onderzoek levert echter een soort “evidence” op waar historici helemaal niet naar op zoek zijn.

Een ander voorbeeld betreft een onderzoek van de Nederlandse Onderwijsinspectie². De onderzoekers rapporteren dat ruim 25% van de scholen ‘rekenzwak’ is. Is hier sprake van evidence? De onderzoekers berekenden standaarddeviaties waarmee ze aan de slag gingen. Omdat er echter geen *absoluut* criterium wordt gehanteerd, is de grens die wordt getrokken volstrekt willekeurig. Volgens zulk een benadering is natuurlijk altijd een groep scholen zwak - in dit geval 25% - en is er altijd een groep gemiddeld (50%) en een groep sterk (25%), óók als het gemiddelde rekenniveau van alle scholen heel hoog zou liggen en de spreiding gering is.

Is hier nu sprake van nieuw feitenmateriaal? Dat is *niét* het geval, want ondanks het gebruik van statistische procedures, is dit ‘feit’ niet meer dan een artefact die het resultaat is van de keuze voor een bepaalde onderzoeksmethode.

Onderzoeksvraag en onderzoeksmethode

Naar welke onderzoeksmethode moet de voorkeur uitgaan? Naar een hard experiment

met aselect samengestelde proef- en controlegroepen van leerlingen en leerkrachten? Of mogen we ook minder hard inzetten en het, bijvoorbeeld op basis van procesonderzoek, meer ‘in de diepte’ zoeken? Deze vraag is nauwelijks te beantwoorden als er niet vooraf duidelijkheid is over de vraag: wat wil je onderzoeken en te weten komen?

Evidence based onderzoek, hier geïnterpreteerd als een ‘hard’ onderzoeksdesign, is niet *beter of slechter* dan ander onderzoek. Of voor dit type onderzoek moet worden gekozen, hangt af van het soort vragen waarop men antwoord wil krijgen. Stel bijvoorbeeld dat we willen weten of het verantwoord is een programma in te voeren dat is gebaseerd op de gedachte dat jonge kinderen ‘abstract’ kunnen leren denken. Voor welk type onderzoek zouden we dan moeten kiezen? Moeten we vele honderden peuters in crèches aselect aan proef- en controlegroepen toewijzen en vergelijkend, constaterend onderzoek doen? Moeten we op basis van gestandaardiseerde testitems onderzoeken hoe peuters op de proefjes reageren en dit vervolgens statistisch verwerken? Of kiezen we voor een opzet waarin op basis van ‘case study’s’³ de kwaliteit en de kenmerken van het abstracte denken worden geanalyseerd? Dit gebeurde bijvoorbeeld in het onderzoek van Baroody e.a.⁴ waarin peuters (één en twee jaar) individueel werden onderzocht. Ze moesten hoeveelheden vergelijken zonder te letten op de kleur en vorm van voorwerpen. De vraag was of de kinderen verzamelingen van bijvoorbeeld drie rondjes en drie vierkanten al of niet als gelijke hoeveelheden beschouwden.

Gelet op de onderzoeksvraag - namelijk of peuters abstract kunnen denken - zou mijn voorkeur uitgaan naar explorerend onderzoek. Dat is onderzoek waarin de onderzoeker in interactie met individuele kinderen de mogelijkheid heeft om op de reacties van de kinderen in te spelen. Hij kan dat bijvoorbeeld doen door variaties aan te brengen in het materiaal en in de vragen en door ondersteunende hints te geven. Bij deze onderzoeksvraag past een klinische onderzoeksmethode in de traditie van Piaget. Dat neemt niet weg dat, voortbouwend op dit onderzoek, in een later stadium de uitvoering van een ‘harder’ opgezet onderzoek heel zinvol kan zijn.

Objectiviteit en relevantie

De methodoloog de Groot⁵ wijst in dit verband op een continuüm met aan de ene zijde *objectiviteit* en aan de andere zijde *relevantie*. Kwalitatief procesonderzoek, zoals het onderzoek naar de vraag of peuters abstract kunnen denken, is een voorbeeld van onderzoek dat in hoge mate wordt gekenmerkt door *relevantie*. Evidence based onderzoek is onderzoek dat vooral wordt gekenmerkt door eisen die doorgaans aan *objectiviteit* worden gesteld.

De Groot verbindt het begrip ‘objectief’ met het begrip ‘formeel’. Een toetsinstrument moet volgens hem in principe een ‘objectief instrument’ zijn. Als voorbeeld bespreekt

hij een toetsinstrument waarmee het inzicht van leerlingen in meetkunde wordt gemeten. 'Objectief' betekent dan dat er richtlijnen moeten zijn voor onder meer de toepasbaarheid (voor wie is het bedoeld?), voor de wijze van afnemen (experimentatie genoemd) en voor de scoring. De Groot operationaliseert het begrip objectief dus vooral aan de hand van eisen die gesteld moeten worden aan toetsen op grond waarvan vaardigheden worden geëvalueerd. Bovendien brengt hij 'objectief' in verband met begrippen als 'acceptabel', 'gewenst' en 'bepaalbaar'. De Groot meent dat in collegiale *discussie* in redelijke mate overeenstemming kan worden bereikt over de vraag wat als 'objectief' kan worden gezien.

Eenvoudig is dat laatste echter ook weer niet. Aan de hand van enkele voorbeelden demonstreert hij de spanning die zich kan voordoen tussen objectiviteit en relevantie. Uitgaande van het werk van Hiele bespreekt hij het toetsen in het meetkunde onderwijs, in het bijzonder het 'construeren' en 'bewijzen'. 'Construeren' is erg belangrijk, maar is lastig objectief te onderzoeken. Bij *construeren* doen de leerlingen iets, maar dat is moeilijk *objectief* te beoordelen: de tekeningen zijn soms onduidelijk, er worden correcties uitgevoerd en het is niet altijd helder wat er precies wordt gerepresenteerd. Echter, wat de leerlingen doen is wél *relevant*.

Hetzelfde geldt voor 'bewijzen'. Men kan natuurlijk de verschillende stappen die de leerling maakt vaststellen, maar dat is iets anders, zegt de Groot, dan evalueren of de leerling inzicht heeft in het bewijsproces als geheel. Dat laatste is juist relevant. Het is trouwens niet verwonderlijk dat voor vele *relevante* centrale psychologische begrippen (angst, nieuwsgierigheid) altijd zwakke operationele definities hebben bestaan, zegt de Groot. Toch zal iedereen erkennen dat dit basale psychologische processen zijn, die onderzocht dienen te worden.

Objectiviteit óf relevantie?

Vroeger waren er twee theorieën of scholen. Ten eerste was er de experimentele psychologie, die zich vooral in het laboratorium bezighield met allerlei metingen, terwijl velen zich afvroegen wat de betekenis van al dat meten was. De tweede school bestond uit niet-experimentele onderzoekers, die relevante problemen onderzochten, waar anderen weer moeilijk greep op konden krijgen vanwege het gebrek aan transparantie en objectiviteit. Toch is de tegenstelling tussen relevantie en objectiviteit niet onoverbrugbaar, meent de Groot. De ontwikkeling van bepaalde instrumenten laat dat zien en hij wijst in dit verband op het werk van de Utrechtse psycholoog Kouwer, bekend van zijn persoonlijkheidstheorie.

We moeten ons dus niet verliezen in de tegenstelling tussen objectiviteit en relevantie, al zal de spanning tussen deze twee begrippen telkens opnieuw een punt van discussie en serieuze zorg zijn. Smaling⁶ bepleit er in dit verband voor 'recht te doen'

aan het object van onderzoek en bespreekt richtlijnen die kunnen bijdragen aan betrouwbaarheid, validiteit en objectiviteit van kwalitatief onderzoek. Vergelijkbaar is het voorstel van Maxwell⁷ om recht te doen aan de 'complexity of local causality', uitgaande van de zogenoemde 'process theory' waarin wordt beoogd causale processen niet slechts te beschrijven, maar de complexiteit ervan te analyseren.

Voorstanders van evidence based onderzoek zouden een liberaler koers kunnen varen als het gaat om onderzoek naar processen die in collegiaal overleg als relevant worden beoordeeld. Procesonderzoekers zouden ernaar kunnen streven om hun kwalitatieve gegevens zodanig te organiseren, dat deze geschikt worden gemaakt voor kwantitatieve bewerking.

- 1 Kagan, J. (2006). *An Argument for Mind*, p. 228 en p. 206. New Haven, London: Yale University Press
- 2 Onderwijsinspectie (2008), www.onderwijsinspectie.nl
- 3 Verschaffel, L. (2008). Lezing Utrecht
- 4 Baroody, A.J., Xia Li & Meng-lung Lai (2008). Toddlers' Spontaneous Attention to Number. In: *Mathematical Thinking and Learning*, p. 240-270
- 5 Groot, A.D. de (1961). *Methodologie, Grondslagen van onderzoek en denken in de gedragswetenschappen*. 's-Gravenhage: Mouton & Co.
- 6 Smaling, A. (1990). Enige aspecten van kwalitatief onderzoek en het klinisch interview. In: *Tijdschrift voor nascholing en onderzoek van het reken-wiskundeonderwijs*, 8 (3), 4-10
- 7 Maxwell, J.A. (2004). Causal Explanation, Qualitative Research, and Scientific Inquiry in Education. In: *Educational Researcher*, 33 (2), 3-11

4 Praktijkgerichte kennisontwikkeling

Alfons ten Brummelhuis | Kennisnet

De kennisparadox illustreert dat de mismatch tussen onderzoek en praktijk besloten ligt in de kwaliteitseisen die we hanteren voor onderzoek dat tegemoet komt aan de kennisvragen van de professionals uit de onderwijspraktijk. Om deze situatie te doorbreken, voert Kennisnet op het gebied van informatie- en communicatietechnologie (ict) onderzoeksactiviteiten uit die er op gericht zijn kennis te produceren die het onderwijs in de praktijk helpt meer effectief gebruik te maken van ict voor leren. Deze bijdrage schetst in het kort de achtergrond, aanpak en tot nu toe behaalde resultaten.

In de afgelopen jaren zijn verschillende analyses uitgevoerd naar de kloof tussen onderzoek en onderwijspraktijk. De moeizame relatie tussen onderzoek en praktijk is in essentie te duiden als een kennisparadox. Met een kennisparadox wordt bedoeld dat er steeds meer kennis wordt ontwikkeld en over het onderwijs wordt uitgestrooid, terwijl de benutting van deze kennis in de praktijk amper toeneemt. Dit is overigens geen indicatie voor gebrek aan wetenschappelijke kwaliteit. Integendeel, de Nederlandse wetenschappelijke output behoort tot de wereldtop¹. Het probleem is dat de beschikbare kennis buiten de kring van wetenschappers amper toegankelijk is en weinig tegemoet komt aan de kennisvragen uit de onderwijspraktijk. Het gevolg is dat het stimuleren van kennisproductie leidt tot een steeds grotere kennisvoorraad in plaats van dat het bijdraagt aan afname van de kennisbehoefte in de praktijk.

Uitgangspunten

- *Scholen zijn vraageigenaar*
Vertrekpunt voor praktijkgerichte kennisontwikkeling zijn de ambities van scholen om onderwijs te verbeteren. Het uitvoeren van onderzoek is geen doel op zich maar een middel om te komen tot (nieuwe) kennis die scholen in staat stelt gefundeerde keuzes te maken. Dit betekent dat de school 'eigenaar' is van de vraagstelling en daarmee klant van kennisontwikkeling. Het gaat om onderzoek van en voor het onderwijs.
- *Bruikbaarheid is kwaliteit*
Er is geen methodische norm die richtinggevend is voor de kwaliteit van praktijkgericht onderzoek. Een experimenteel design is op voorhand niet beter dan een case studie. Elk wetenschappelijk paradigma kan van waarde zijn. Kwaliteit van onderzoek wordt afgemeten aan de fit tussen onderwijsvraag en de methodische aanpak die in de gegeven context de vraageigenaar de meest bruikbare kennis oplevert.
- *Valorisatie en cumulatie van kennis*
Om te voorkomen dat onderzoek wordt gedaan naar wat al bekend is of verschillende scholen hetzelfde wiel uitvinden, is een programmatische aanpak vereist.

Deze aanpak is gericht op verspreiding en benutting van bestaande kennis en systematische kennisopbouw zodat er een samenhangend geheel ontstaat van kennisvalorisatie en -cumulatie.

Werkwijze

Bovengenoemde uitgangspunten zijn vertaald in de volgende werkwijze.

- *Onafhankelijke vraagverheldering*

Vraagverheldering is gericht op beantwoording van de vraag: wat wil je precies met je onderwijs of (nieuw) leerarrangement bereiken. In termen van kennis gaat het heel praktisch om inzichten die een school nodig heeft om te besluiten een bepaalde onderwijsaanpak te continueren, op te schalen, bij te stellen of te stoppen. In de werkwijze die Kennisnet hanteert, is er voor gekozen de ondersteuning die scholen wordt geboden voor de vraagverheldering te scheiden van de onderzoeksrol. Dit om te voorkomen dat inhoudelijke of methodische voorkeuren van onderzoekers de vraagstelling van scholen onbedoeld beïnvloeden. Deze fase is het meest cruciaal voor borging van de latere kennisbenutting in de praktijk.

- *Onderzoeksaanpak afstemmen op vraagstelling*

De relatie tussen kenmerken van een onderzoeksvraag en de opzet van onderzoek is toegelicht in tabel 1. De tabel toont een opbouw in hardheid van bewijs. Verder illustreert de tabel dat lang niet elke kennisvraag geschikt is voor de meest harde bewijsvoering. Dit betekent dat als de meest harde bewijsvoering als norm wordt gezet om evidence based onderwijs² te stimuleren, er veel onderwijsvragen onbeantwoord blijven. De focus van kennisproductie ligt dan eenzijdig op onderwerpen die al uitgekristalliseerd zijn in de onderwijspraktijk. Kennisopbouw over innovatieve onderwijsbenaderingen die per definitie in de aanvangsfase weinig helder zijn, blijven dan onderbelicht. Zij vereisen immers uitvoering van onderzoek met een mindere graad van hardheid.

- *Benutten van resultaten*

Doorslaggevend voor de verspreiding en de benutting van resultaten uit onderzoek is de aansluiting bij wat scholen in de praktijk willen bereiken. Het meest efficiënt is eerst de beschikbare kennis over wat wel en niet werkt te bundelen en te verbinden met vraagstukken van scholen³. Indien beschikbare kennis tekort schiet is er aanleiding voor nieuwe kennisontwikkeling.

Kenmerken vraag	Kenmerken onderzoek
1. Systematisch in kaart brengen van onderwijsaanpak	1. Procesbeschrijving: good practice - case studie
2. Uitproberen van een onderwijsidee en al doende verbeteren	2. Procesbenadering waarbij onderzoeker tevens co-ontwerper is: ontwerpgericht onderzoek
3. Van een praktijksituatie willen weten hoe anderen denken over kwaliteit, opbrengsten en kansrijkheid	3. Gepercipieerde opbrengsten (kwantitatief en/of kwalitatief) verzameld in survey of panelstudie
4. Inzicht in vooraf gedefinieerde opbrengsten van een uitgekristalliseerde onderwijssituatie	4. Meting achteraf van vooraf verwachte opbrengsten: evaluatie onderzoek
5. Inzicht in vooraf gedefinieerde opbrengsten van een uitgekristalliseerde onderwijssituatie in vergelijking met 'regulier' onderwijs	5. Voor- en nameting in (quasi) experimenteel design: effect meting

Tabel 1: Classificatie van onderwijsvraagstukken en onderzoeksbenaderingen

Resultaten

Uit evaluatie van projecten die volgens bovengenoemde kenmerken in de afgelopen jaren zijn uitgevoerd, komt naar voren dat bijna negen van de tien scholen van mening zijn dat gegenereerde kennis voldoet aan de behoeften van scholen en bijdraagt aan de oplossing van het oorspronkelijke probleem van de school⁴. Verder is de oriëntatie op (leer)resultaten in het primaire proces kenmerkend voor deze praktijkgerichte kennisontwikkeling⁵. Op de website onderzoek.kennisnet.nl zijn videofragmenten te vinden waarin onderzoekers en scholen een impressie geven van het uitgevoerde onderzoek en de behaalde resultaten. Op deze website is eveneens een overzicht te vinden van onderzoek dat momenteel in uitvoering is. De ervaringen tot nu toe leren dat praktijkgerichte kennisontwikkeling voorziet in een behoefte en scholen in de gelegenheid stelt hun ambities beter en sneller te realiseren. Responspercentages van minimaal 80%⁸ en responsaantallen van ruim 8000 docenten⁶ weerleggen dat het onderwijsveld onderzoeksmoe is. De resultaten onderstrepen dat de behoefte aan kennis die aansluit bij de vragen van het onderwijs groot is evenals de bereidheid van scholen om aan dergelijk onderzoek mee te werken.

Toekomst

Onderwijs dat uitsluitend is gebaseerd op bewezen werking heeft als risico dat de nadruk komt te liggen op instandhouding van beproefde werkwijzen uit het verleden en weinig perspectief biedt voor innovatie. Immers, vernieuwing van onderwijs is niet mogelijk zonder het risico te nemen van succes of falen. Innovatie is per definitie onvoorspelbaar⁷. De uitdaging voor de toekomst is dan ook naast de benutting van beproefde kennis nieuwe kennis te ontwikkelen en aan te wenden voor het oplossen van hardnekkige vraagstukken die spelen in het onderwijs.

Dit vraagt op de eerste plaats meer aandacht voor de potentiële gebruikers van kennis in het onderwijs en een accentverschuiving in kennisontwikkeling. Minder geredeneerd vanuit de onderzoeker als producent van kennis en meer vanuit de school als ontvanger en gebruiker van kennis. De ervaringen die Kennisnet heeft opgedaan met praktijkgerichte kennisontwikkeling op het domein van leren met ict zijn naar verwachting breder toepasbaar in het onderwijs⁸. Een impuls voor benodigde (basis) faciliteiten en voorzieningen kan zo'n proces van praktijkgerichte kennisproductie versterken. Dit zal de kwaliteit en productiviteit van het onderwijs ten goede komen en tegelijkertijd het onderwijsveld versneld helderheid verschaffen over de bijdrage van ict bij het oplossen van belangrijke onderwijsvraagstukken zoals taalachterstanden, voortijdig schoolverlaten, lerarentekort en talentenontwikkeling van leerlingen.

1 Brief aan de Tweede Kamer over kennisinvesteringen (2008). Kamerstuk 22-10-2008

2 Zie bijvoorbeeld www.onderwijsbewijs.nl

3 Zie publicaties in de Kennisnet onderzoeksreeks: Wat weten we over..... Beschikbaar via onderzoek.kennisnet.nl

4 Graaff, R. de, Akkerman, S.F. & Berg, I. van den (2008). Effecten in kaart. Utrecht: IVLOS

5 Meijer, J. & Eck, E. van (2008). Leren met meer effect: rapportage van het onderzoek. Amsterdam: SCO-Kohnstamm Instituut

6 Gennip, H., Rens, C. van, Smeets, E. (2008). Didactiek in Balans. Nijmegen: ITS.

7 WRR (2008). Innovatie vernieuwd (www.wrr.nl)

8 Zie: Leren met meer effect, beschikbaar via onderzoek.kennisnet.nl

5 Allemaal aan de RCT?

Marian de Graaf, Hans Meij en Ton van Yperen | Nederlands Jeugdinstituut

De databank Effectieve Jeugdinterventies van het Nederlands Jeugdinstituut maakt kennis over effectieve jeugdinterventies toegankelijk, zodat hulpverleners en cliënten kunnen profiteren van de meest recente inzichten uit wetenschap, onderzoek en praktijk. Steeds vaker wordt ervoor gepleit dat professionals in jeugdzorg en onderwijs uitsluitend evidence based zouden moeten werken. Maar hoe zit het met hun autonomie? En wat is de rol van de beroepsverenigingen?

De laatste jaren is de roep om effectieve jeugdzorg vooral vertaald naar de stelling dat hulpverleners ‘dus’ evidence based moeten werken. Die stelling is zijn eigen leven gaan leiden. We zouden allemaal aan de Randomized Controlled Trials¹ (RCT’s) moeten, er zou alleen nog maar met evidence based interventies gewerkt mogen worden en van de professional zou niet veel meer overblijven dan een robot².

Enkele jaren geleden werd door het Nederlands Jeugdinstituut de databank Effectieve Jeugdinterventies in het leven geroepen. Het oorspronkelijke doel was het bevorderen van de effectiviteit van de jeugdzorg. De databank maakt kennis over effectieve jeugdinterventies in Nederland toegankelijk voor professionals, managers en beleidsmakers, zodat hulpverleners en cliënten kunnen profiteren van de meest recente inzichten uit wetenschap, onderzoek en praktijk. Inmiddels is de databank flink uitgebreid met interventies uit andere sectoren, waaronder het onderwijs. En in de sector van het onderwijs woedt nu dezelfde discussie³. Ook hier dreigt het gevaar van doorschieten. Want waar gaat het eigenlijk om? Wij zullen deze vraag beantwoorden vanuit drie gezichtspunten: (1) de functie van de databank Effectieve Jeugdinterventies om een gezonde opwaartse druk te creëren, (2) de autonomie van de professional en (3) de noodzaak van een collectieve ambitie.

Een gezonde opwaartse druk

In de databank Effectieve Jeugdinterventies is veel kennis samengebracht over wat wel werkt en wat niet, minder of averechts werkt. Met deze kennis kan de professional zijn voordeel doen. Zo is de leerkracht erbij gebaat te weten dat straffen niet helpt, maar positieve controle (positief gedrag belonen en negatief gedrag negeren) wel.

De kennis in de databank Effectieve Jeugdinterventies over wat werkt is verpakt in interventies die ‘in theorie effectief’ of ‘bewezen-effectief’ zijn. In veel discussies wordt ‘bewezen-effectief’ gelijk gesteld met bewijs uit RCT’s. De databank is echter veel genuanceerder opgebouwd. Bij de beoordeling van interventies wordt een beoordelingssysteem gebruikt dat verschillende niveaus van evidentie van effectiviteit erkent en honoreert. Er zijn veel soorten onderzoek die de werkzaamheid van een interventie kunnen laten zien. De ene soort onderzoek biedt meer zekerheid over

de werkzaamheid van een interventie dan de andere. Voor de databank is dat geen reden om minder hard bewijs te negeren. Bewijs met een lage zekerheidsgraad is ook een indicatie voor effectiviteit. Het beoordelingssysteem van de databank Effectieve Jeugdinterventies werkt als een ladder: hoe groter de zekerheid dat de interventie werkt, hoe hoger de interventie op de ladder staat. De boodschap van de databank is om zo hoog mogelijk op die effectladder te klimmen en niet te snel genoeg te nemen met alleen onderzoek naar de tevredenheid van cliënten of onderzoek zonder vergelijkingsgroepen.

Door te laten zien op welke sport van de effectladder een interventie staat, creëert de databank een 'opwaartse druk'. Deze druk wordt bekrachtigd doordat een erkende interventie elke vijf jaar opnieuw wordt beoordeeld. Bij zo'n herbeoordeling wordt niet alleen gekeken in hoeverre de interventie nog up-to-date is, maar ook of de interventie al een stapje hoger is gekomen op de ladder.

De jeugdinterventies in de databank zijn concrete, geprotocolleerde programma's. Waar het echter om gaat zijn de principes die maken dat de interventie werkt. Het mag geen trucje zijn. Er moet minstens een theorie zijn, die aannemelijk maakt dat de interventie voor de betreffende doelgroep werkt. Overigens zet het NJi ook daarom per doelgroep en per werksoort op een rij wat volgens onderzoek werkzame principes zijn. Wat werkt er wel en niet bij de aanpak van pesten op school, bij het voorkomen van schooluitval of bij het bestrijden van onderwijsachterstanden? Het is een enorme klus om dat op een rij te zetten. Op www.jeugdinterventies.nl ziet u daar nu en de komende tijd de resultaten van.

De bedoeling van de Databank Effectieve Jeugdinterventies is dus om in de jeugdsector de aandacht te vergroten voor vakkennis over wat er wel en niet werkt. Belangrijk uitgangspunt en ideaal daarbij is dat de professional zoveel mogelijk gebruikmaakt van die kennis. De databank zorgt intussen ook voor een 'opwaartse druk' om omhoog te klimmen op de effectladder.

Autonomie en monitoring

Wat moet de professional nu concreet met de interventies die in de databank staan? Het uitgangspunt is dat de professional zelf verantwoordelijk is voor het al dan niet gebruiken van interventies uit de databank. Hij moet er bijvoorbeeld zelf voor kiezen of hij in het kader van de voor- en vroegschoolse educatie gebruikmaakt van Opstap, Kaleidoscoop of Sporen. Het is verstandig dat hij zich zoveel mogelijk houdt aan een werkwijze waarvan is aangetoond dat de kans op effect groot is. De professional kan echter redenen hebben om het bij een bepaalde cliënt of groep cliënten anders te doen. Dat hoeft geen probleem te zijn, maar zijn professionele autonomie verplicht hem dan wel om na te gaan of cliënten zijn gebaat bij zo'n eigen koers. Om daarachter

te komen, moeten de afwijkende werkwijze en de resultaten daarvan worden vastgelegd. Vergelijking met collega's die met dezelfde doelgroep werken, kan aanwijzingen opleveren over de wijze waarop het (nog) beter kan.

In de jeugdzorg worden er op dit moment op veel plaatsen monitors gebouwd die hulpverleners zicht bieden op de effectiviteit van hun praktijk. Dat werkt heel stimulerend. Zo bleek uit een monitor van Intensieve Orthopedagogische Gezinsbehandeling dat in vijf jaar tijd de effectiviteit van het werk van de hulpverleners meer dan verdubbelde, terwijl de behandelduur afnam. Wanneer praktijkwerkers het niet zo nauw nemen met belangrijke behandelprincipes, confronteert dit type monitor hen met de consequenties daarvan.

Kortom: als we het over effectief werken hebben, gaat het er niet om dat de professional de 'robot' wordt die erkende interventies uitvoert uit bijvoorbeeld de databank Effectieve Jeugdinterventies. De professional heeft de autonomie om te beslissen wat hij met welke cliënten het beste kan doen en heeft de deskundigheid om hierover te oordelen. Maar dat verplicht hem ook om te kijken of hij met zijn aanpak de beste resultaten boekt. Autonomie en monitoring horen bij elkaar. Met andere woorden: effectief werken staat niet haaks op professionele deskundigheid en autonomie, maar doet daarop juist een beroep.

Noodzaak van een collectieve ambitie

Het bestaan van de databank Effectieve Jeugdinterventies veroorzaakt bij sommigen de vertwijfeling dat er straks alleen nog maar evidence based interventies gefinancierd zullen worden. Er zijn verschillende redenen waarom dat niet goed zou zijn. We gaan nu niet op al die redenen in, maar er is er één die in dit verband aandacht verdient. Het is niet aan de financier of de beleidsmaker om vast te stellen welke interventie een professional moet uitvoeren. Financiers en beleidsmakers mogen niet op de stoel gaan zitten van de professional. Maar ... zij moeten ook niet op die stoel *hoeven* te gaan zitten.

Om dat te voorkomen kunnen de beroepsverenigingen een rol spelen. Hun stem missen wij namelijk in dit debat. De beroepsverenigingen zouden veel meer een eigen verantwoordelijkheid moeten nemen voor de opwaartse druk in effectief werken. Zij kunnen dat doen door van professionals te eisen dat zij volgens de beste stand van kennis werken en zich regelmatig laten bijscholen in de wat-werkt-principes. Als deze ambitie bij beroepsverenigingen ontbreekt, kunnen financiers en beleidsmakers in het gat stappen, en dat is jammer. In de medische sector trekken beroepsverenigingen dat veel meer dan in de jeugdsector naar zich toe. Zo constateerde de inspectie onlangs dat er veel mis is met de hygiëne in operatiekamers. Eerder bleek in een onderzoek naar het optreden van post-operatieve infecties dat tijdens bevallingen

met een keizersnede de deur van de operatiekamer soms wel 50 tot 70 keer open en dicht ging. Naar aanleiding daarvan heeft de beroepsvereniging van de medici een richtlijn opgesteld om de hygiëne in operatiekamers te vergroten. In de sectoren van jeugdzorg en onderwijs zouden de beroepsverenigingen ook zo actief moeten zijn.

Samenvattend

Jeugdigen en hun ouders hebben het recht om het beste te krijgen wat er aan kennis beschikbaar is. De laatste jaren is dat vooral vertaald naar de stelling dat hulpverleners 'dus' evidence based moeten werken. Waar het echter om gaat is dat er een gezonde opwaartse druk moet zijn in de ontwikkeling van de effectiviteit van het aanbod. De Databank Effectieve Jeugdinterventies helpt om die druk vorm te geven. Het is een etalage van interventies die goed zijn en laat onder meer zien op welke sport van de effectladder die interventies staan. 'Doe er je voordeel mee', is de boodschap.

De professional is autonoom in de beslissing om interventies of kennis uit die databank wel of niet te gebruiken, maar die autonomie verplicht hem om de resultaten van zijn handelen zichtbaar te maken en te vergelijken met de resultaten van anderen. Financiers of beleidsmakers die alleen nog maar evidence based interventies willen financieren, gaan op de verkeerde stoel zitten. In plaats daarvan zouden beroepsverenigingen veel duidelijker moeten kiezen voor de collectieve ambitie om het aanbod - via opwaartse druk - zo effectief mogelijk te maken.

- 1 Een Randomized Controlled Trial is een experimenteel onderzoek, waarbij de onderzoeksgroep wordt vergeleken met een controlegroep op basis van een voormeting, een nameting en een meting enige tijd na afloop van de interventie. Idealiter zijn de onderzoeksgroep en de controlegroep willekeurig samengesteld.
- 2 Voor een overzicht van de discussie in de jeugdzorg, zie: Faas, M. (2008). Effectiviteit, evidence based en de jeugdzorg (deel 2). De jeugdzorger als robot? In: Sozio, 84; 14-19
- 3 Gravemeijer, K.P.E. & Kirschner, P.A. (2007): Naar meer evidence based onderwijs? In: Pedagogische Studiën, 84; 463-472; Ros, B. (2008). Het beste bewijs? In: Didaktief, 8; 4-7

6 Naar structurele coalities van scholen en onderzoekers

Kars Veling | directeur Johan de Witt scholengroep

In het Talentencentrum van de Johan de Witt scholengroep worden onderwijspraktijk en wetenschappelijk onderzoek met elkaar verbonden met het doel manieren te vinden om leerlingen op te leiden tot een diploma op een zo hoog mogelijk niveau. Het ontwerpen vervult in dit proces een onmisbare, verbindende functie. Zonder dat verbeteringen beargumenteerd worden uitgetoet, zullen onderwijspraktijk en onderzoek elkaar maar moeilijk vinden. Onderwijs, ontwerp en onderzoek zijn in een cyclus met elkaar verbonden.

In de vele discussies over de kwaliteit van het onderwijs lijkt één opvatting onomstreden te zijn, namelijk dat onderwijs (meer) moet worden gebaseerd op wetenschappelijke kennis. Van vele kanten wordt gepleit voor 'evidence based education'.

Onderwijs Bewijs

Prachtig. Meer geld voor onderwijsonderzoek is wat mij betreft welkom. Maar meer wetenschappelijk onderzoek levert niet vanzelfsprekend meer wetenschappelijk gefundeerd onderwijs op. Een hardnekkig probleem is dat de resultaten van onderzoek moeilijk in de scholen doordringen. De verspreiding van wetenschappelijke kennis in het onderwijs is moeilijk te realiseren. Het subsidieprogramma *Onderwijs Bewijs* heeft daarom de doelstelling "(meer) verbindingen te leggen tussen het onderwijsveld en wetenschappelijke onderzoek." Een projectvoorstel wordt daarom alleen maar aanvaard wanneer het is ingediend door een consortium waarin tenminste één onderwijsinstelling participeert.

De door het ministerie opgestelde leidraad voor indieners van plannen maakt echter ook meteen duidelijk waarom deze doelstelling zo lastig is te realiseren. Het is de bedoeling van *Onderwijs Bewijs* dat "eenduidig kan worden vastgesteld" wat wel en niet werkt. Dit vereist onderzoek dat statistisch houdbare conclusies oplevert over oorzaken en gevolgen. De leidraad stelt: "Om een oorzakelijk effect aan te tonen zijn zuivere experimenten (met aselechte toewijzing) het meest geschikt." Alleen in uitzonderingsgevallen maken aanvragen die afwijken van deze opzet een kans op honorering.

Deze gedachtegang is helder en plausibel. Maar we weten dat aan experimentele onderzoeken met generaliseerbare uitkomsten veel eisen worden gesteld. Zo zijn er veel proefpersonen nodig, ook al omdat ze - het liefst door middel van loting - moeten worden ingedeeld in een experimentele groep en een controlegroep. Om voldoende data te kunnen verzamelen, zijn doorgaans meerdere scholen nodig. Deze scholen moeten in staat en bereid zijn om met een aselekt gekozen groep leerlingen

een bepaalde interventie uit te voeren, het liefst zo dat de leerlingen niet weten dat ze in de experimentele groep zitten. Leerlingen van zowel de experimentele groep als van de controlegroep moeten meewerken aan enquêtes of tests. Een experimenteel onderzoek vraagt dus veel van scholen. Ik vind het dan ook niet verwonderlijk dat onderwijsonderzoekers erover klagen dat scholen steeds minder bereid zijn om aan onderzoek mee te werken.

Daar komt bij dat het niet zo duidelijk is dat scholen die aan een onderzoek meewerken daar zelf direct beter van worden. De doorlooptijd van een onderzoek is doorgaans langer dan de termijn waarop een school zou willen leren van opgedane ervaringen. Verder noopt een goed opgezet onderzoek tot het preciseren en daarmee tot het isoleren van variabelen. Niet zelden levert dit proces bij onderwijsmensen een gevoel van vervreemding op. Is het zinvol onderwijs te onderzoeken wanneer bijvoorbeeld 'de motivatie van leerlingen' of 'de pedagogische competentie van docenten' wordt geabstraheerd?

Hiermee is natuurlijk niet gezegd dat experimenteel onderwijsonderzoek geen belangrijke functie kan vervullen. Kennis van samenhang tussen onderwijsinterventies en leereffecten kan zeker bijdragen aan de verbetering van het onderwijs op de langere termijn. Maar in en rondom de scholen moet meer gebeuren.

Talentencentrum

Binnen de Johan de Witt scholengroep wordt sinds een aantal jaren gebouwd aan een structurele samenwerking tussen onderwijspraktijk en wetenschap. Samen met de Universiteit Leiden (in het bijzonder professor Paul Vedder en later ook professor Mien Segers) hebben we een Talentencentrum ingericht met het doel effectieve manieren te vinden om leerlingen op te leiden tot een diploma op een zo hoog mogelijk niveau. In het innovatieproject *Durven, delen, doen* van de VO-raad heeft de Johan de Witt scholengroep de mogelijkheid gekregen om het Talentencentrum binnen de school te verbreden en is voor een aantal jaren ook het SCO-Kohnstamm Instituut hierbij betrokken.

De samenwerking tussen school en onderzoekers krijgt gaandeweg een steeds duidelijker structuur. In het begin vergeleek ik het Talentencentrum nogal eens met een laboratorium, maar goed beschouwd is deze metafoor niet bruikbaar. In een laboratorium worden proefopstellingen ingericht met de bedoeling wetenschappelijke vragen te beantwoorden. Een school heeft echter een missie die niet is gericht op het bijdragen aan wetenschappelijke kennis. Onderwijs is er voor de leerlingen! Het is dan ook beter het Talentencentrum te typeren als een *atelier*, als een werkplaats waar onderwijs wordt 'gemaakt' in een combinatie van ambachtelijkheid en experimentele openheid.

Praktijk, ontwerp en onderzoek

In een atelierschool (wat uiteraard ook een samenwerkingsverband van scholen kan zijn) worden onderwijspraktijk en wetenschappelijk onderzoek verbonden. Cruciaal is daarbij een derde functie, namelijk die van het *ontwerpen*. Onderwijs, ontwerp en onderzoek zijn alledrie onmisbaar in een school die zichzelf ontwikkelt. De onderwijspraktijk vormt de basis en levert ook de drive om te komen tot reflectie en verbetering. Het onderzoek zorgt voor de kritische functie: Doen we wat we zeggen? En bereiken we ook echt wat we beogen? De noodzakelijke verbindende functie is die van het ontwerpen, het beargumenteerd inrichten van mogelijke verbeteringen in de onderwijspraktijk. Zonder deze functie zullen onderwijspraktijk en onderzoek elkaar moeilijk vinden. De onderwijspraktijk heeft immers sores genoeg om daarnaast nog aandacht te hebben voor abstract wetenschappelijk onderzoek. En onderzoekers zullen gauw teleurgesteld zijn in scholen die zich niet als laboratorium willen laten gebruiken.

Onderwijs, ontwerp en onderzoek moeten in een cyclus met elkaar worden verbonden. Nieuwsgierige onderwijzers ervaren een probleem en staan open voor verandering. Ontwerpers (natuurlijk kunnen dit onderwijzers en onderzoekers zijn) helpen om te bedenken wat wellicht tot verbetering zou kunnen leiden. Onderzoekers volgen of een verandering het beoogde effect heeft. De terugkoppeling van de bevindingen naar de onderwijspraktijk kan weer nieuwe ideeën oproepen. Enzovoort.

In het kader van het Talentencentrum zijn meerdere ontwikkelingen via deze cyclus in gang gezet. Door een verbetering van de intake willen we de plaatsing van leerlingen in een type brugklas verbeteren. Door klassen met een specifiek programma in te richten willen we proberen leerlingen te laten schakelen naar een hoger niveau. Door POP-gesprekken met leerlingen willen we hen helpen om kansen beter te benutten. Door acties ter verbetering van het pedagogische klimaat willen we bereiken dat leerlingen zich meer uitgedaagd en gesteund voelen. Door tests en extra leerstof willen we ervoor zorgen dat leerlingen die meer kunnen ook meer gaan presteren. Of dat allemaal lukt? We gaan het zien. En we gaan leren van onze ervaringen.

Verbindingen op niveau van onderzoeksprogramma's

Het is goed om verschillende typen onderwijsonderzoek van elkaar te onderscheiden. Het hierboven bedoelde monitoronderzoek heeft andere kenmerken dan experimenteel onderzoek dat is gericht op generaliseerbare stellingen over oorzaken en gevolgen. Deze constatering roept echter ook een belangrijke vraag op. Gaat het onderscheiden van deze twee typen onderzoek niet opnieuw ten koste van het gebruik van wetenschappelijke kennis in het onderwijs? Dat zal zeker het geval zijn wanneer onderwijsonderzoekers zich in twee verschillende klassen of - erger nog - kampen gaan verdelen, één voor fundamenteel en één voor toegepast onderzoek. Dan drijven de drie voor onderwijsontwikkeling noodzakelijke functies gemakkelijk

weer uit elkaar. De eigen kritische inbreng van wetenschappelijk onderzoek kan dan verschrompelen, omdat de cyclus van ontwerp en monitoronderzoek de verbinding met het fundamentele onderzoek kwijtraakt.

Naar mijn mening hebben de landelijke pedagogische centra te kampen met dit probleem. Zij opereren dichtbij de scholen. Zij hebben vanouds de opdracht de functies van ontwerp en onderzoek te vervullen, maar in de praktijk houden zij zich vooral bezig met de taken begeleiding, advies en scholing. De afstand tot de universitaire onderzoeksinstellingen met wetenschappelijke ambities is tamelijk groot.

Prof. Monique Volman pleit in het artikel in *Onderzoek de school in!?* terecht voor structurele samenwerking in wat zij aanduidt als de “driehoek van schoolteam, onderzoekers en begeleiders/adviseurs”. Een belangrijk argument hiervoor is dat een deugdelijke toetsing van de effectiviteit van onderwijsinterventies bijna altijd meerdere jaren kost. Het zou kapitaalsvernietiging zijn om na een te korte periode weer ergens opnieuw te beginnen.

Er is ook een wetenschapstheoretisch argument voor de door Volman bepleite structurele samenwerking. Experimenteel onderzoek is gericht op de toetsing van een veronderstelde samenhang van variabelen. Maar wetenschappelijke vooruitgang vraagt om meer dan het sprokkelen van significante correlaties. Onderzoek krijgt wetenschappelijke betekenis in het kader van een *onderzoeksprogramma*. Zo'n onderzoeksprogramma genereert zinvolle vraagstellingen en maakt het ook mogelijk de uitkomsten van empirisch onderzoek te duiden.

Het constituerende belang van onderzoeksprogramma's vertoont een parallel met het belang van samenhang in onderwijsontwikkeling voor de onderwijspraktijk. Scholen die 'maar eens wat proberen' komen niet ver. Onderzoekers die wachten op de opdracht om de effecten van zo'n probeersel te meten, voegen waarschijnlijk niet veel toe, zelfs niet als ze aan voldoende data kunnen komen.

Mijn stelling is dat het voor het onderwijsonderzoek, het onderwijsontwerp en de onderwijspraktijk² de moeite loont om een verbinding te maken op het niveau van onderzoeksprogramma's respectievelijk onderwijsontwikkelingsprogramma's. Zulke verbindingen vormen een stevige basis voor meerjarige samenwerking. Korte termijn monitoronderzoek kan dan ook data opleveren voor een langere termijn onderzoek. Suggesties voor het ontwerp van onderwijsverbetering krijgen een steviger basis doordat ze zijn ingebed in een onderzoeksprogramma dat de moeite waard is.

Voorwaarden

In wetenschapstheoretische termen kan een onderzoeksprogramma annex onderwijsontwikkelingsprogramma gebaseerd zijn op twee uitgangspunten, te weten een

specifieke *constituering* van het onderwijs als object van onderzoek en ontwikkeling, en een specifiek *kennisideaal*, dat wil zeggen: een opvatting over de richting waarin wordt gezocht naar variabelen die relevant zijn voor kennisverwerving en voor verbetering van de onderwijspraktijk³. Het voert te ver om hier de mogelijkheden voor dergelijke onderzoeks- en ontwikkelingsprogramma's te verkennen. Misschien zijn er aanknopingspunten te vinden in een grondgedachte die binnen ons Talentencentrum groeit. De ontwikkeling van de school en ook het daaraan gelieerde onderzoek lijkt samenhang te vinden in de leerontwikkeling van kinderen (wat we in de loop van de tijd *leerbiografie* zijn gaan noemen). Kenmerkend voor deze interesse in de leerbiografie, is de aandacht voor de tijdsdimensie. Naast de constituering van het object, zijn we ook gezamenlijk geïnteresseerd in *schoolse interventies* die beogen de leerbiografie te optimaliseren. De al eerder genoemde acties die binnen het Talentencentrum zijn geïnitieerd (intake, POP, leerlijnen, schakelmogelijkheden, leerlijnen, pedagogisch klimaat), passen in dit kader.

Meerjarige coalities van onderzoeksinstituten en scholen kunnen veel betekenen, maar dan wel onder een aantal voorwaarden. Scholen - dat wil zeggen docenten en schoolleiders - moeten bereid zijn hun onderwijspraktijk serieus ter discussie te stellen. Scholen en onderzoeksinstituten moeten gezamenlijk investeren in het inrichten van een cyclus van onderwijspraktijk, ontwerp en onderzoek. En onderzoeksinstituten moeten al bij het bedenken van onderzoeksprogramma's verbinding maken met onderwijsinstellingen, zodat monitoronderzoek en grootschaliger experimenteel onderzoek elkaar kunnen versterken.

- 1 Een bundel artikelen (uitgave VO-raad Utrecht, 2008) naar aanleiding van een conferentie, georganiseerd door de VO-raad met de KPC-groep en ADEF.
- 2 Voor alle duidelijkheid zij opgemerkt dat de genoemde verschillende functies uiteraard in organisaties en in personen verenigd kunnen zijn.
- 3 Voor een fundering van deze gedachte zie: Veling, K. (1983). *Methodologie en de grondslagen van een pluriforme sociologie*. Assen: Van Gorcum

7 De noodzaak van de onderzoekende leerkracht

Monique de Haan | basisschool Het Baken, Veenendaal

Het is in het onderwijs nog niet vanzelfsprekend dat men onderzoek doet naar wat werkt en waaróm iets werkt. Dat gebeurde wel op basisschool Het Baken in Veenendaal, waar vergelijkbare groepen kinderen de leerstof volgens verschillende onderwijsconcepten kregen aangeboden. Een uitgelezen situatie om de effecten van verschillende onderwijsvisies op een wetenschappelijk verantwoorde manier naast elkaar te zetten. In de praktijk blijkt dat echter niet eenvoudig te zijn. Wat moet er in de school gebeuren om evidence based te werken?

De wet op het basisonderwijs (1985) staat een basisschool voor met doorlopende leer- en ontwikkelingslijnen voor kinderen van vier tot en met twaalf jaar. Toch is er op veel basisscholen tussen de groepen 2 en 3 sprake van een kloof in werkwijze, aanbod en benadering. Vaak wordt zowel in de groepen 1 en 2 als in de groepen 3 tot 8 nog vrij traditioneel onderwijs gegeven. Dat wil zeggen dat in groep 1 en 2 procesgericht en in de groepen 3 tot 8 methodisch en productgericht wordt gewerkt. Met name de overgang van groep 2 naar groep 3 - de voormalige grens tussen kleuterschool en lagere school - is in de praktijk én in de wetenschap onderwerp van gesprek en onderzoek.

Doorgaande lijn

Op veel basisscholen zijn er kinderen die niet aan de leerstof in groep 3 kunnen beginnen, omdat ze niet aan de vereiste norm voldoen. Hierbij dient zich de vraag aan of het kind zich moet aanpassen aan het onderwijssysteem of dat het onderwijssysteem zich moet voegen naar de behoeften van het kind. In het laatste geval moet duidelijk worden waaruit die behoeften bestaan en hoe het onderwijs daaraan tegemoet kan komen.

In het Onderwijsverslag 2005 van de Onderwijsinspectie staat hierover het volgende: "De Wet Primair Onderwijs (1985) beoogt voor leerlingen een onvertraagde ontwikkeling. Leerlingen moeten de school binnen een tijdvak van 8 jaar kunnen doorlopen. Ruim de helft van de basisscholen garandeert de ononderbroken ontwikkeling van hun leerlingen nog niet voldoende. In verreweg de meeste gevallen ligt dit aan het onderwijs in de kleutergroepen en aan de aansluiting op elkaar van de aanpak in de groepen 1 en 2 en de werkwijze in groep 3. Een groot deel van de leraren begint in groep 3 met alle leerlingen op bladzijde 1 van de (lees-)methode en weet geen weg met leerlingen die al verder zijn."

Het vraagstuk over de ononderbroken ontwikkelingslijn leeft ook op basisschool Het Baken in Veenendaal, een grote school met ruim 700 kinderen, die zijn verdeeld over 29 groepen waaronder 12 onderbouwgroepen. Het Baken had in twee opeenvolgen-

de schooljaren formatieruimte om een groep te formeren waarin kinderen met een zogenoemde 'kleuterverlenging' bij elkaar in één klas werden geplaatst. In overleg met de ouders werd besloten deze groep als 2+ en het daaropvolgende jaar als 3+ te benaderen. Niet de methode, maar de ontwikkeling van het kind bepaalt in deze plusgroepen wanneer de leerstof wordt aangeboden. Dit betekent dat de klassikale instructie en de vaste normen waaraan kinderen moeten voldoen, worden losgelaten en dat het onderwijs aansluit bij het niveau van het kind. De kinderen in deze groep leren via betekenisvol spel, vanuit de zone van de naaste ontwikkeling, lezen en rekenen. Elke stap voorwaarts is een overwinning. Er wordt in beide plusgroepen, net zoals in de reguliere groepen 1 en 2, gebruikgemaakt van het programma *Basisontwikkeling*, een ontwikkelingsstimuleringsprogramma voor- en vroegschoolse educatie. Doordat de overige groepen 3 programmagericht werken, deed er zich een situatie voor die zich goed leende voor vergelijkend onderzoek waarvoor de volgende probleemstelling is geformuleerd:

Heeft het brede aanbod, afgestemd op de onderwijsbehoeften van de leerlingen in een speelleergroep een bredere ontwikkelingsperspectief dan het door de methode bepaalde aanbod in de groepen drie?

Acceptatie

In het kader van de opleiding tot Remedial Specialist (VU Amsterdam i.s.m. Fontys Hogeschool), kreeg ik het gereedschap in handen om dit onderzoek op een wetenschappelijk verantwoorde manier op te zetten. Dankzij de begeleiding van docenten van het Onderwijscentrum van de Vrije Universiteit werd het onderzoek theoretisch gevoed vanuit de wetenschap: bestuderen, selecteren en vervolgens beschrijven wat er in de literatuur over de probleemstelling bekend is.

Tijdens de beginfase van het onderzoek bleek dat leerkrachten er moeite mee hadden om zich te focussen op de probleemstelling, die zij nogal vaag en/of te theoretisch vonden. Sommige collega's reageerden defensief: "Doen we het dan niet goed?" Verder wilden sommige leerkrachten graag uitleg over de samenstelling van de groepen en in welke groep de meeste zorgleerlingen zitten. Anderen vonden dat ook de leerkrachtcompetenties meegenomen moeten worden in het onderzoek... Allemaal reële vragen die elk een eigen onderzoek waard zijn.

Al met al was de acceptatie van dit onderzoek bij de start wisselend. De onderbouwcollega's en de directeur waren enthousiast; zij hebben hun volledige medewerking verleend en mij in gesprekken van veel informatie voorzien. Ze waren geïnteresseerd in het verloop van het onderzoek en in de theoretische achtergronden. Ook de ouders reageerden positief. Zij waren met ons een avontuur aangegaan waarin hun kinderen de hoofdrol speelden en zij vonden het plezierig dat de groep zo intensief werd gevolgd. Mede daardoor was de respons op de enquête onder de ouders enorm hoog (89%!).

Er waren ook collega's die afwijzend reageerden op het onderzoek. Een positief gevolg hiervan was, dat de discussie over de kwaliteit van ons onderwijs in een stroomversnelling terecht kwam. In de loop van het onderzoek kwam, in verband met de voortgang van de plusgroep, de vraag 'wat willen wij met onze school?' steeds terug. Hoe gaan we als school verder met de plusgroepen? Kunnen we een aparte stroom binnen de school handhaven? Of is het beter deze groep langzamerhand weer op te nemen in het systeem? Waar blijft dan de winst? Of moet het hele systeem veranderen? Veel vragen die onrust en onzekerheid veroorzaakten.

De discussie die over deze vragen ontstond, maakte duidelijk dat mensen vanuit hun eigen referentiekader verschillende betekenissen toekennen aan begrippen. Het verschil in kennis en inzicht zorgde voor vertroebeling van de *situatie* en op een gegeven moment ook van de *relatie*. Dit bemoeilijkte een open gesprek. Het was dus belangrijk om het hele team van dezelfde informatie te voorzien, zodat zij allemaal dezelfde taal zouden kunnen spreken. Het ging hierbij niet meer alleen om dit onderzoek, maar over 'ons onderwijs'. Het was een meerwaarde van mijn onderzoek, die pas achteraf duidelijk werd en die ik van tevoren niet had voorzien.

Uitkomsten

Bij de analyse van het onderzoek zijn de kinderen vergeleken op het gebied van welbevinden en de gegevens uit het Cito leerlingvolgsysteem. De resultaten van de plusgroepen waren positief: de gemiddelde scores waren op een aantal gebieden (met uitzondering van spelling) hoger dan in de controlegroepen. Zo werd er winst gemaakt op het gebied van rekenen en begrijpend lezen.

De kinderen hebben in de 2+ groep door middel van de tussendoelen geletterdheid leren lezen. De betekenis en het oefenen in die betekenisvolle context zijn hierbij uitgangspunt. Kinderen gaan lezen om de informatie of het verhaal te volgen - het is functioneel. Begrip vormt hier de uitdaging om leeskilometers te maken en zo de techniek van het lezen steeds beter onder de knie te krijgen.

Het onderzoek op Het Baken bevestigt wat de literatuur al langer aangeeft: ons jaar-klassensysteem waar kinderen moeten voldoen aan de gestelde norm moet ombuigen naar een onderwijsvorm waar het kind en zijn ontwikkeling centraal staan.

De opbrengsten van het onderzoek zijn tijdens een bijeenkomst aan de ouders van de plusgroepen gepresenteerd. De ouders waren echter zo geïnteresseerd, dat zij vroegen om een uitgebreide presentatie van het gehele onderzoek, inclusief de theorie. Deze uitgebreide presentatie, waarin de probleemstelling, het theoretisch kader en de resultaten aan bod komen, zal eerst aan het team worden gegeven. De teamleden kunnen op basis van deze informatie zelf conclusies trekken. Het zou natuurlijk mooi zijn als het team tot dezelfde slotsom komt als het onderzoek. Dat zou het begin kunnen zijn van een proces waarin we de visie op onderwijs en leren opnieuw formuleren in onze eigen 'Baketaal'.

Toekomst

Zoals uit het hierboven beschreven voorbeeld blijkt, is het in de praktijk lastig en tijdrovend om in de school onderzoek te doen naar onderwijszaken. Het ontbreekt het onderwijs eenvoudig aan tijd en mankracht. Leerkrachten willen graag weten waarom een onderzoek op een bepaalde manier wordt uitgevoerd, maar zij hebben daarover niet alle kennis in huis en hebben geen tijd om zich uitgebreid in te lezen. Dit probleem kan uitsluitend worden opgelost als leerkrachten en middenmanagement hiervoor worden gefaciliteerd. Wanneer scholen evidence based moeten werken, zal het ministerie van OCW hierin haar verantwoordelijkheid moeten nemen.

Ook is het noodzakelijk dat leerkrachten zelf een onderzoekende houding ontwikkelen, zodat ze zich afvragen: wat werkt? En waaróm werkt het? Deze houding helpt hen om verantwoorde keuzes te maken. Opleidingsinstituten besteden in toenemende mate aandacht aan het ontwikkelen van zo'n onderzoekende attitude en nemen zowel in bachelor- als masteropleidingen een onderzoeksopdracht op. Zo komt het *leren onderzoeken* via de studenten de scholen binnen. Dit biedt kansen om leerkrachten, samen met studenten, te leren onderzoeken. Het management van de school zal de leerkrachten hiertoe moeten stimuleren en hiervoor ruimte moeten bieden in tijd en verantwoordelijkheid: leer onderzoeken zodat je onderzoekend kunt (blijven) leren!

8 Onderwijsbeleid op wetenschappelijke leest geschied

IJsbrand Jepma en Anne Luc van der Vegt | [onderwijsonderzoekers Sardes](#)

In de discussies over evidence based onderwijs gaat het meestal om het effect van bepaalde werkwijzen of materialen. Maar hoe zit het met onderwijsbeleid? In hoeverre is dat evidence based? Kan wetenschappelijk worden aangetoond dat onderwijsbeleid effectief is? In het licht van het rapport van de commissie Dijsselbloem is dit een relevante vraag.

In de discussie over evidence based onderwijs gaat het nagenoeg uitsluitend over de bewezen waarde van werkwijzen en leermaterialen in het onderwijs. Een andere interessante discussie is in hoeverre het onderwijsbeleid 'evidence based' is. Kan worden aangetoond dat beleid effectief is?

Spannende beleidsvragen

Als het gaat om onderwijsbeleid, levert dit spannende beleidsvragen op, zoals:

- Leidt de brede school tot een bredere ontwikkeling van talenten van leerlingen?
- Bevordert de lumpsum financiering in het basisonderwijs dat 100% van de doelgroepkinderen worden bereikt met voor- en vroegschoolse educatie (VVE)?
- Garandeert de invoering van budgetfinanciering bij Passend Onderwijs de kwaliteit van onderwijszorg voor zorgleerlingen?
- Realiseert de 1.000-lesurennorm in het voortgezet onderwijs acceptabele onderwijsprestaties van leerlingen?

Het is de moeite waard om stil te staan bij zulke vragen, temeer daar het onderwijsbeleid van de afgelopen decennia de laatste tijd veel kritiek krijgt te verduren; niet alleen van scholen, maar steeds meer van de overheid zelf. De commissie Dijsselbloem¹ stelt in haar rapport: "De overheid heeft haar kerntaak, het zeker stellen van deugdelijk onderwijs, ernstig verwaarloosd."

Wat is goed onderwijsbeleid? Het antwoord lijkt voor de hand te liggen: goed onderwijsbeleid bevordert betere leerresultaten en minder schooluitval van leerlingen. Dit antwoord past bij de VBTB-filosofie (Van Beleidsbegroting Tot Beleidsverantwoording), die in 1999 is geïntroduceerd. Het succes van beleid meet je af aan de maatschappelijke effecten. We kijken niet alleen naar *output* (bijvoorbeeld het aantal VVE-plaatsen), maar vooral naar *outcome* (minder taalachterstanden bij jonge kinderen).

Toch is deze redenering een beetje te simpel. Dijsselbloem stelt terecht dat de overheid en de scholen elk hun eigen verantwoordelijkheden hebben. Doordat het niet eenvoudig is voor de overheid om *directe* invloed uit te oefenen op de onderwijskwaliteit, is een verband tussen beleid en effecten lastig aan te tonen. Misschien verklaart dat wel het feit dat we de laatste tijd wat minder horen over VBTB.

Van beleidsdoelen tot effecten

In de discussies over evidence based onderwijs gaat het doorgaans om causale relaties tussen een bepaalde onderwijsmethode en leerlingresultaten. Wanneer het gaat om de gevolgen van beleid, spreken bestuurskundigen als Bressers en Klok² liever van doel-middel relaties dan van causale relaties. Waarom? Goed onderwijsbeleid biedt op zich geen garantie voor effecten bij leerlingen. De keten die de verbinding vormt tussen beleidsdoelen en maatschappelijke effecten, heeft een aantal schakels, die alle voldoende stevig moeten zijn om effecten te kunnen bewerkstelligen.

Deze keten begint bij beleid. De overheid (landelijk of lokaal) formuleert bepaalde beleidsdoelen, bijvoorbeeld: 'De taalachterstanden van leerlingen moeten in vier jaar tijd met 25 procent afnemen'. Om dit beleid *uit te voeren*, neemt de overheid bepaalde maatregelen: scholen worden gestimuleerd een eigen taalbeleid te ontwikkelen door middel van een ondersteuningstraject, dat resulteert in een taalbeleidsplan. Dit noemen we de *beleidsprestatie*. Als een beleidsprestatie is geleverd, wil dit nog niet zeggen dat de gestelde doelen zijn bereikt. Daarvoor is het ook nodig dat het beleid een *veldproces* op gang brengt. Scholen besteden bijvoorbeeld meer tijd aan het taalonderwijs, ze gaan de taalprestaties systematischer monitoren en ze organiseren extra hulp voor leerlingen met taalproblemen. Dat is dan de output van het beleid. Het resultaat van deze maatregelen is, hopelijk, dat deze *effect* hebben op de leerlingresultaten. De taalachterstanden nemen af en daarmee is het gestelde beleidsdoel bereikt. Dat is de outcome.

Toegepast op het voorbeeld van taalbeleid ziet de keten 'van beleid tot effect' er als volgt uit:



Dat deze keten werkelijk tot stand komt, is allerm minst vanzelfsprekend. Met name de schakel tussen *beleidsprestatie* en *veldproces* is cruciaal. Vanaf dat moment hebben de beleidsmakers geen directe invloed meer op het behalen van de gestelde doelen. Het veldproces wordt bepaald door allerlei factoren, zoals de visie van de school, het schoolmanagement, de expertise van de leerkrachten, de beschikbare tijd voor verbetering en de invloed van stakeholders, zoals ouders.

Scholen staan voor de keuze hoe ze het overheidsbeleid moeten invoeren. Zij maken een afweging waarin verschillen soorten argumenten een rol spelen:

- (a) *Professionele* argumenten: wat sluit het beste aan bij de professionele opvattingen van het schoolteam over het eigen vak?
- (b) *Historische* argumenten: hebben we als school eerder zoiets gedaan? Wat waren de ervaringen?
- (c) *Pragmatische* argumenten: welke methode is eerder beschikbaar voor de docenten, de leerlingen?
- (d) *Financiële* argumenten: welke methode of werkwijze is het goedkoopst?
- (e) *Wetenschappelijke* argumenten: wat zeggen onderzoek en literatuur over de effecten?

Weging van al die argumenten kan leiden tot veldprocessen die niet helemaal of helemaal niet passen bij de doelen van de overheid. De schakel tussen beleidsprestatie en veldproces is daarom vaak de zwakste schakel in de keten. Wie kent binnen het onderwijs niet het begrip ‘beleidsarm’ innoveren?

Alles meten in één onderzoek?

Vaak wordt van beleidsonderzoek verwacht dat het zicht geeft op de volledige keten, van beleid tot en met effecten. Dat is een hoge verwachting, die hoge eisen stelt aan het onderzoeksdesign. Het lukt onderwijsonderzoekers lang niet altijd om aan die eisen te voldoen. Dat komt deels doordat goede onderzoeksdesigns kostbaar zijn en deels doordat de overheid resultaten wil zien op korte termijn. Zo zijn er voor longitudinaal onderzoek amper middelen en tijd. Er wordt dan soms toch maar gekozen voor een meting achteraf, zonder controlegroep, als het beleid al enkele jaren loopt. Bovendien worden niet altijd gegevens verzameld die zijn toegesneden op de vraag. In plaats daarvan gebruikt men beschikbare gegevensbestanden met onderwijsprestaties, die achteraf in verband worden gebracht met een beleidsmaatregel.

In een recente review van het beleidsgericht onderzoek primair onderwijs (BOPO) werd geconstateerd dat ambitieus beleidsonderzoek vaak geen conclusies mogelijk maakt omtrent effecten³. Zo blijkt uit een studie van het ITS naar de effecten van de brede school⁴, dat het niet mogelijk is om conclusies te trekken over de brede school in het algemeen. Dat is begrijpelijk, omdat de brede school allerlei verschijningsvormen heeft. Ook al hebben de onderzoekers geprobeerd het begrip af te bakenen, dan nog is de diversiteit groot. Wel geeft het onderzoek antwoord op deelvragen. We krijgen zicht op de implementatie van de brede school en het lijkt erop dat brede-schoolactiviteiten tot leerwinst kunnen leiden.

Voor onderzoek naar voor- en vroegschoolse educatie geldt hetzelfde. Naar aanleiding van een lokaal onderzoek kopt de krant “VVE werkt niet”, terwijl het om een specifieke gemeentelijke invulling gaat en er vaak van alles mankeert aan de implementatie. Liever enkele goede landelijke studies naar effectiviteit, dan allerlei lokale studies

onder een handjevol kinderen. Kleinschalige gemeentelijke onderzoeken kunnen zich beter beperken tot implementatie. In combinatie met wat bekend is uit grootschalig effectonderzoek, kunnen dan zinnige beleidsaanbevelingen worden gedaan.

Onderzoeksagenda

Overheidsbeleid is vaak te complex om de volledige keten van beleidsdoel tot effect met één allesomvattende studie te kunnen aantonen. Dat is met name het geval als beleid op veel verschillende manieren kan worden geïmplementeerd, zoals bij de brede school, VVE of Passend Onderwijs. In die gevallen is een deterministische aanpak verstandiger, waarbij deelonderzoeken worden uitgevoerd naar verschillende schakels in de keten. Er worden dan aparte onderzoeken verricht naar de totstandkoming van *beleidsprestaties*, naar *veldprocessen* en naar *effecten*.

Bij het onderzoek naar de eerste schakels in de keten worden niet alleen onderwijskundigen betrokken, maar worden - afhankelijk van de ingezette beleidsinstrumenten - ook andere wetenschappelijke disciplines ingezet, zoals juristen, economen/ econometristen of bedrijfskundigen. Bij dit type onderzoek wordt er doorgaans niet naar gestreefd de werkelijkheid in het keurslijf van (quasi-)experimenteel onderzoek te dwingen. Beleidsvragen over het functioneren van het Nederlandse onderwijssysteem, vereisen veelal een (brede) internationale systeemvergelijking. Bij onderzoek naar de effecten van schaalgrootte in het onderwijs, kan worden aangesloten bij de bestaande 'natuurlijke variatie' in het onderwijsveld. Er zijn immers scholen van verschillende grootte. Voor het kiezen van de juiste onderzoeksmethode is er kortom een zeker mate van onderzoekspragmatisme en creativiteit nodig.

Naast dit beleidsonderzoek, wordt wetenschappelijk onderzoek uitgevoerd naar de effecten van concrete materialen en werkwijzen in het onderwijs. Dit gebeurt waar mogelijk op een experimentele manier, maar er zijn ook alternatieven. Andere bijdragen in deze special houden zich bezig met de vraag wat de beste onderzoeksmethoden zijn. De combinatie van beleidsgericht onderzoek en wetenschappelijk onderzoek geeft zicht op de volledige keten van beleid tot effecten. Wetenschappelijk onderzoek naar werkwijzen en materialen laat zien wat effectief is, het beleidsonderzoek laat zien in hoeverre het beleid is geïmplementeerd.

Verschiede soorten onderzoeken kunnen op elkaar worden afgestemd door middel van een onderzoeksagenda voor de evaluatie van (voorgenomen) onderwijsbeleid. Deze onderzoeksagenda wordt interactief samengesteld. Zowel politici, beleidsmakers, wetenschappers als praktijkdeskundigen (afvaardiging van schooldirecteuren én leerkrachten) worden hierbij betrokken. Een goede samenwerking tussen onderzoekers en beleidsmakers kan ook voorkomen dat het onderwijsveld wordt opgezaald met torenhoge verwachtingen en ambities van de politiek. Goed onderzoek kan

namelijk ook een corrigerende werking hebben. Wanneer doelstellingen niet worden gehaald, kan het zijn dat de scholen niet doen wat ze zouden moeten doen, maar het kan ook een gevolg zijn van het feit dat de overheid geen goede doelstellingen formuleert.

Meer ruimte voor onderzoek, goed doordacht beleid op de langere termijn (duurzaam onderwijsbeleid) en minder ruimte voor (wisselende) ideologie; dat past ook heel goed bij de aanbevelingen van de commissie Dijsselbloem.

- 1 Commissie Parlementair Onderzoek Onderwijsvernieuwingen (2008). Tijd voor onderwijs: Eindrapport. Den Haag: Tweede Kamer
- 2 Bressers, J.Th.A. & Klok, P.J. (2003). De inhoud van het beleid. In: A. Hoogerwerf & M. Herweijer (red.). Overheidsbeleid, een inleiding in de beleidswetenschap. Alphen a/d Rijn: Kluwer
- 3 Doolaard, S. & Leseman, P.P.J. (2008). Versterking van het fundament: integrerende studie naar aanleiding van de opbrengsten van de onderzoekslijn 'Sociale en institutionele context van scholen' uit het onderzoeksprogramma BOPO 2005-2008. Groningen: GION
- 4 Claassen, A., Knipping, C., Koopmans, A. & Vierke, H. (2008) Variatie in brede scholen en hun effecten. Nijmegen: ITS

9 Vier vragen over evidence based werken in het onderwijs

Jo Kloprogge | adviseur Sardes

Waar komt het begrip ‘evidence based’ vandaan? En hoe komt het dat dit begrip in het Nederlandse onderwijs zo snel ingang heeft gevonden? Wanneer is een aanpak wel of niet evidence based? En is het eigenlijk wel waar dat wat evidence based wordt genoemd ook werkelijk evidence based is? Jo Kloprogge gaat in op deze vier vragen.

Het begrip ‘evidence based’ is in vrij korte tijd in Nederland bekend en gevreesd geworden. Evidence based behoort te worden uitgesproken met een licht Amerikaans accent en met een ondertoon van hetzij respectvolle waardering, hetzij intuïtieve afschuw. Het moet eigenlijk meteen aan de uitspraak te herkennen zijn of je vóór of tegen evidence based bent en of je deel uitmaakt van de hardliners of gematigden aan de pluskant, dan wel van de verwerpers of licht verontrusten aan de minkant. Het is verstandig om zorgvuldig positie te kiezen, want de roep om evidence based zal in de onderwijswereld ongetwijfeld nog een aantal jaren klinken.

Waar komt evidence based vandaan?

Over de herkomst van evidence based wordt vaak gezegd dat het uit de Verenigde Staten afkomstig is. Dat zal wel zo zijn, maar het zegt eigenlijk nauwelijks iets. Alle mogelijke trends en hypes zijn immers wel te vinden in de VS, van uiterst softe tot uiterst harde benaderingen. Meer plausibel is dat evidence based is overgenomen uit de medische wereld. Wikipedia verwijst hier ook nadrukkelijk naar: “Evidence based practice (EBP) is het uitvoeren van een handeling door een beroepsbeoefenaar op zo’n wijze dat de uitvoering is gebaseerd op de best beschikbare informatie over doelmatigheid en doeltreffendheid. Binnen de geneeskundige en paramedische beroepen is EBP een belangrijke stroming, die vaak wordt aangeduid als *evidence based medicine* (EBM). Het evidence based practicemodel vindt als beoordelingsinstrument steeds meer ingang bij overheden en verzekeraars die het beroepsmatig handelen economisch willen toetsen.”

Waarom vindt evidence based zo snel ingang in het Nederlandse onderwijs?

Dat evidence based zo snel ingang heeft gevonden in het Nederlandse onderwijs heeft waarschijnlijk te maken met de kater na het ‘Nieuwe Leren’ en de kritische opmerkingen van de commissie Dijsselbloem over het geringe evidence based gehalte van onderwijskundige vernieuwingen. Het Nieuwe Leren breidde zich immers enkele jaren geleden als een olievlek uit over scholen, zonder dat er een wetenschappelijke onderbouwing aan ten grondslag lag, en zonder dat het werd ingebed in een experimentele situatie, waar het in een beperkte en gecontroleerde omgeving zijn waarde had kunnen bewijzen.

Dat ging dus niet goed. Overigens is de blaam hiervoor wel erg eenzijdig bij de scholen gelegd. Als de filmpjes die werden getoond bij het afscheid van KPC directeur Carel van den Heuvel in april 2005 nog eens worden vertoond, zal menig onderwijscoryfee met enige gêne haar of zijn standpuntbepaling over het Nieuwe Leren terugzien.

Wanneer is iets wel of niet evidence based?

Deze vraag zal ons nog wel jarenlang bezighouden. Zowel in wetenschappelijke kring als in het werkveld lopen de meningen hierover sterk uiteen. In wetenschappelijke kring gaat het vooral om de vraag of je iets evidence based mag noemen zonder dat deze conclusie is gebaseerd op een onderzoek met een experimenteel design. Kan evidentie die tot stand is gekomen op basis van een minder rigoureuus design doorslaggevend zijn om iets als werkzaam te verklaren of niet? Dit is een heel akelige vraagstelling om de eenvoudige reden dat zuiver experimentele designs heel moeilijk toepasbaar zijn in het onderwijs. Strak vasthouden aan dit beginsel zou er wel eens toe kunnen leiden dat de nieuwe editie van het boekje *What works!* uit niet meer dan een inleiding en een slotbeschouwing bestaat.

Maar hanteren we minder strikte methodologische regels, dan rijst de vraag waar we dan precies de grens leggen tussen methodologieën die we overtuigend vinden en methodologieën die tekort schieten. Als we niet uitkijken is onderwijskundige praktijk evidence based als we een of twee wetenschappers vinden die zeggen dat je die praktijk best als evidence based kunt beschouwen, omdat er toch wel een aardig verhaal aan ten grondslag ligt of omdat zij het wel een goed idee vinden. Dan kunnen we alles wel weer evidence based noemen, want in de wetenschap zijn er altijd wel een paar experts te vinden die bereid zijn een keurmerk af te geven, al dan niet gestimuleerd door het vooruitzicht een kleine evaluatie aan het nieuwe project op te mogen hangen.

De kunst zal dus zijn om een balans te vinden tussen overtuigende, maar in het onderwijs nauwelijks uitvoerbare, strikte methodologische designs en makkelijk uitvoerbare keurmerkprocedures zonder echte overtuigingskracht. Misschien moet er eens een stevige conferentie worden belegd om die balans te zoeken.

Is evidence based wel evidence based?

Je zou verwachten dat de eis om in het onderwijs evidence based te werken, stevig wordt geschraagd door bewijs dat de werkwijzen die evidence based worden geacht, ook inderdaad in grote lijnen de beoogde effecten halen. In de medische wetenschappen is dat in ieder geval in grote lijn wel het geval. Hoewel een niet te onderschatten aantal burgers zijn heil zoekt in twijfelachtige geneeskundige praktijken, zouden we toch niet graag overgeleverd zijn aan artsen en ziekenhuizen die de

resultaten van wetenschappelijk onderzoek naast zich neerleggen.

Maar in het onderwijs treffen we veel minder bewijs aan dat evidence based werkt. De redenen daarvoor hoeven we niet ver te zoeken. Allereerst is het volume aan onderwijskundig onderzoek dat jaarlijks wordt uitgevoerd, een haast verwaarloosbare fractie van het volume aan medisch onderzoek. De fondsen voor onderwijsonderzoek zijn na het verscheiden van de Stichting voor Onderzoek van het Onderwijs miniem en ook nog eens opgedeeld in kleine snippertjes. De eens toch wel bloeiende onderwijswetenschappen zitten al jaren in het verdomhoekje en kunnen nauwelijks substantieel informatie over evidence based opleveren.

Maar bij onderwijs spelen ook culturele factoren een rol. Als in de VS uit onderzoek blijkt dat een sterk leiderschap op school helpt om goede leerlingprestaties te realiseren, is het maar de vraag of dit resultaat ook zonder meer op Nederland toepasbaar is. Tegen leiderschap wordt hier toch wel anders aangekeken dan in veel andere landen. Hiermee is niet gezegd dat onderzoeksresultaten per definitie niet overdraagbaar zijn, maar een grote portie voorzichtigheid is wel op zijn plaats.

We hebben in Nederland zelfs een interessant voorbeeld van een onderwijspraktijk die bij introductie niet alleen niet evidence based was, maar zelfs strijdig met veel van wat we aan onderwijskundige kennis in de ransel dachten te hebben. Het betreft hier de schakelklassen waarbij kinderen met taalachterstanden een jaar lang apart worden geplaatst om vooral Nederlands te kunnen leren. De toets van evidence based, hetzij op basis van onderzoek hetzij op basis van een theoretische toetsing, hadden de schakelklassen nooit kunnen doorstaan. En toch zijn de stevige evaluaties die zijn uitgevoerd erg positief. Tegen alle wijsheid in pakken de schakelklassen goed uit.

Dit alles is geen pleidooi om het principe van evidence based los te laten. Het is belangrijk om onderwijsonderzoek, onderwijspraktijk en onderwijsbeleid aan elkaar te verbinden en onderzoeksresultaten zeer serieus te nemen. Maar dit is wel een pleidooi om het principe van evidence based alleen daar te hanteren waar dat kan, en ook dan goed af te wegen wat je wel en niet uit het wetenschappelijk onderzoek kunt afleiden.

De toekomst

We zullen in de komende vijf tot tien jaar nog wel ernstig met het begrip evidence based worstelen. Bestuurders op landelijk en lokaal niveau beginnen nu al te vragen 'hoe evidence based' vernieuwingsvoorstellen van scholen of andere educatieve instellingen zijn. Alvorens subsidies worden toegekend, wordt steeds vaker de vraag gesteld of de bestemming van het geld de toets van evidence based kan doorstaan. Bestuurders willen immers graag het risico vermijden dat ze kostbare belastingcenten

hebben uitgedeeld aan een onderwijspraktijk, zonder dat zij zich ervan hebben vergewist of deze wel onderbouwd kan worden door onderzoek. Aangezien de onderbouwing door onderzoek bijna altijd afwezig of zeer gebrekkig is, zal snel worden uitgeweken naar het principe van de expertbeoordeling. Wie in het bezit is van een titel, 'hoogleraar' of 'doctor' voor zijn naam kan zetten of een reputatie heeft van 'niet al te dom in onderwijsland', mag zich voorbereiden op allerlei vragen om een bepaalde aanpak al dan niet het stempel van evidence based te geven. Men zal graag bereid zijn een zakje zilverlingen over tafel te gooien om een wetenschappelijke vrijbrief te krijgen om een project te kunnen subsidiëren dat bijvoorbeeld ouders bij de school betreft, segregatie tegengaat of het rekenonderwijs verbetert.

Na een paar jaar is deze trend uitgewoed en maakt deze plaats voor een nieuwe bestuurlijke wending. Dan zal echt blijken of we verder willen met het wetenschappelijk onderbouwen van ons praktisch en beleidsmatig onderwijskundig handelen. Daarvoor is naar mijn idee nodig dat:

- het volume aan onderzoek wordt verveelvuldigd
- het onderzoek drastisch wordt gereorganiseerd, en
- de samenwerking tussen onderzoek, praktijk en beleid in het onderwijs in de vorm van kenniskringen, onderwijsbewijsachtige procedures en andere verbanden wordt verbreed en geïntensiveerd.

Langs die weg kunnen we over een aantal jaren misschien echt wetenschappelijk onderbouwd onderwijswerk doen, ongeacht of dat dan nog evidence based heet of anders wordt genoemd. Voorshands moeten we nog vooral aanmodderen.

me un porc, un
es. Quelquefois

je les renvoye

me. Cher un d

a à bouff:

Des chaloupe

. Pour bouff

er au sein cor

