

Wat klopt er niet bij John Hattie?

Bob Slavin is hoogleraar aan het 'Institute for Effective Education' aan de Universiteit van York en directeur van het 'Centrum voor Onderzoek en Hervorming in het onderwijs' aan de Johns Hopkins-universiteit in Baltimore. Niet vreemd dus dat hij met belangstelling kennis nam van de publicaties John Hattie, zijn collega aan de Universiteit van Melbourne. Met name zijn boek *Visible Learning* maakte Hattie in veel landen bekend. Ook in Nederland is die publicatie onder de titel *Leren zichtbaar maken* bekend geworden en graag geciteerd.

Slavin werd met name getripped door de claim van Hattie dat hij elk onderzoeksgebied dat raakte aan onderwijzen en leren bij zijn onderzoek had betrokken. Dat leidde ertoe dat Hattie op deze manier factoren ordende van 1 tot 138, op basis van de gemeten effectgrootte of geconstateerde prestaties. In zijn presentaties weet Hattie die eenvoud van dit overzicht heel helder over te brengen, tot vreugde van veel leraren.

Slavin geeft aan dat hij geneigd is om alles wat te mooi lijkt om waar te zijn het voordeel van de twijfel te geven. Het maakte hem echter wel nieuwsgierig naar de achterliggende studies die die grote effectomvang zouden onderbouwen. Hij ging daarom samen met zijn collega Marta Pellegrini van de Universiteit van Florence aan het werk om de achterliggende bewijzen van Hatties claims op te sporen en te onderzoeken.

Uit dat onderzoek bleek dat Hattie niet alleen gebruik maakte van meta-analyses (d.w.z. overzichten van grotere aantallen onderzoeken) waarin aanzienlijke tegenstrijdigheden, zg. bias, voorkwamen. Hij bleek dergelijke meta-analyses ook nog eens te hebben samengevoegd tot meta-meta-analyses. Dat versterkte die tegenstrijdigheden nog meer. Dat betekent dat de factoren die het ene onderzoek probeerde uit te sluiten, juist in een andere studie het onderzoeksonderwerp konden zijn. Dat maakt dat conclusies van beide onderzoeken niet zomaar kunnen worden samengevoegd, alsof ze hetzelfde pretenderen. Slavin en Pellegrini concluderen daarom dat zijn aanpak en de daarop gebaseerde conclusies verregaand onjuist zijn.

Slavin geeft aan dat hij goed begrijpt waarom leraren Hatties conclusies zo aantrekkelijk vinden. Hij kan er goedvol over vertellen en weet de door hem opgestelde rangorde naar effectgrootte met kleurtjes heel inzichtelijk te maken. Effectgroottes tussen 0,00 en 0,15 heten 'ontwikkelingseffecten', d.w.z. dat ze voortkomen uit de natuurlijke ontwikkeling van kinderen. Tussen 0,15 en 0,40 heten ze leraareffecten, d.w.z. dat deze effecten voortkomen uit het feit dat een leraar iets toevoegt, maakt niet uit wat. Tussen 0,40 en 1,20 is er sprake van gewenste effecten. Hattie benadrukt hiermee dat alle factoren en programma's die een effect hebben lager dan die 0,40 genegeerd kunnen worden. Die voegt volgens hem niets toe aan wat een leraar toch al gedurende een schooljaar toevoegt door les te geven. Hij vergelijkt dat dan met kinderen die niet naar school zouden gaan. Bij die kinderen zijn wel effecten tussen 0,00 en 0,15 te vinden, volgens hem. "When teachers claim that they are having a positive effect on achievement or when a policy improves achievement, this is almost always a trivial claim: virtually everything works. One only needs a pulse and we can improve achievement." (Hattie, 2009, p. 16). Hiermee gaat hij volgens Slavin voorbij aan de essentie van wat effectmeting is, namelijk dat de gevonden effecten worden vergeleken met die van vergelijkbare leerlingen in een controlegroep, die de betreffende aanpak of omstandigheden niet hebben meegemaakt. Was dit wel het geval geweest dan kon een geconstateerde effectgrootte van bv. 0,15-0,25 heel belangrijk zijn, benadrukt hij.

De centrale claims van Hattie zijn:

- Bijna alles werkt
- Elk effect dat kleiner dan 0,40 kun je negeren
- Het is mogelijk om tot een betekenisvolle rangorde van onderwijskundige factoren te komen door deze met elkaar te vergelijken op basis van de gemiddelde uitkomsten van meta-analyses.

Dat klinkt aantrekkelijk door de eenvoud en begrijpelijkheid. Helaas is het onjuist.

Veel van de gebruikte onderzoeken hadden deze opzet niet (bv. omdat het explorerend onderzoek betrof) en ook de kwaliteit van zowel de onderzoeken als de daarop gebaseerde meta-analyses die Hattie wereldwijd verzamelde en gebruikte verschillen nogal. Of de onderzoeksgroep was te klein, of er was geen voormeting en/of nameting of de te vergelijken groepen bleken niet at random te zijn gekozen. Hij is niet kritisch daarmee omgegaan en heeft de onderliggende tegenstrijdigheden niet meegenomen bij zijn conclusies. Hij verwerpt die kritiek zelfs: "There is...no reason to throw out studies automatically because of lower quality" (Hattie, 2009, p. 11) Bovendien is door het samenvoegen en middelen van al die meta-analyses de relatie tussen gevonden effect en de aanpak en omstandigheden waarop dat effect is gebaseerd helemaal buiten beeld. Als de gebruikte onderzoeken niet hetzelfde doel en/of concrete onderzoeksonderwerp of vergelijkbare

omstandigheden betroffen, dan is het niet verantwoord hierop cijfermatige conclusie te trekken en te doen alsof dat effect 'wetenschappelijk' is bewezen.

Slavin heeft het onderwerp dat derde op Hattie's ranglijst staat (feedback geven) uitgekozen voor nadere analyse. Volgens die ranglijst heeft feedback geven een gemiddeld effect van 0,79. Dit is gebaseerd op 12 meta-analyses. Daaruit koos Slavin de analyses met de hoogste gemiddelde scores om nader te onderzoeken. Het hoogste gemiddelde effect (1,24) werd gemeld bij een review van onderzoeken naar het effect van verschillende aanpakken die leerlingen in het speciaal onderwijs moesten versterken en bemoedigen, uitgevoerd in 1985-1986. Het betrof 25 studies over in totaal 35 leerlingen die individueel begeleid werden. Het is bekend dat de effecten van individuele begeleiding grotere effecten oplevert dan groepsaanpakken.

De analyse met het op een na hoogste effect (1,13) betrof een review uit 1982. De thema's daarin waren instructie-aanwijzingen, deelname en corrigerende feedback. Een van de onderliggende studies betrof een onderzoek op één onderwerp bij twee leerlingen dat leidde tot een effectgrootte van 11,81. Dat is te vergelijken met toename van het IQ van 100 naar 277... Het is later gecorrigeerd naar 4,99, wat nog te vergelijken is met een toename van 75 IQ-punten. Veel van de andere onderzoeken uit dit review waren met elkaar verbonden, zonder dat de verschillen in input werden geëxpliciteerd, of hadden geen controlegroep en/of geen voor en nameting.

Een meta-analyse van Rummel and Feinberg (1988), had een effectgrootte van 0,60. Dat is waarschijnlijk het grappigste onderdeel van de Hattie & Timperley (2007) meta-meta-analyse. Het bestaat geheel uit korte laboratoriumstudies over de mate waarin betaald worden voor of op een andere manier gestimuleerd worden om op een later tijdstip deelname aan een activiteit, waarvoor de betrokkene al gemotiveerd was, te beperken. Zij benoemen het als een positief effect als de deelnemer inderdaad minder deed tijdens de activiteiten waarvoor hij werd betaald om dat te doen. De reviewers besloten de studies positief te labelen als hun bevindingen klopten met de theorie, d.w.z. dat feedback en bevestiging de latere deelname aan de aanvankelijk gewilde activiteit verminderde. In werkelijkheid wijst die gevonden positieve effectgrootte op een negatief effect van die feedback op de uitvoering van die activiteit.

Zo kan Slavin nog veel meer voorbeelden noemen. Het komt er steeds op neer dat Hattie meta-meta-analyses grote aantallen meta-analyses van allerlei aard bij elkaar pakt, zonder al te veel te letten op de betekenis of de kwaliteit van de originele studies of van de door anderen gemaakte meta-analyses daarvan.

In onderzoeksoverzichten die Slavin daar tegenover stelt, zoals het What Works Clearinghouse (2007), of zijn eigen Best-Evidence Syntheses (www.bestevidence.org) of het Evidence for ESSA (www.evidenceforessa.org) blijkt dat effectscores van 0,40 (wat voor Hattie het omslagpunt is) bijna nooit voorkomen. Dit komt doordat in deze drie onderzoeksverzamelpunten elk onderzoek heel kritisch wordt gescreend op de vraag of de eisen die je aan wetenschappelijk onderzoek moet stellen, zijn gerespecteerd. Zo niet, dan wordt het daar niet opgenomen. Helaas is Hattie niet zo kritisch geweest, waardoor aan zijn bevindingen weinig waarde kan worden toegekend.

Het is daarom niet aan te raden te doen alsof Leren zichtbaar maken ons richting kan geven bij de keuzes die we moeten maken om (het effect van) ons onderwijs te verbeteren, d.w.z te zorgen dat alle leerlingen zich ononderbroken kunnen ontwikkelen, zoals artikel 8 van de WPO voorschrijft.

Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45 (5), 283-292.

Hattie, J. (2009). *Visible learning*. New York, NY: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81-112.

Lysakowski, R., & Walberg, H. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal*, 19 (4), 559-578.

Rummel, A., & Feinberg, R. (1988). Cognitive evaluation theory: A review of the literature. *Social Behavior and Personality*, 16 (2), 147-164.

Skiba, R., Casey, A., & Center, B. (1985-86). Nonaversive procedures I the treatment of classroom behavior problems. *The Journal of Special Education*, 19 (4), 459-481.

What Works Clearinghouse (2017). *Procedures handbook 4.0*. Washington, DC: Author.

Robert (Bob) Edward Slavin (1950, [Bethesda, Maryland, VS](#)) is een sociaal wetenschapper die gespecialiseerd is in onderzoek naar de ontwikkeling van kinderen. Hij is [hoogleraar](#) aan het 'Institute for Effective Education' aan de [Universiteit van York](#) en directeur van het 'Centrum voor Onderzoek en Hervorming in het onderwijs' aan de [Johns Hopkins-universiteit](#) in [Baltimore](#). Ook is hij voorzitter van de 'Success for All' stichting waarvan hij medeoprichter is.

Dolf Janson, op basis van <https://robertslavinsblog.wordpress.com/2018/06/21/john-hattie-is-wrong/>