

Troeven en limieten van evidence-based aanpak in het onderwijs

Auteur: Olivier Mottint

Dit artikel verscheen in “*De democratische school*”, nr. 91, september 2022

Deel 1

De evolutie van "evidence-based policy", van geneeskunde tot onderwijs: grondslagen, eerste resultaten, eerste spanningsvelden

In de Angelsaksische landen is “*evidence-based policy*”, empirisch onderbouwd beleid, al goed ingeburgerd. Andere landen volgen, vooral in Europa, elk in eigen tempo. Aanvankelijk beperkt tot het domein van de geneeskunde, wordt deze nieuwe beleidsopvatting nu toegepast op bredere terreinen van overheidsoptreden, met bijzondere aandacht voor het onderwijs. Volgens de voorstanders van deze aanpak is het de bedoeling om af te stappen van ideologie en eindelijk het pragmatisme te laten zegevieren door hervormingen te baseren op de resultaten van wetenschappelijk onderzoek, het instrument bij uitstek om beleidsmakers duidelijk te maken "wat werkt". Hebben wij eindelijk het geheim van een doeltreffend overheidsbeleid ontdekt, of is dit slechts een technocratische fata morgana? Het lijkt nuttig de balans op te maken van deze aanpak in een tijd waarin "*evidence*" (bewijsmateriaal) steeds vaker wordt ingeroepen om onderwijsbeleid te legitimeren.

Evidence-based medicine op zoek naar de gouden standaard...

Zoals opgemerkt door Draelants en Revaz (2022, p. 57) "*is het idee om beleid en praktijk te baseren op gedegen kennis net zo oud als de Verlichting*" en "*heeft men niet gewacht op de opkomst van evidence-based medicine*". Pas sinds de jaren 1990 nam de tendens stelselmatig toe, om beleidshervormingen te baseren op "*wetenschappelijke bewijzen*". In het Verenigd Koninkrijk en de Verenigde Staten wensten de beleidsmakers voortaan de gegevens van evidence-based medicine, die in het vorige decennium waren ontstaan, om te zetten in gezondheidsbeleid en -praktijken (Baron, 2018; Draelants & Revaz, 2022).

Evidence-based medicine (EBM) ontstond uit de wens om de medische praktijk te baseren op de resultaten van kwantitatief onderzoek. Het beoogde doel is zorgverleners te helpen de "*beste beslissingen*" voor patiënten te nemen door hen de "*best mogelijke wetenschappelijke informatie*" te verschaffen, door de doeltreffendheid van verschillende concurrerende behandelingsmethoden voor een bepaalde aandoening te vergelijken. Om de efficiëntie van gezondheidsbeleid en -praktijken te meten, steunt de *evidence-based medicine* op een hiërarchie van bewijsmateriaal (Draelants & Revaz, 2022; Woolf, 1989). Onderaan de piramide staan de intuïtieve, situationele meningen en oordelen van gezondheidswerkers, die in deze nieuwe benadering als onbetrouwbaar worden beschouwd. Bovenaan staan kwalitatieve studies (beschrijvende studies, case studies, enz.), gevolgd door "onvolmaakte" (niet-gerandomiseerde, enz.) statistische studies. Gerandomiseerde gecontroleerde onderzoeken (RCT's) [1] zijn de meest rigoureuze statistische studies en worden beschouwd als het meest betrouwbare bewijsmateriaal (de "*gouden standaard*"). De replicatie van zulke studies over eenzelfde interventie maakt het eventueel mogelijk meta-analyses uit te voeren: deze geven een overzicht van de stand van het onderzoek over dit specifieke onderwerp, waardoor de effectieve intensiteit

van een interventie nauwkeuriger kan worden gekwantificeerd. De statistische grootte die de "kracht" van een behandeling of een interventie kenmerkt, is de "effectgrootte" ("effect-size").

Voorstanders uit diverse hoeken ...

Onder impuls van het door neoliberalen bepleite *New Public Management* grepen politieke besluitvormers snel naar de *evidence-based medicine* als een middel om een efficiënt gezondheidsbeleid te voeren, de overheidsuitgaven tot een minimum te beperken en tegelijkertijd de resultaten ("*outcomes*") te optimaliseren. Volgens hen was het een instrument om het overheidsoptreden te rationaliseren, en zo komaf te maken met de aan de staat toegeschreven inefficiëntie en "bureaucratische rompslomp" (Draelants & Revaz, 2022).

Andere politieke actoren, wat verder naar links in het politieke spectrum, schaarden zich ook achter deze nieuwe aanpak. Zij hoopten daarmee het imago van de staat door een meer "efficiënte" aanpak op te krikken, te beschermen tegen de meest virulente aanvallen en zodoende het vertrouwen van de burger-gebruikers in de overheid en de politieke leiders te herstellen (Chatterjee, 2022; Draelants & Revaz, 2022; Ozga, 2009).

Verknochte voorstanders van wetenschappelijke rationaliteit vormen een derde categorie. Verspreid over het gehele politieke spectrum, tot aan de linkerzijde van links [2], zijn zij de mening toegedaan dat een *evidence-based* beleid, door het terzijde schuiven van professionele uitspraken en praktijken gebaseerd op "mythen", of overtuigingen zonder afdoende wetenschappelijke onderbouwing, zal bijdragen tot kwalitatief betere openbare diensten voor alle burgers, ook de meest kwetsbaren.

Van geneeskunde tot onderwijs...

In navolging van de medische sector, zou *evidence-based* beleid ook de weg vinden naar de onderwijssector, vanaf het begin van de jaren 2000 in pionierslanden zoals het Verenigd Koninkrijk (Draelants & Revaz, 2022; Wiseman, 2010). Weliswaar met enige verschillen in het ontwerp- en uitvoeringsproces: zo is het om praktische of ethische redenen [3] niet altijd mogelijk om in het onderwijs gerandomiseerde gecontroleerde proeven (RCT's) uit te voeren.

Wanneer dergelijke RCT's niet beschikbaar of haalbaar zijn, refereren beleidsmakers vaak naar grote internationale onderzoeken (PISA, PIRLS, TIMSS...). Deze grote studies vergelijken onderwijsstelsels met elkaar en worden daarom geacht de beste presterende ("*benchmarks*" en andere "*best practices*") te identificeren. Sommige economen spelen een belangrijke rol in de ontwikkeling van die *evidence-based education* (Draelants & Revaz, 2022): voorstanders van de "menselijk kapitaal theorie" die onderwijs zien als een "investering" om individuen en staten klaar te stomen om te "*presteren*" in de "*kenniseconomie*". Zij moedigen beleidsmakers aan om "*efficiënte*" hervormingen door te voeren, die zowel de door de arbeidsmarkt vereiste competenties effectief ontwikkelen, als de overheidsuitgaven tot een minimum beperken. Vanuit dit oogpunt wordt het politieke gebruik van wetenschappelijke onderzoeksgegevens beschouwd als de beste weg naar een doeltreffend onderwijsbeleid en onderwijspraktijken met bewezen "*rendabiliteit*".

Reeds in het begin van de jaren 2000 werd *evidence-based education* door middel van grote hervormingen in de Verenigde Staten ingevoerd. Het programma "*No Child Left Behind*" (NCLB) bijvoorbeeld bestaat uit een hybride hervorming die verantwoordingsmechanismen (invoering van gekwantificeerde doelstellingen, door scholen te bereiken op straffe van sancties) combineert met stimulansen om kwantitatief onderzoek te ontwikkelen en de

aanbevelingen daarvan op scholen toe te passen. Het programma “*Success for All*” (SFA) wordt gekenmerkt door de invoering van specifieke onderwijsscripts waarvan de prestaties zijn aangetoond door “*rigoureuze statistische studies*” en “*replica*” (herhaalde experimenten) (Slavin e.a., 1996). In Frankrijk richtte onderwijsminister Blanquer in 2018 een *Conseil Scientifique de l'Education Nationale* (CSEN) op, belast met de ontwikkeling van een “*evidence-based pedagogie*” (Rosier, 2018), hoewel deze zich niet strikt beperkt tot kwantitatief onderzoek.

In Vlaanderen bepleit het recente rapport van de Commissie *Beter Onderwijs* (Brinckman & Versluys, 2021) onder meer de invoering van “directe instructie”, een pedagogische methode die volgens internationaal kwantitatief onderzoek effectief is, in tegenstelling tot methoden die gericht zijn op het ontdekken van kennis door leerlingen. In het Franstalig onderwijs (van België) kan de hervorming van 1994 die leerlingen automatisch van het eerste naar het tweede jaar van het secundair onderwijs laat overgaan, worden beschouwd als een van de eerste “*evidence-based*” onderwijsinitiatieven. Deze (gedeeltelijke) afschaffing van het zittenblijven werd destijds namelijk doorgevoerd op basis van onderzoeksresultaten [4] (Draelants & Revaz, 2022). Het *Pacte pour un Enseignement d'Excellence* bevat ook expliciete verwijzingen naar het *evidence-based* perspectief. Onderzoekers van de universiteit van Mons en van het (Franstalig) katholiek onderwijs bieden opleidingen aan rond expliciet onderwijs “*in overeenstemming met de evidence-based aanpak*” [5]. Een team van de Universiteit van Luik is bezig met het opzetten van *Positive Behavior Support* (PBS), een programma afgeleid van *evidence-based education*, en dat reeds in veel landen werd ingevoerd (Verenigde Staten, Verenigd Koninkrijk, Australië, Nieuw-Zeeland, Spanje, Denemarken, enz.). Zoals de naam aangeeft, beoogt dit programma “positief gedrag” bij leerlingen te stimuleren door te zorgen voor de systematisering van de meest “*doeltreffende*” ingrepen op dit gebied.

Gemengde eerste resultaten: wanneer ‘*what works*’ niet werkt...

De feitelijke implementatie van *evidence-based medicine* was lang niet altijd succesvol. De ontwikkeling van systemen gebaseerd op rigoureuze kwantitatieve studies, liepen in de praktijk soms uit op tegenvallende resultaten. Zoals opgemerkt door Draelants en Revaz (2022, p. 36) “*is het pikant om vast te stellen dat dit soort beleid, dat zich beroept op de retoriek van ‘wat werkt’, niet noodzakelijk werkt en vaak op uitvoeringsproblemen stuit*”.

In het onderwijsveld is de stand van zaken niet anders: Bressoux (2017, p. 125) schrijft dat “*de twijfels (...) over een top-down model van onderzoek naar het veld niet ongegrond zijn. Zo verwijst Marc Gurgand naar een opmerkelijke observatie van de econoom Brian Jacob. In diens verslag over de veralgemening van 77 toepassingen die sinds 2002 werden uitgevoerd door het Institute of Education Sciences (VS) volgens wetenschappelijke criteria (randomisatie, enz.) en als doeltreffend werden beoordeeld, constateert Jacob dat, na uitvoering en evaluatie in hun eigen schoolcontext, slechts 7 ervan positieve effecten hadden.*”

Verskillende hypothesen (Biesta, 2007, 2010; Draelants & Revaz, 2022) kunnen deze kloof tussen de beloften van empirisch onderbouwd beleid en de daadwerkelijke resultaten ervan verklaren. Ze worden goed samengevat door Bressoux (2017, p. 125): “*een gebrekkige uitvoering (slecht begrepen programma's, terughoudendheid om ze uit te voeren, onvoldoende ondersteuning, enz.), of het feit dat veel programma's, ondanks hun correcte uitvoering, in feite slecht waren aangepast aan sommige van de omstandigheden die in de generalisatiefase werden aangetroffen*”. Draelants en Revaz (2022, p. 31) concluderen dat “*het bewijs van de superioriteit van methoden van aanpak van het evidence-based type nog moet worden geleverd*”.

Uiteraard mag deze vaststelling niet leiden tot een diskwalificatie van het kwantitatief wetenschappelijk onderzoek in de geneeskunde. Maar wel tot een nederige revisie aangaande evidence-based medicine. Hoewel statistisch redeneren een krachtig instrument is om "algemene wetten" vast te leggen en "hoogstwaarschijnlijk doeltreffende" maatregelen te identificeren, leidt het onvermijdelijk tot blinde vlekken, die voortkomen uit specifieke, complexe en oorzakelijke verbanden (Draelants & Revaz, 2022). Als een statistische studie met grote betrouwbaarheid kan vaststellen dat een interventie I een gemiddeld effect E teweegbrengt op een populatie P, kan ze die conclusie alleen bereiken door individuen, bijzondere situaties en contexten samen te voegen en zo de complexiteit te verminderen, door specifieke kenmerken "vrijwillig te negeren". (Draelants & Revaz, 2022; Weisberg, 2014). Zij kan dus niet verklaren waarom dezelfde interventie I contrasterende of zelfs tegengestelde effecten (E1 en E2) heeft op de individuen P1 en P2. Zij is ook blind voor de kwalitatieve aard van de causale ketens die verklaren waarom interventie I dit effect E teweegbrengt. Zo hebben gerandomiseerde gecontroleerde proeven bijvoorbeeld aangetoond dat lithiumzouten bepaalde psychiatrische stoornissen kunnen genezen, maar geen verklaring bieden voor de achterliggende biochemische mechanismen. Toch zouden wellicht, dank zij het inzicht in deze mechanismen, behandelingen kunnen worden geïndividualiseerd, geïmplementeerd, gereguleerd en geoptimaliseerd in specifieke situaties en/of voor bepaalde patiënten.

Naar een rehabilitatie van beroepservaring en kwalitatieve benaderingen?

Ten gevolge daarvan wordt *evidence-based medicine* in zekere zin "herbekeken", zoals Draelants & Revaz (2022) duidelijk maken door te spreken over *evidence-based medicine 2.0*. De overdonderende beweringen over de bijna exclusieve legitimiteit van RCT's maken nu plaats voor de rehabilitatie van het oordeel van gezondheidswerkers, als waardevolle aanvulling om de bijzonderheden van individuen en situaties te begrijpen en ernaar te handelen. Rekening houden met de mening van beroepsbeoefenaars biedt ook het voordeel dat hervormingen die gebaseerd zijn op kwantitatieve gegevens, beter door hen worden onthaald en dus beter zullen worden uitgevoerd. Dit is niet verwonderlijk wanneer men weet dat "*uit verschillende empirische onderzoeken blijkt dat juist wanneer de actoren op het terrein als passieve uitvoerders van hervormingen worden beschouwd, deze op een mislukking uitdraaien*" (Draelants & Revaz, 2022, p.53).

EBM 2.0. heeft ook kwalitatief onderzoek geherwaardeerd, dat de statistische informatie nuttig aanvult. Dit kan bijvoorbeeld bijdragen tot een beter begrip van de complexiteit van de betrokken processen, tot een beter begrip van de aard van de causale ketens die zich ontwikkelen tussen een interventie en de waargenomen effecten ervan, en tot een beter begrip van de contrasterende doeltreffendheid van interventies in verschillende situaties. Op het gebied van *evidence-based education* beginnen dergelijke tendensen van beredeneerde rehabilitatie van de praktijkervaring (van leerkrachten, leidinggevenden, enz.) schuchter de kop op te steken, ook al kijken sommige onderzoekers die voorstanders zijn van EBE, in hun geschriften jammer genoeg nog minachtend neer op elke opmerking of oordeel uit het veld, en wordt die snel terzijde geschoven met het argument dat "*ze niet wetenschappelijk gefundeerd zijn*". Ook de herwaardering van gemengde onderzoeksmethoden, waarbij kwantitatieve en kwalitatieve studies worden gecombineerd komt vaker voor, en de effecten daarvan zijn positief. Zo had de uitvoering van het programma "Language skills" in Franse kleuterscholen – dat de taalvaardigheid van leerlingen uit de arbeidersklasse beoogt te ontwikkelen via voorlezen door de ouders – veel baat bij de combinatie van een kwantitatieve én een kwalitatieve benadering. Dank zij de kwalitatieve studie werd duidelijk waarom het programma

bij sommige leerlingen minder goed aansloeg dan bij andere, en kon het instrument bijgevolg worden aangepast (Draelants & Revaz, 2022; Pin & Barone, 2021).

Laten we hopen dat deze eerste keerpunten in de evolutie van *evidence-based education* verdere ontwikkelingen kennen. Hoewel statistische studies geenszins een diskwalificatie verdienen, ziet het ernaar uit dat men toch meer rekening moet houden met het oordeel van professionals en met kwalitatieve studies om te komen tot een toepassing van onderzoeksresultaten die beter afgestemd is op de complexiteit en diversiteit van situaties "op het terrein".

Olivier Mottint

Referenties

- Baron, J. (2018). A Brief History of Evidence-based Policy. *The Annals of the American Academy of Political and Social Science*, 678 (1), 40-50.
- Biesta, G. J. J. (2007). Why 'What Works' Won't Work: Evidence-based Practice and the Democratic Deficit of Educational Research. *Educational Theory*, 57 (1), 1–22.
- Biesta, G. J. J. (2010). Why 'What Works' Still Won't Work: From Evidence-based Education to Value-Based Education. *Studies in Philosophy & Education*, 29 (5), 491-503.
- Bressoux, P. (2017). Practice-Based Research : une aporie et des espoirs. Een kritische blik op het artikel van Anthony S. Bryk. *Education et didactique*, 11 (3), 123-134.
- Brinckman, P. & Versluys, K. (Eds.) (2021). *Naar de kern: de leerlingen en hun leer-kracht*. Rapport van de Commissie Beter Onderwijs, Oktober 2021.
- Bryk, A. S. (2015). Accelerating How We Learn to Improve. *Educational Researcher*, 44 (9), 467-477.
- Chatterjee, E. (2022). Numbers Without Experts: The Populist Politics of Quantification. In C. Newfield, A. Alexandrova & S. John (Eds.), *Limits of the Numerical : The Abuses and Uses of Quantification* (pp. 21-46). Chicago : University of Chicago Press.
- Draelants, H. & Revaz, S. (2022). *L'évidence des faits : La politique des preuves en éducation*. Paris : PUF.
- Ozga, J. (2009). Governing Education Through Data in England: From Regulation to Self-Evaluation. *Journal of Education Policy*, 24 (2), 149-162.
- Pin, C. & Barone, C. (2021). L'apport des méthodes mixtes à l'évaluation : Une expérimentation randomisée dans le champ des politiques éducatives. *Revue française de science politique*, 71 (3), 391-412.
- Rosier, L. (2018). *Un conseil scientifique pluridisciplinaire pour l'école*. On-line beschikbaar op de website van het Franse dagblad *Le Monde* : https://www.lemonde.fr/education/article/2018/01/09/un-conseil-scientifique-pluridisciplinaire-pour-l-ecole_5239234_1473685.html
- Weisberg, H. I. (2014). *Willfull Ignorance : The Mismeasure of Uncertainty*. Hoboken, NJ : Wiley.
- Wiseman, A. W. (2010). The Uses of Evidence for Educational Policymaking: Global Contexts and International Trends. *Review of Research in Education*, 34 (1), 1–24.
- Woolf, S. H. (Ed.) (1989). *Guide to Clinical Preventive Services: Report of the U.S. Preventive Services Task Force*. Darby, PA : DIANE Publishing.

Voetnoten

[1] RCT's (gerandomiseerde gecontroleerde experimenten) zijn een soort kwantitatieve wetenschappelijke studie waarbij de deelnemers willekeurig worden ingedeeld in groepen die overeenkomen met elke behandeling (medisch, educatief, enz.) die wordt getest. Deze "randomisatie" zorgt ervoor dat de proefpersonen gelijkmatig worden verdeeld voordat zij de interventie "ondergaan". De doeltreffendheid van de interventie wordt gemeten door vergelijking met een "controlegroep" die een placebo krijgt of een behandeling waarvan de doeltreffendheid reeds bekend is.

[2] Normand Baillargeon bv, een onderwijsfilosoof met nauwe banden met het anarchosyndicalisme en een vurig verdediger van een cultuur van rationaliteit, heeft herhaaldelijk gepleit voor het principe van evidence-based onderwijs.

[3] Zowel in de geneeskunde als in het onderwijs is het bijvoorbeeld niet aanvaardbaar RCT's uit te voeren wanneer er alle reden is om aan te nemen dat de verwachte voordelen van de interventie vrijwel zeker en groot zijn. Men kan er dan niet opzettelijk voor kiezen de personen in de controlegroep deze zeer waarschijnlijke voordelen te onthouden

[4] Onder de onderzoeken die destijds werden aangehaald, vermelden we vooral dit van Marcel Crahay, die de ondoeltreffendheid van zittenblijven aantoonde. Dergelijke onderzoeken vielen in goede aarde bij beleidsmakers die inzagen welk financieel voordeel een vermindering van het zittenblijven kon opleveren (Draelants & Revaz, 2022).

[5] <http://lenseignement.catholique.be/cecafoc/formation.html?mid=7033>

Deel 2

Bewijzen PISA, TIMSS, PIRLS ... onmiskenbaar de meerwaarde van evidence-based aanpak in onderwijs?

Evidence-based onderwijs steunt op twee pijlers: de grote internationale enquêtes die de prestaties van leerlingen evalueren en quasi-experimenteel onderzoek. In dit artikel behandelen we de internationale studies. Zij zijn bij het grote publiek het meest bekend. De publicatie van de PISA-ranglijst bijvoorbeeld zorgt om de drie jaar voor opschudding in de media. Overall in de geïndustrialiseerde wereld verheugt men zich of maakt men zich zorgen over de paar plaatsen die het nationale (bij ons het Vlaamse en het Franstalige) onderwijssysteem heeft gewonnen of verloren. Deze focus op de landenrangschikking is betwistbaar. De verschillen tussen landen zijn relatief klein, soms volstaan een paar punten (op een schaal van 0 tot 1000) om tientallen plaatsen in de ranglijst op te schuiven. Op deze basis is "voortgang" of "achteruitgang" nogal relatief. Temeer omdat deze schommelingen soms kunnen verklaard worden door methodologische vertekeningen (Stewart, 2013). De obsessie voor de Pisa-ranking is des te meer misplaatst omdat de kloof tussen de prestaties van "rijke" en "arme" leerlingen binnen hetzelfde onderwijssysteem die veel groter is dan de verschillen in gemiddelde scores tussen landen, veel minder aandacht krijgt. [1].

Wat ook vaak aan de media-aandacht ontsnapt, zijn de analyses en beleidsaanbevelingen die door internationale instellingen worden ontwikkeld naar aanleiding van de grote internationale studies (PISA, PIRLS, TIMSS). Deze analyses en aanbevelingen beweren "evidence-based" te zijn - d.w.z. steunend op de statistische verwerking van de in deze enquêtes verzamelde gegevens - en beogen de politieke leiders richtsnoeren te geven voor "goed bestuur" in het onderwijs. Zijn de analyses en aanbevelingen van grote instellingen zoals de OESO daardoor automatisch betrouwbaar? En hoe betrouwbaar zijn dan onze analyses van Ovds? Wij gebruiken immers PISA-gegevens om aan te tonen dat de schoolmarkt de belangrijkste verklarende factor is voor de ongelijkheid tussen scholen. Kan een evidence-based aanpak, gebaseerd op grote internationale onderzoeken, een waardevol instrument zijn om in alle landen een ambitieus en rechtvaardig onderwijsbeleid uit te werken? In dit artikel bespreken we de limieten en enkele valkuilen bij het gebruik van de grote internationale enquêtes voor het uitstippelen van onderwijsbeleid.

De vooronderstellingen van het paradigma van de "efficiënte school"

Een eerste punt van waakzaamheid betreft de invloeden op en vanuit de grote internationale onderzoeken. Deze onderzoeken, hoofdzakelijk uitgevoerd door twee internationale instellingen op het raakvlak van de economische, politieke en academische wereld, zijn niet vrij van theoretische, methodologische en ideologische keuzes die in vraag kunnen worden gesteld. De OESO [2], die de PISA-enquête [3] (over leesvaardigheid, wiskundige en wetenschappelijke geletterdheid) opstelt, is een organisatie waarvan de hele geschiedenis in het teken staat van de bevordering van het economisch liberalisme. Als studie bureau van de rijke industrielanden pleit de OESO in haar analyses stevast voor de uitbreiding van de marktlogica, de vermindering van de overheidsuitgaven en de ontwikkeling van de privé-sector ten koste van de openbare diensten. (Halimi, 2004; Jennar & Kalafatides, 2007; Laval & Weber, 2002). Het IEA [4], dat de bekende onderzoeken PIRLS [5] (lezen) en TIMSS [6] (wiskunde en natuurwetenschappen) en de minder bekende onderzoeken ICCS [7] (burgerschapsvorming) en ICILS [8] (computervaardigheid) coördineert, is minder politiek geïntereerd [9], maar maakt deel uit van

hetzelfde *"liberale paradigma van de efficiënte school"* (Normand, 2006), waarin de logica van de efficiëntie ("kosten-batenanalyse") wordt gecombineerd met de prioritaire zorg om de ontwikkeling van "menselijk kapitaal", dat individuen en staten in staat moet stellen "competitief" te zijn in de "economie van morgen" (Laval & al., 2012; Normand, 2006; Saillard, 2022). Zoals Draelants en Revaz in navolging van Ogien (2020) aangeven, is dit gemeenschappelijke paradigma van de OESO en het IEA *"ingebed in de context van de neoliberale dominantie (...). In dit kader beoogt de kwantificeringsactiviteit te voldoen aan één eis: efficiëntie, door te trachten elke toewijzing van middelen optimaal te maken op basis van de meting van de verhouding tussen het gewenst resultaat en de verwachte kosten"*. Bijgevolg, zo voegen zij eraan toe, *"kan evidence-based beleid (...) worden geanalyseerd als een beweging die wordt geschraagd door een nieuwe onderwijsethiek die waarde hecht aan prestaties, vergelijken, competitie, markt"* (p. 47). Betekent dit dat de resultaten van deze grote internationale onderzoeken waardeloos zijn? Zeker niet, maar het spoort ons aan om kritisch te kijken naar de gebruikte indicatoren, de analyses en de aanbevelingen.

Partiële en vertekende indicatoren

Wat de indicatoren betreft, wijzen veel critici terecht op de utilitaristische opvatting van kennis waarop enquêtes zoals PISA zijn gebaseerd, met name bij wiskundige en wetenschappelijke geletterdheid (Bodin, 2005; Laval & Weber, 2002; Normand, 2006; Saillard, 2022). Zo testen deze enquêtes vooral het vermogen van de leerlingen om competenties "in het dagelijks leven" te gebruiken, met weinig aandacht voor de beheersing van meer conceptuele kennis. Ideologische vooringenomenheid - voor neoliberalen gaan competenties boven kennis – speelt daarbij een rol. Pragmatisme ook. Om internationale vergelijkingen te maken dient men relatief algemene vragen te stellen, vrij van de bijzonderheden (in termen van kennis) van nationale leerplannen. Men kan zich dan afvragen, zoals de Britse statisticus Harvey Goldstein doet (Stewart, 2013), of deze meting van alleen de kleinste gemene deler van onderwijssystemen in staat is om hun algemene "prestaties" te beoordelen.

Een ander voorbeeld van het niet waardenvrij karakter van indicatoren is te vinden in de ICCS-enquête over burgerschapsonderwijs. De vragen die aan leerlingen worden gesteld gaan over individuele rechten, formele vrijheden, gelijke kansen, respect voor minderheden, houding ten opzichte van verkiezingen en corruptie, of het vermogen om met sociale media om te gaan. De sociaaleconomische invalshoek die is gekozen om "goed burgerschap" te definiëren is die van de "duurzame ontwikkeling". De enquête verzamelt de mening van jongeren over solidariteit en milieu, en vraagt hen naar het ethisch gedrag dat zij individueel in verschillende situaties zouden moeten aannemen. Er wordt niet gepeild naar hun kennis van de macro-economische mechanismen die ongelijkheden veroorzaken en milieucrisis versterken, of ban fundamentele begrippen als kapitalisme, socialisme, marxisme, markteconomie of planeconomie. De index voor burgerschap, die tot stand komt met de resultaten van de enquête, bevat interessante aspecten maar is eenzijdig. De ICSS-enquête evalueert niet in welke mate jongeren in staat zijn de fundamentele beginselen van het economisch stelsel waarin zij leven te begrijpen en in vraag te stellen.

Uit deze beperktheden en vertekeningen afleiden dat deze grote internationale onderzoeken ons enkel informeren over de bekwaamheid tot "liberaal handelen" van leerlingen, is dan weer te kort door de bocht. De PISA-testen meten competenties die liberalen koesteren omdat ze in het kader van de kapitalistische economie ingezet kunnen worden, maar competenties op het gebied van leesvaardigheid, wiskundige en wetenschappelijke geletterdheid zijn absoluut ook noodzakelijk (maar onvoldoende!) voor het uitoefenen van een kritisch burgerschap.

De beperktheid van de genoemde internationale onderzoeken komt ook tot uiting in hun beperkte inhoudelijke focus. Ze zeggen bijvoorbeeld niets over de kennis van leerlingen op gebied van geschiedenis, geografie, literatuur, filosofie of politieke economie, net zoals ze hele gebieden van wetenschap en wiskunde negeren. Om die reden heeft Ovds er zich altijd voor behoed om de PISA-ranking te erkennen als ultieme maatstaf voor de kwaliteit of het niveau van onderwijs. We hebben wel gebruik gemaakt van de schatkamer aan PISA-gegevens om de mate van segregatie en ongelijkheid van onderwijssystemen te meten omdat leerlingen binnen een zelfde onderwijssysteem grosso modo ten opzichte van het onderzoek gelijk zijn [10].

Politieke instrumentalisatie

Een andere kritiek die men op deze grote internationale onderzoeken kan leveren, heeft betrekking op de politieke instrumentatie ervan. Ten eerste stelt men vast dat, zelfs wanneer de PISA-resultaten de OESO er logischerwijs toe zouden moeten brengen sommige van haar liberale dogma's op onderwijsvlak in vraag te stellen, de OESO in feite haar toewijding aan die dogma's alleen maar verdubbelt. Zo illustreert PISA dat de marktmechanismen in het onderwijs segregatie genereren en sociale ongelijkheid versterken. De aanbevelingen van de OESO gaan echter niet in de richting van de vermindering van de marktwerking. De OESO pleit eerder om de overheidsfinanciering voor privéscholen te verhogen (OESO, 2012b) en om het concurrentiemechanisme tussen scholen te optimaliseren (OESO, 2014). PISA stelt vast dat de meest "kwalitatieve" onderwijssystemen diegenen zijn die de leraren beter betalen eerder dan diegenen die de verloning op verdienste baseren. De OESO stelt echter dat het niet "*haalbaar*" is om de verloning van alle leraren te verhogen en dat "*een verloningssysteem, langdurig gebaseerd op prestaties, een geweldige uitdaging vertegenwoordigt*" (OESO, 2012a). Aangezien een verhoging van de overheidsuitgaven haaks staat op een neoliberale visie, zullen structurele maatregelen, die overheidsinvesteringen zouden vereisen, uit de aanbevelingen van de OESO worden geweerd met het argument dat er andere, "efficiëntere" maatregelen zijn. In plaats van de klasgrootte te verminderen, zal de OESO focussen op maatregelen waarbij de leraren zich in toenemende mate moeten verantwoorden, in fase met de doctrine van het *New Public Management*. In een interview in *Le Monde* (Baumard, 2014) verklaarde Andreas Schleicher, hoofd van het PISA-onderzoek, dat "*het Franse onderwijs niet relevant is*" omdat het te veel gericht is op kennis en te weinig op competenties: "*het kan de moderne wereld niet schelen wat je weet*". Hij verklaarde ook dat 40 leerlingen in een klas niet noodzakelijk een probleem hoefde te zijn. Men ziet dus dat de OESO sommige analyses van "op feiten gebaseerde gegevens" omarmt en andere verwerpt omdat ze in strijd zijn met de neoliberale onderwijscatechismus.

Evenzo gebruiken nationale (of regionale) regeringen regelmatig bepaalde PISA-gegevens om vooraf vastgestelde beleidsstandpunten te legitimeren, terwijl ze andere gegevens (uit PISA of andere bronnen) proberen te negeren omdat die minder in overeenstemming zijn met hun beleidsagenda (Davies, 2012; Mons & Pons, 2013). Mons en Pons (2013, p. 16): "*De technische rijkdom en axiologische plasticiteit van PISA (die gemobiliseerd kan worden om zowel neoliberale als sociaaldemocratische aanbevelingen te ondersteunen) maakt een grote verscheidenheid aan toepassingen en aanpassingen mogelijk*". De politici hebben dan vrij spel om die PISA-gegevens te selecteren die in overeenstemming zijn met hun politieke strategie en met de belangen die zij verdedigen. In België bijvoorbeeld heeft geen enkele regering (of regeringspartij) de bevindingen van PISA over het oorzakelijk verband tussen de quasi-marktwerking en de sociale ongelijkheid in het onderwijs, aangegrepen om ambitieuze maatregelen tegen sociale segregatie voor te stellen. Met de stempel "evidence-based" kunnen OESO-aanbevelingen en "wetenschappelijk gelegitimeerde" onderwijshervormingen van nationale (regionale) regeringen worden voorgesteld als "neutraal" en "onbetwistbaar" en zijn

ze daardoor soms in staat elke kritiek in de kiem te smoren (Draelants & Revaz, 2022; Normand, 2006), terwijl ze wel degelijk het resultaat zijn van ideologische en politieke keuzes

De dreiging van "depolitisering door kwantificering"

Wanneer evidence-based onderwijs zich baseert of beweert zich te baseren op grote internationale onderzoeken, bestaat het risico van "depolitisering door kwantificering" (Desrosières, 2008; Draelants & Revaz, 2022; Laval & Weber, 2002; Normand, 2006). Analyses en hervormingen worden dan, omdat ze zouden "gebaseerd zijn op onbetwistbare cijfers", aan de burgers gepresenteerd als politiek neutraal. Er zou geen sprake meer zijn van progressieve of conservatieve onderwijsvoorstellen of hervormingen, maar gewoon van effectieve of ineffectieve hervormingen. Deze subtiele depolitisering wordt bereikt door de subjectiviteit te verbergen die aanwezig is bij de keuze van de indicatoren, de oriëntatie van de aanbevelingen en de selectie van bepaalde bewijzen met uitsluiting van andere om onderwijshervormingen te rechtvaardigen.

Het is noodzakelijk dat de keuze van de indicatoren - en dus van de leerinhouden en de attitudes van de leerlingen die gemeten werden - kritisch wordt onderzocht want zij zijn bepalend voor de doelstellingen van het onderwijs die door de ontwerpers van het onderzoek worden naar voren geschoven. Het is ook belangrijk de aannames en vooronderstellingen die de aanbevelingen sturen, aan het licht te brengen, opdat "*in naam van een gedepoliteerde universele rationaliteit geen fenomeen van mondialisering en vermarkting van het onderwijs*" plaatsvindt, dat uiteindelijk "*een nieuwe mondiale en neoliberale onderwijsorde*" zou instellen (Laval & Weber, 2002).

Dit alles sluit echter een kritisch gebruik van de gegevens uit deze grote internationale onderzoeken niet uit. Deze gegevens zouden ten minste aan deze twee voorwaarden moeten voldoen:

- rekening houden met de beperkingen van indicatoren, vooral wanneer deze worden gebruikt om internationale ranglijsten op te stellen, veeleer dan om interne ongelijkheid van onderwijssystemen te beoordelen;
- voorzichtigheid met betrekking tot analyses, aanbevelingen en legitimaties met de stempel "evidence-based", die soms veeleer het product zijn van een ideologische interpretatie van de resultaten dan de uitkomst van een objectieve en strenge wetenschappelijke procedure.

Olivier Mottint

Referenties

Bodin, A. (2005). Ce qui est vraiment évalué par PISA en mathématiques. Ce qui ne l'est pas. Un point de vue français. *Communication faite dans le cadre du colloque « Teaching mathematics: beyond the PISA survey »*, Paris, du 6 au 8 octobre 2005.

Bressoux, P. (2017). Practice-Based Research : une aporie et des espoirs : Une revue critique de l'article d'Anthony S. Bryk. *Education et didactique*, 11 (3), 123-134.

Davies, P. (2012). The state of evidence-based policy evaluation and its role in policy formation. *National Institute Economic Review*, 219, 41-52.

Desrosières, A. (2008). *Gouverner par les nombres : l'argument statistique II*. Paris : Presses des Mines/Transvalor.

Draelants, H. & Revaz, S. (2022). *L'évidence des faits : La politique des preuves en éducation*. Paris : PUF.

Halimi, S. (2004). *Le Grand Bond en arrière*. Paris : Fayard.

Hirtt, N. (2020). *Schoolse ongelijkheid, ultiem overblijfsel van het unitaire België? Een statistische analyse van de oorzaken van schoolse ongelijkheid in Vlaanderen en in de federatie Wallonie-Bruxelles op basis van PISA 2018*. On-line op website van Ovds : <https://www.skolo.org/nl/2020/03/03/nieuwe-studie-pisa-2018/> of (in het Frans) <https://www.skolo.org/CM/wp-content/uploads/2020/02/PISA-2018-FR.pdf>

Jennar, R.-M. & Kalafatides, L. (2007). *L'AGCS: Quand les États abdiquent face aux multinationales*. Paris : Liber/Raisons d'agir.

Laval, C. & Weber, L. (2002). *Le nouvel ordre éducatif mondial : OMC, Banque mondiale, OCDE, Commission Européenne*. Paris : Syllepse.

Laval, C., Vergne, F., Clément, P. & Dreux, G. (2012). *La nouvelle école capitaliste*. Paris : La Découverte.

Mons, N. & Pons, X. (2013). Pourquoi n'y a-t-il pas eu de « choc PISA » en France ? Sociologie de la réception d'une enquête internationale (2001-2008). *Revue française de pédagogie*, 182, 9-18.

Normand, R. (2006). L'école efficace ou l'horizon du monde comme laboratoire. *Revue des sciences de l'éducation*, 32 (1), 53-70.

OCDE (2012a). La rémunération fondée sur la performance améliore-t-elle la qualité de l'enseignement? *PISA à la loupe*, 16 (5).

OCDE (2012b). Les chèques éducation permettent-ils d'accroître l'égalité des chances ? *PISA à la loupe*, 20 (9).

OCDE (2014). Concurrence entre les établissements d'enseignement : quand est-elle bénéfique ? *PISA à la loupe*, 42 (8).

Ogien, A. (2020). Peut-on se dépendre du pouvoir du chiffre ? *Droit et société*, 105 (2), 479-489.

Saillard, C. (2022). *Les Français et les mathématiques : un regard sociologique sur les indicateurs du niveau des élèves comme instrument de gouvernance des systèmes éducatifs*. <https://variances.eu/?p=6773>

Stewart, W. (2013). Is PISA fundamentally flawed ? *TES Magazine*, 26 juli 2013. <https://www.tes.com/magazine/archive/pisa-fundamentally-flawed-0>

Voetnoten

[1] Dit is zeker in België het geval waar de kloof tussen de gemiddelde scores van de 25% rijkste en de 25% armste leerlingen meer dan 100 punten bedraagt.

[2] OESO : Organisatie voor Economische Samenwerking en Ontwikkeling

[3] PISA : Programme for International Student Assessment

[4] IEA : International Association for the Evaluation of Educational Achievement

[5] PIRLS : Progress in International Reading Literacy Study

[6] TIMSS : Third International Mathematics and Science Study

[7] ICCS : International Civic and Citizenship Education Study

[8] ICILS : International Computer and Information Literacy Study

[9] ... maar wel partnerschappen met de Wereldbank en de Europese Commissie

[10] Voor een meer uitgewerkte argumentatie, zie Hirtt (2020, p. 2).

Deel 3

Is evidence-based onderwijs een wonderoplossing voor ambitieus en rechtvaardig onderwijs?

In het vorige deel hebben we een eerste vorm van evidence-based onderwijs belicht: gebaseerd op de grote internationale onderzoeken. Nu belichten we het experimenteel onderzoek en de meta-analyses die dit onderzoek samenvatten. A priori lijkt het heel zinvol om de resultaten van wetenschappelijk onderzoek te gebruiken ter verbetering van de onderwijspraktijk. Vanuit die optiek lijkt het een goede werkwijze om vooral de meest solide onderzoeken te valideren, die de doeltreffendheid van een interventie testen door middel van een experiment waarbij een interventiegroep en een controlegroep met elkaar worden vergeleken. Een stap verder is dat dit onderzoek wordt gerepliceerd door andere onderzoekers, in andere contexten, om na te gaan of de effecten van een interventie behouden blijven en overdraagbaar zijn op andere situaties. Dit alles vormt het abc van een wetenschappelijk proces, al moet men toegeven dat externe variabelen in klaslokalen minder snel worden uitgesloten dan in een scheikunde-lab of tijdens een therapeutische proef. Dank zij dergelijk onderzoek zou men pedagogische hypothesen kunnen testen en komaf maken met hoogoplopende polemieken over de doeltreffendheid of billijkheid van een bepaalde démarche. Dit zou een nieuw licht kunnen werpen op - maar niet noodzakelijk een oplossing bieden voor - debatten die tot dusver hoofdzakelijk gebaseerd waren op ervaringen of indrukken. Ervaringen en indrukken die zeer zeker hun legitimiteit hebben en in aanmerking moeten worden genomen (zie het eerste artikel van dit dossier), maar die, aangezien zij onvermijdelijk voortvloeien uit de bijzondere context van de persoon die ze vertelt, niet ipso facto kunnen worden veralgemeend.

Tenzij men anti-wetenschappelijk tewerk zou willen gaan of van mening is dat niets in de pedagogie door wetenschappelijk onderzoek kan worden begrepen, is er geen reden om zich te verzetten tegen de bevindingen aangedragen door de evidence-based aanpak. Op voorwaarde dat aanvullend kwalitatief onderzoek niet wordt uitgesloten, zodat de bijzonderheden en de complexiteit van situaties in aanmerking kunnen worden genomen. En op voorwaarde dat de school niet gereduceerd wordt tot een "prestatiefabriek" en dat andere dimensies van de schoolervaring in rekening worden gebracht. En tenslotte, mits enige nederigheid aan de dag wordt gelegd bij het formuleren van aanbevelingen aan leraren.

Wat leert (quasi-)experimenteel onderzoek over pedagogie?

Het is uiteraard onmogelijk om de resultaten van duizenden studies in een paar regels samen te vatten. Bij wijze van voorbeeld en om de diversiteit van de behandelde thema's te illustreren, brengen we hier een opsomming van enkele lessen uit dit experimentele, evidence-based onderzoek.

- hoe sneller en hoe frequenter de feedback is die aan de leerlingen gegeven wordt, hoe beter voor het leerproces, vooral als deze feedback bestaat uit een expliciete correctie van mentale processen in plaats van het louter registreren van fouten die moeten worden "herwerkt" (Kluger & DeNisi, 1996);
- voor dezelfde hoeveelheid instructietijd leidt het "spreiden" van het leren (d.w.z. afwisselen met pauzes en reactiveren over een langere periode) tot een betere verwerving op lange termijn dan de "massale toediening" over een korte periode (Donovan & Radosevich, 1999);

- in het basisonderwijs hebben programma's voor woordenschatopbouw (Stahl & Fairbanks, 1986) en programma's voor vlotheidstraining door het herhaald lezen van kleine teksten (Therrien, 2004) een significant positief effect op de ontwikkeling van vlotheid en begrip;
- leerlingen elkaar iets laten uitleggen (Rosenshine & Meister, 1994) is zeer bevorderlijk voor het leren; gestructureerde settings voor coöperatief leren zijn veel doeltreffender dan methodes die de nadruk leggen op concurrentie tussen leerlingen (emulatieschema's, enz.) (Johnson & al., 1981, 2000; Roseth & al., 2006);
- flipped classes (Bissonnette & Gauthier, 2012; Lo & Hew, 2017) en andere variaties op de digitale schoolmodus (Galand, 2020) toonden tot dusver geen significante meerwaarde voor het leren aan;
- de STAR-studie (Kerckhofs, 2021), toont de grote en blijvende positieve effecten aan van kleinere klassen in de basisschool.

Meer controversieel: de superieure doeltreffendheid van instructie in vergelijking met ontdekking inzake de verwerving van schoolinhoud...

Onderzoek uit de evidence-based stroming lijkt ook aan te tonen dat instructiegerichte methoden - expliciet onderwijs, aanvankelijk sterk geleid door de leraar, stapsgewijs opgebouwd, waarbij in elke fase de verworvenheden worden geverifieerd, met systematische feedback, vaak met een fase van wederzijds of coöperatief onderwijs, en vervolgens overgaand naar geleidelijk steeds minder begeleiding - over het algemeen doeltreffender zijn voor het leren op school dan methoden die uitgaan van ontdekking, onderzoek of project (zie bijvoorbeeld Baker & al., 2002, Bereiter & Kurland, 1981-1982; Bissonnette & al., 2010; Borman, 2003; Gauthier & al., 2005; Gersten & Baker, 2001; Herman & al., 1999; Shymansky & al., 1990; Stockard & al., 2018; Swanson & Sachse-Lee, 2000). De onderzoeken die tot deze conclusie leidden, zijn van verschillende aard.

Allereerst zijn er grootschalige vergelijkende studies (bijvoorbeeld het *Follow Through-project* of het rapport van het *National Reading Panel* in de Verenigde Staten) waarbij (tien)duizenden leerlingen betrokken zijn die vele jaren worden gevolgd. Andere richten zich op de uitvoering van specifieke programma's (*Success for All*, *Direct Instruction*, enz.) waarvan de doeltreffendheid vervolgens wordt gemeten en vergeleken met die van andere onderwijsbenaderingen. Nog andere studies zijn retrospectief, zoals bijvoorbeeld het geval was in Wisconsin. Daar besloot men, na de bekendmaking van de resultaten van het STAR-onderzoek over de voordelen van kleinere klassen, in de jaren negentig kleine klassen in te voeren aan het begin van de schoolloopbaan. Onderzoek naar de impact van deze hervorming bevestigde de voordelen van kleine klassen, maar stelde ook vast dat de voordelen in sommige klassen groter waren dan in andere. De onderzoekers probeerden toen de redenen voor deze contrasterende voordelen te achterhalen door de leraren te observeren wier klassen de grootste of de minste vooruitgang boekten. Uit deze observaties en interviews bleek dat in de klassen met de meeste vooruitgang een groot aantal individualiseringsmomenten hadden plaatsgevonden en een bijzonder expliciet, stapsgewijs voortschrijdend en gestructureerd onderwijs (Molnar & al., 2001) werd toegepast.

Ten slotte zijn er "micro"-studies over het "effect van het werkende voorbeeld" (worked-example effect). Bij dit type experimenteel onderzoek worden twee gerandomiseerde groepen leerlingen vergeleken. Aan de eerste groep legt men uit hoe een specifiek probleem kan opgelost worden door het oplossingsproces voor te doen. De tweede groep wordt voorzien van leermiddelen waarmee zij zelf het oplossingsproces kunnen ontdekken. Leerlingen van beide

groepen krijgen de opdracht soortgelijke problemen op te lossen. Ook dit type onderzoek lijkt de gemiddeld betere effectiviteit van expliciete instructie tegenover de methode van zelfontdekking te bevestigen (zie bijvoorbeeld Atkinson & al., 2000 of Clark & al., 2006).

Onderzoekers verklaren de superioriteit van de expliciete instructie door twee factoren. De leerling die zelf in een ontdekkings situatie wordt geplaatst, bevindt zich in een situatie van een "dubbele taak": tegelijkertijd de oplossing van het probleem vinden én zich de oplossingsaanpak eigen maken, waardoor hij in een toestand van "cognitieve overbelasting" komt. De spanwijdte van het werkgeheugen wordt overschreden. De ontdekkingstaak "interfereert" dan met het proces van het verwerven van de oplossingsaanpak. De leerling die zelf een probleem oplost, doet dat onvermijdelijk met omwegen, pogingen, door valse pistes te volgen, door de fasen van zijn aanpak willekeurig te ordenen. De oplossing van het probleem stelt hem niet altijd in staat om uit deze losse elementen een helder en gestructureerd "script" samen te stellen waarmee hij soortgelijke problemen kan oplossen.

Kritieken, gewogen en (meestal) te licht bevonden ...

De claim over de superieure doeltreffendheid van een expliciet instructieve aanpak, gaf aanleiding tot diverse kritische commentaren. We overlopen enkele kritieken.

Ten eerste is er kritiek op de betrouwbaarheid van bepaalde experimentele protocollen omdat ze onvoldoende rigoureuus worden bevonden. Uiteraard is elk onderzoek van een afzonderlijke experimentele werkwijze voor kritiek vatbaar: geen enkele is perfect, de keuze van de evaluatiecriteria is altijd aanvechtbaar. In de menswetenschappen is het immers onmogelijk één variabele compleet uit te sluiten. Enige voorzichtigheid is dus geboden bij de mogelijke conclusies uit dit onderzoek. De draagwijdte van deze kritiek wordt echter beperkt door het feit dat er honderden studies zijn waarin de doeltreffendheid van onderwijsbenaderingen wordt vergeleken, en dat zij allemaal naar éénzelfde conclusie convergeren en meestal gebaseerd zijn op een methodologie die voldoet aan de normen op dit gebied.

Een andere kritiek luidt dat de aangetoonde doeltreffendheid van instructieve benaderingen slechts oppervlakkig is. Als leerlingen die aan dit soort instructies werden onderworpen, de hun voorgelegde oefeningen in post-tests beter oplossen dan anderen, zou dit komen doordat zij "domweg" zijn geconditioneerd om als automaten oplossingspatronen te reproduceren, zonder een diepgaand begrip van de onderliggende concepten te ontwikkelen. Deze kritiek wijst op een zeer reëel risico: leraren die een instructieve aanpak hanteren en alleen geïnteresseerd zijn in de "prestaties" van hun leerlingen op de examens, zouden hen wel degelijk kunnen trainen in het mechanisch uitvoeren van procedures die zij niet begrijpen. Dit werd waargenomen in sommige Amerikaanse klaslokalen waar het programma *Success for All* werd uitgevoerd (Kozol, 2005, 2006). Dit risico is echter niet onvermijdelijk, aangezien de expliciete aanpak zowel concepten als procedures kan betreffen. Het experiment van Schwonke et al. (2009) over meetkunde gaat in die richting. In dit onderzoek kreeg een eerste groep leerlingen expliciet onderricht zowel betreffende de oplossingsprocedure als de onderliggende concepten; de tweede groep moest dezelfde procedure en dezelfde concepten zelf ontdekken met behulp van feedback. Uit het experiment bleek vooreerst dat de leerlingen in de eerste groep minder leertijd nodig hadden om de aangeleerde procedure goed te beheersen. In een tweede fase vroegen de onderzoekers de leerlingen van beide groepen om nieuwe problemen van hetzelfde type op te lossen door hardop hun aanpak en de concepten waarop deze is gebaseerd uit te leggen. Opnieuw legden de leerlingen in de groep met de expliciete instructies een beter begrip van geometrische concepten aan de dag, dan de leerlingen in de "ontdekkende" groep.

Een andere kritiek luidt dat de doeltreffendheid van instructiemethoden alleen werd aangetoond voor elementair leren en welomschreven procedures. Het klopt dat een groot deel van dit onderzoek gericht was op de verwerving van elementaire lees-, schrijf- en rekenvaardigheden. Studies over het aanleren van complexere vaardigheden zijn weliswaar minder talrijk, maar bevestigen toch de doeltreffendheid van de instructieve methoden (zie bijvoorbeeld Marin & Halpern, 2011).

Een laatste terugkerende kritiek is dat inzake pedagogie alles vooral contextgebonden is en dat elke poging om een doeltreffend pedagogisch instrument te veralgemenen botst op de grote diversiteit van de situaties. In een debat met een andere onderzoeker (Anthony Bryk) geeft Bressoux (2017, p. 125) een interessante reactie op dit argument: *"Wie heeft ooit beweerd dat de handeling van het lesgeven geen bijzondere component heeft? Maar is er niet ook een generieke component, die een groot aantal leerhandelingen gemeen heeft? Uit onderzoek naar de doeltreffendheid van leerkrachten blijkt bijvoorbeeld dat een leerkracht die doeltreffend is in één vak (bv. wiskunde) doorgaans ook doeltreffend is in andere vakken (bv. Frans), dat een leerkracht die doeltreffend is in één klas doorgaans doeltreffend is in alle klassen waarin hij of zij lesgeeft, en dat een leerkracht die in één jaar doeltreffend is, doorgaans ook doeltreffend is in de volgende jaren. Dit zijn allemaal feiten die ingaan tegen de hypothese van een opkomend, onvoorspelbaar fenomeen. Ik ben het roerend eens met Bryk wanneer hij stelt dat het moeilijk is om duidelijk gecontroleerde en gelokaliseerde wetenschappelijke experimenten te generaliseren. Maar moeilijk betekent niet onmogelijk, en er zijn al enkele gevallen van geslaagde generalisatie bekend."*

Zullen evidence-based instructieve praktijken dan volstaan om een democratische school tot stand te brengen?

De democratisering van de scholen vereist noodzakelijkerwijs dat aan alle leerlingen de kennis wordt overgedragen waardoor zij de wereld kunnen begrijpen. Gezien de resultaten van het onderzoek lijken expliciete pedagogieën, sterk gestructureerd en gestuurd door de leraar, een van de essentiële ingrediënten te zijn om deze democratisering van leerprocessen te bevorderen. Dit houdt in dat wij moeten afstappen van een zeker dogmatisme dat expliciete onderwijsmethoden veroordeelt onder het voorwendsel dat die de leerlingen slechts een "geestdodende training" te bieden hebben. Het is noodzakelijk om van deze karikatuur af te stappen. Maar het is evenzeer noodzakelijk om niet te vervallen in het tegenovergestelde dogmatisme, dat instructieve methoden tot hét wondermiddel voor pedagogie en onderwijshervorming zou verheffen. Dit zou ons inderdaad blootstellen aan vier mogelijke ontsporingen.

1. De uitsluiting van elke andere pedagogische benadering.

Ongetwijfeld meegesleept door hun enthousiasme, roepen sommige onderzoekers op tot het verbieden van "constructivistische" pedagogieën, omdat ze inherent ondoeltreffend zouden zijn. Hoewel instructieve methoden gemiddeld doeltreffender zijn dan op ontdekking gebaseerde methoden, blijft het een feit dat een sterk gestructureerde praktijk van "actieve pedagogie", waarbij exploratie en explicatie rigoureuus worden verwoord, en waarbij het leren nauwgezet wordt gepland, ook tot uitstekende resultaten op het gebied van schoolonderwijs kan leiden, zoals blijkt uit het onderzoek van Reuter (2007) op een Freinet-school. Uit ander onderzoek blijkt ook dat leerlingen, zodra zij een bepaald niveau van deskundigheid hebben bereikt, beter leren door te proberen zelf problemen op te lossen dan via expliciete methodologie (Kalyuga & al., 2001, 2003).

2. Pedagogie (en het onderwijs) herleiden tot een kwestie van prestatie.

Efficiëntie in termen van overdracht van schoolkennis is zeker een fundamenteel criterium voor pedagogische keuzes. Als dit echter het enige criterium zou zijn, wordt de school gereduceerd tot louter technische overwegingen, die alleen de "prestaties" van de leerlingen voor ogen hebben. Het uitvoeren van projecten die de in de klas geleerde kennis tot leven brengen, samenwerken om probleemsituaties op te lossen, debatteren over een cultureel werk in de klas of gewoon het gezamenlijk lezen van een gedicht zijn misschien niet de meest "efficiënte" werkwijzen om competenties te ontwikkelen. Blijft het feit dat deze praktijken onmisbaar zijn voor het nastreven van andere doelstellingen: het ontwikkelen van solidariteit en samenwerking, het ontdekken dat kennis in de loop van de geschiedenis is opgebouwd door mensen te confronteren met concrete problemen (Delire, 2021), of het besef dat de in de klas verworven kennis en vaardigheden het mogelijk maken de wereld te begrijpen en te veranderen, en niet alleen voor de examens te slagen. Dank zij andere van deze zogenaamd "inefficiënte" praktijken zullen de leerlingen plezier kunnen beleven aan literatuur, "schoolse" onderwerpen leuk vinden, in culturele werken redenen tot verontwaardiging en engagement ontdekken, of manieren om naar de wereld te kijken die de zekerheden die zij tot dusver hadden, aan het wankelen brengen.

3. De enge standaardisatie van het onderwijs.

In Angelsaksische landen leidde de aangetoonde gemiddeld superieure doeltreffendheid van instructiemethoden tot de invoering van "kant-en-klare" pedagogische hulpmiddelen die horen bij deze pedagogische tendens en tot steeds meer gedetailleerde pedagogische voorschriften. Leraren reduceerden hun beroep geleidelijk tot de strikte toepassing van evidence-based onderwijsscripts. In Vlaanderen laat het rapport van de Commissie "*Beter Onderwijs*", dat de invoering van directe instructie bepleit, vrezen voor een dergelijke ontsporing. Dit betekent niet dat wij zonder meer gekant zijn tegen het strategische gebruik van "recepten" in de pedagogie - zulke recepten zijn er in elk beroep, van vliegtuigpiloten tot chirurgen - of dat de pedagogische vrijheid onbeperkt zou moeten zijn. Maar deze ontwikkeling naar enge pedagogische voorschriften bedreigt leraren met de uitholling van hun beroep, en leerlingen met de standaardisering van het schoolleven, dat gereduceerd wordt tot de eeuwigdurende herhaling van soortgelijke pedagogische scenario's. Dit sluit elke mogelijkheid uit voor de leraar om rekening te houden met de eigenheid van situaties, net zoals het elke opening van de school naar de buitenwereld in de weg staat; de creativiteit van de leraar en zijn inbreng in de pedagogische handeling wordt eenvoudigweg weggevaagd. Beklemd door technische voorschriften dreigt de schoolervaring dan de vorm aan te nemen van een monotoon "*pedagogisch Taylorisme*", een "*regime van intensieve training tot prestaties*" (Hargreaves, 2003) dat weinig bevorderlijk is voor emancipatie. Het is zeker belangrijk dat leerkrachten via de opleiding een betere kennis verwerven van de resultaten van wetenschappelijk onderzoek over pedagogie en we kunnen stellen dat we daar nog lang niet aan toe zijn. Maar deze wetenschappelijke opleiding zou hen beter in staat moeten stellen diverse pedagogische middelen, elk met hun eigen diverse doelstellingen, strategisch in te zetten, in plaats van leraren te reduceren tot de rol van uitvoerders.

4. Geen zicht op de effecten van de structuren van onderwijssystemen.

Evidence-based onderzoekers schrijven regelmatig slogans die erop hameren dat scholen en leraren "het verschil maken". Zij gaan daarin zelfs zo ver dat zij volledig lijken te vertrouwen op het "schooleffect" (leiderschap van de directie, enz.) en het "leerkrachteffect" (pedagogie, handhaving van discipline, enz.) om het schoolsysteem doeltreffender en meer egalitair te maken. Dit komt doordat het experimentele onderzoek waarop zij zich baseren alleen factoren binnen de onderwijsstelsels kan bestuderen. Het boek "*Visible Learning*" (Hattie, 2009), dat een van de belangrijkste referenties is van de evidence-based stroming, bestudeert de invloed van 138 variabelen op het leren op school, maar in deze lange lijst komen de onderwijsmarkt

en de effecten daarvan op onderwijsongelijkheden niet voor. Maar, zoals Nico Hirtt (2020, p. 19) aantoonde, "zijn de structurele kenmerken van onderwijssystemen (quasi-markt en streaming) de belangrijkste factoren in de ongelijkheid van deze systemen. Zij verklaren meer dan de helft van de intra-Europese verschillen op dit gebied". Door deze structurele factoren te verwaarlozen en zich uitsluitend op de pedagogie te richten, komen leerkrachten niet alleen onder druk te staan, maar kunnen zij ook niet nadenken over de structurele hervormingen die nodig zijn om een ambitieus schoolsysteem voor alle leerlingen op te bouwen.

Om te besluiten

Al geven de resultaten van experimenteel onderzoek aan - en dit is belangrijk - dat expliciete pedagogie onontbeerlijk is om een ambitieuze en egalitaire school op te bouwen, twee valkuilen moeten worden vermeden.

En eerste valkuil zou er in bestaan dat men een exclusieve en enge toepassing van deze pedagogieën zou voorschrijven en dus elk beroep op actieve pedagogieën en op hun mogelijke bijdrage buiten het strikte criterium van doeltreffendheid zou uitsluiten. Net als Goigoux (2011, p. 24) verdedigen wij integendeel een zeker pedagogisch eclecticisme, "*dat de verworvenheden van actieve pedagogie wil verzoenen met de vereisten van expliciete en gestructureerde pedagogie. Het combineert fasen van declaratief onderwijs (uiteenzetting van regels, procedures of begrippen), fasen van geleid oplossen onder nauw toezicht van de leraar en fasen van vallen en opstaan, verkennen of ontdekken (gebruik van probleemsituaties), waarbij de meeste aandacht uitgaat naar fasen van training, oefening of spel die het onthouden van begrippen en het automatiseren van procedures bevorderen*".

We moeten er ons voor hoeden overdreven verwachtingen te stellen in pedagogische hervormingen en te denken dat instructiegerichte pedagogieën de wonderoplossing zijn voor alle schoolproblemen waarmee zoveel leerlingen uit de arbeidersklasse worden geconfronteerd. De verbetering van de pedagogische praktijk is een niet te onderschatten hefboom maar pedagogie speelt niet de hoofdrol bij sociale ongelijkheid in het onderwijs. Zoals Nico Hirtt (2020) aantoonde, zijn alle onderwijssystemen in Europa waar marktmechanismen een grote rol spelen sterk gesegregeerd en daardoor gekenmerkt door grote sociale ongelijkheid. Zolang de marktmechanismen en de daaruit voortvloeiende segregatie niet aangepakt worden, zal een hervorming die uitsluitend betrekking heeft op de pedagogie in het beste geval enkel matige en in het slechtste geval minimale resultaten opleveren. Deze vaststelling lijkt ons zowel voor het Vlaams als voor het Franstalig onderwijs relevant.

Olivier Mottint

Referenties

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. W. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70, 181–214.

Baker, S., Gersten, R. & Lee, D.S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, 103 (1), 51-73.

Bereiter, C. & Kurland, M. (1981-1982). A Constructive Look at Follow Through Results. *Interchange on Educational Policy*, 12 (1), 1-22.

Bissonnette, S. Richard, M., Gauthier, C. & Bouchard, C. (2010) Quelles sont les stratégies d'enseignement efficaces favorisant les apprentissages fondamentaux auprès des élèves en

difficulté de niveau élémentaire ? Résultats d'une méga-analyse. *Revue de recherche appliquée sur l'apprentissage*, 3 (1).

Borman, G.D., Hewes, G.M., Overman, L.T. & Brown, S. (2003). Comprehensive school reform and achievement : A meta-analysis. *Review of Educational Research*, 73 (2), 125-230.

Bressoux, P. (2017). Practice-Based Research : une aporie et des espoirs. Une revue critique de l'article d'Anthony S. Bryk. *Education et didactique*, 11 (3), 123-134.

Clark, R.C., Nguyen, F., and Sweller, J. (2006). *Efficiency in learning: evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.

Delire, J. M. (2021). Mathématiques et Histoire, une association porteuse de sens ? *L'École démocratique*, 87, 18-21.

Donovan, J. J. & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect : Now you see it, now you don't. *Journal of Applied Psychology*, 84 (5), 795-805.

Galand, B. (2020). Le numérique va-t-il révolutionner l'éducation ? *Cahiers du Girsef*, 120.

Gauthier, C., Mellouki, M., Siard, D., Bissonnette, S. & Richard, M. (2005). Quelles sont les pédagogies efficaces ? Un état de la recherche. *Les Cahiers du débat*, janvier 2005.

Gersten, R. & Baker, S. (2001). Teaching expressive writing to students with learning disabilities : A meta-analysis. *The Elementary School Journal*, 101 (3), 251-272.

Goigoux, R. (2011). Une pédagogie éclectique au service des élèves qui ont le plus besoin de l'école. *La nouvelle revue de l'adaptation et de la scolarisation*, 52, 21-30.

Hargreaves, A. (2003). *Teaching in the Knowledge Society : Education in the Age of Insecurity*. New York: Teachers College Press.

Hattie, J. (2009). *Visible Learning : A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York : Routledge.

Herman, R., Aladjem, D., McMahon, P., Masem, E., Mulligan, I., O'Malley, A., Quinones, S., Reeve, A., & Woodruff, D. (1999). *An Educators' Guide to Schoolwide Reform*. Washington : American Institutes for Research.

Hirtt, N. (2020). *Sociale ongelijkheid, ultiem overblijfsel van het unitaire België ?* <https://www.skolo.org/nl/2020/03/03/nieuwe-studie-pisa-2018>

Johnson, D. W., Johnson, R. T. & Stanne, M. B. (2000). *Cooperative learning methods : A meta-analysis*.

https://www.researchgate.net/publication/220040324_Cooperative_learning_methods_A_meta-analysis

Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D. & Skon, L. (1981). Effects of cooperative, competitive and individualistic goal structures on achievement : A meta-analysis. *Psychological Bulletin*, 89 (1), 47-62.

Kalyuga, S., Ayres, P., Chandler, P. & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, 38 (1), 23-31.

Kalyuga, S., Chandler, P., Tuovinen, J. & Sweller, J. (2001). When Problem Solving Is Superior to Studying Worked Examples. *Journal of Educational Psychology*, 93 (3), 579-588.

Kerckhofs, J.-P. (2021). *Met veel of weinig leerlingen in de klas, maakt het een verschil?* <https://www.skolo.org/nl/2021/04/08/met-veel-of-weinig-leerlingen-in-eeen-klas-maakt-het-eeen-verschil/>

- Kluger, A. N. & DeNisi, A. S. (1996). The effects of feedback interventions on performance. : A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254-284.
- Kozol, J. (2005). Confections of Apartheid: A Stick-and-Carrot Pedagogy for The Children of Our Inner-City Poor. *Phi Delta Kappan*, 87 (4), 265-275.
- Kozol, J. (2006). Success for All : Trying to Make an End Run around Inequality and Segregation. *Phi Delta Kappan*, 87 (8), 624-626.
- Lo, C. K. & Hew, K. F. (2017). A Critical Review of Flipped Classroom Challenges in K-12 Education : Possible Solutions and Recommendations for Future Research. *Research and Practice in Technology Enhanced Learning*, 12, article n°4.
- Marin, L. M. & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity*, 6 (1), 1-13.
- Molnar, A., Smith, P., Zahorik, J., Halbach, A., Ehrle, K., Hoffman, L.M. & Cross, B. (2001). 2000-2001. *Evaluation Results of the Student Achievement Guarantee in Education (SAGE) Program*. <https://nepc.info/sites/default/files/epru-0201-104.pdf>
- Reuter, Y. (2007). *Une école Freinet : Fonctionnement et effets d'une pédagogie alternative en milieu populaire*. Paris : L'Harmattan.
- Rosenshine, B. & Meister, C. (1994). Reciprocal teaching : A review of research. *Review of Educational Research*, 64 (4), 479-530.
- Roseth, C. J., Fang, F., Johnson, D. W. & Johnson, R. T. (2006). *Effects of cooperative learning on middle school students : A meta-analysis*. Paper voorgesteld op de Annual Meeting of the American Educational Research Association, San Francisco.
- Schwonke, R., Renkl, A., Kreig, C., Wittwer, J., Alven, V. & Salden, R. (2009). The worked-example effect: Not an artefact of lousy control conditions. *Computers in Human Behavior*, 25, 258-266.
- Shymansky, J. A., Hedges, L. V. & Woodworth, G. (1990). A Reassessment of the Effects of Inquiry-Based Science Curricula of the 60's on Student Performance. *Journal of Research in Science Teaching*, 27 (2), 127-144.
- Stahl, S. A. & Fairbanks, M. M. (1986). The effects of vocabulary instruction : A model-based meta-analysis. *Review of Educational Research*, 56 (1), 72-110.
- Stockard, J., Wood, T. W., Coughlin, C. & Rasplika Khoury, C. (2018) The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Review of Educational Research*, 88 (4), 479-507.
- Swanson, H. L. & Sachse-Lee, C. (2000). A Meta-Analysis of Single-Subject-Design Intervention Research for Students with LD. *Journal of Learning Disabilities*, 33 (2), 114-136.
- Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading : A meta-analysis. *Remedial and Special Education*, 25 (4), 252-260.