

# Evidence-based VVE? De gebrekkige empirische onderbouwing van Voor- en Vroegschoolse Educatie

Geert Driessen

---

## Evidence-based ECE? The inadequate empirical foundation of Early Childhood Education

**G** *Orthopedagogiek: Onderzoek en Praktijk*, 55 (5-6), 127-141  
© Garant | ISSN 2211-6273 | mei-juni 2016

### SAMENVATTING

Voor- en Vroegschoolse Educatie (VVE) betreft educatieve stimuleringsprogramma's voor peuters en kleuters die vanwege factoren gelegen in de thuissituatie onderwijsachterstanden dreigen op te lopen. VVE kent al een lange geschiedenis en er wordt jaarlijks ruim 250 miljoen euro aan besteed. Het is dan ook merkwaardig dat de empirische basis ervoor onduidelijk is, zeker in het licht van een trend naar meer *evidence-based* overheidsbeleid. In dit artikel wordt op zoek gegaan naar de oorsprong van VVE en beoordeelt op kritische wijze de kwaliteit van het bewijs dat wordt opgevoerd voor de positieve werking ervan. Bovendien worden drie recente Nederlandse publicaties op het gebied van VVE samengevat en gewogen. Het bewijs voor positieve effecten blijkt vooralsnog te ontbreken.

**Kernwoorden:** VVE, onderwijsachterstanden, beleid, evidence-based, onderzoek, compensatieprogramma

### SUMMARY

Early Childhood Education (ECE) concerns educational stimulation programs aiming at young children who are at risk of educational disadvantage. ECE in the Netherlands has a long history and each year the Dutch government invests some 250 million euros. Therefore it is remarkable that its empirical foundation is unclear, especially in light of a trend towards more evidence-based policy. This paper investigates the origin of ECE and critically examines the quality of the evidence supplied. In addition, three recent Dutch publications on ECE are summarized and reviewed. For the time being, the evidence for positive effects proves to be lacking.

**Keywords:** Early Childhood Education, educational disadvantage, policy, evidence-based, research, compensation program

### OVER DE AUTEURS

Dr. Geert Driessen is onderwijskundige en pedagoog en was tot recent als onderzoeker op het gebied van onderwijsachterstanden verbonden aan het ITS van de Radboud Universiteit. *E-mail:* g.driessen@its.ru.nl  
*Web:* www.geertdriessen.nl

### ABOUT THE AUTHOR

Dr. Geert Driessen is an educational researcher specialized in educational disadvantage at Radboud University Nijmegen. *E-mail:* g.driessen@its.ru.nl *Web:* www.geertdriessen.nl

Hierna volgt de oorspronkelijke auteursversie van dit artikel. Voor de uiteindelijk gepubliceerde versie wordt verwezen naar:

*Orthopedagogiek: Onderzoek en Praktijk*, 55 (5-6), 127-141, mei-juni 2016.

Uitgegeven bij Garant.

## **1. Voor- en Vroegschoolse Educatie**

Voor- en Vroegschoolse Educatie (VVE) vormt momenteel het belangrijkste onderdeel van het Onderwijsachterstandbeleid (OAB). Dit beleid is erop gericht onderwijsachterstanden die het gevolg zijn van sociale, economische en culturele factoren gelegen in het gezin waarin een kind opgroeit te voorkomen of bestrijden. OAB bestaat al veertig jaar, zij het onder verschillende benamingen (Driessen, 2013). Omdat bleek dat aanpakken op latere tijdstippen in de schoolloopbaan niet zoveel effect (meer) hadden, is het zwaartepunt in het beleid verschoven naar de voor- en vroegschoolse fase. Jonge kinderen die tot de doelgroep van het beleid horen, komen tussen 2.5- en 6-jarige leeftijd in aanmerking voor een (aanvullend) educatief programma, dat gericht is op de stimulering van de ontwikkeling. Dit ter compensatie van het gemis aan (de 'juiste') educatieve stimulering thuis. Een dergelijk programma wordt uitgevoerd op kinderdagverblijven, peuterspeelzalen en in de kleutergroepen van de basisschool. VVE op kinderdagverblijven en peuterspeelzalen (de voorschoolse fase) valt onder de verantwoordelijkheid van de gemeenten; VVE op basisscholen is de verantwoordelijkheid van de schoolbesturen. De formele doelgroep bestaat uit kinderen van laagopgeleide ouders (de zogenoemde 'gewichtenleerlingen'). Voor deze kinderen ontvangen gemeenten en besturen vanuit het ministerie van OC&W een vergoeding. Daarnaast zijn zij echter vrij om ook (nog) andere doelgroepen te definiëren, hetgeen met name gemeenten ook zeer veelvuldig doen. Punt is dan wel dat zij dat uit eigen middelen moeten bekostigen. Ook bij de uitvoering hebben gemeenten en besturen, binnen bepaalde randvoorwaarden, veel vrijheid. Ze kunnen daarvoor kiezen uit een ruim aanbod aan educatieve stimuleringsprogramma's waarvan verwacht wordt dat die zich richten op de vier ontwikkelingsdomeinen taal, rekenen, sociaal-emotioneel, en motorisch. Geschat wordt dat circa 70000 kinderen deelnemen aan voorschoolse educatie, en 35000 aan vroegschoolse educatie (CPB, 2016). Het jaarlijkse budget voor voorschools VVE, gebaseerd op het formele doelgroepcriterium, bedraagt 254 miljoen euro, ofwel 3628 euro per kind. Dit is los van het bedrag dat gemeenten uit eigen middelen inzetten. Hoeveel aan voorschools VVE wordt besteed is niet duidelijk; dit budget zit in de lumpsum die schoolbesturen van het ministerie ontvangen. In totaal zijn er de afgelopen decennia vele miljarden in VVE gepompt. Zonder adequaat bewijs dat het werkt. Daarom is VVE niet onomstreden. Hierna wordt kort ingegaan op enkele probleempunten. Het belangrijkste is wel de empirische onderbouwing van de effectiviteit: welk bewijs is er überhaupt dat VVE werkt? Dat vormt het volgende onderwerp. Onlangs zijn er enkele spraakmakende Nederlandse studies gepubliceerd die expliciet ingaan op die effectiviteit. Deze worden aan een kritische beschouwing onderworpen. Afsluitend wordt de balans opgemaakt.

## **2. Probleempunten**

Het VVE kent een aantal probleempunten (Driessen, 2012; Driessen, Veen & Van Daalen, 2015). Een eerste betreft de kinderen die het onderwerp van het beleid vormen, *de doelgroep*. Officieel bestaat deze uit kinderen van laagopgeleide ouders. Daar ontvangen de schoolbesturen en gemeenten geld voor. In de praktijk definiëren zij echter zelf ook eigen doelgroepen. Zo worden vaak ook kinderen van migranten tot de doelgroep gerekend, kinderen waarvan de ouders geen (goed) Nederlands spreken, kinderen van ouders die aan de drank of drugs zijn verslaafd, kinderen die in een achterstandswijk wonen, en ook kinderen met gedrags- en leerproblemen en -stoornissen. Dit laatste duidt erop dat het oorspronkelijke achterstandsperspectief (vanuit het OAB: milieufactoren) vermengd wordt met het zorgperspectief (vanuit het Passend Onderwijs:

individuele ontwikkelingskenmerken), en dat lijkt steeds meer te gebeuren.<sup>1</sup> Bij dit alles spelen drie punten een rol. Omdat elke gemeente zijn eigen doelgroepen definieert, is er landelijk nogal wat variatie ontstaan. Het gevolg daarvan is ongelijke behandeling: in sommige gemeenten krijgt een kind wel ondersteuning, in andere een vergelijkbaar kind niet. Wat ook problematisch is, is dat voor het identificeren van de doelgroepen (screenings)instrumenten worden gebruikt die niet voor dit doel zijn ontwikkeld en waarvan de betrouwbaarheid en validiteit onbekend dan wel discutabel is. Van het criterium voor de landelijk gedefinieerde doelgroep, een laag ouderlijk opleidingsniveau, is inmiddels vastgesteld dat de predictieve validiteit zeer onder de maat is (Driessen, 2015). Het gevolg van dit alles is willekeur en een groot aantal vals-positieven, ofwel kinderen die ten onrechte ondersteuning ontvangen, en ook – en dat is waarschijnlijk ernstiger – vals-negatieven, ofwel kinderen die ten onrechte geen steun ontvangen.

Een tweede punt betreft de kwaliteit van de uitvoerders, *de leidsters* (c.q. pedagogisch medewerkers). De kern van het VVE betreft stimulering van de taalontwikkeling. Oorspronkelijk vormden vooral kinderen van laagopgeleide migranten de doelgroep. Het is dan van groot belang dat de leidsters zelf de Nederlandse taal goed beheersen. Daar schortte het tot recent echter nogal aan. Om die reden zijn leidsters getoetst op hun taalvaardigheid en indien ze niet voldeden op taalcursus gestuurd. Naar verwachting is het niveau momenteel wat gestegen, maar of dat echt voldoende is, is maar zeer de vraag. Ook is gebleken dat leidsters in het algemeen wel voldoen wat betreft hun pedagogische vaardigheden; duidelijk lager scoren ze echter met betrekking tot hun educatieve vaardigheden. Daar zou in de opleiding veel meer aandacht voor moeten zijn. Momenteel is formeel gezien mbo-niveau voldoende voor deze functie, maar gelet op de zwaarte en toenemende complexiteit lijkt hbo meer voor de hand te liggen. Leidsters geven zelf bijvoorbeeld ook aan dat ze niet voldoende zijn toegerust voor de opvang van het groeiende aantal kinderen met leer- en gedragsproblemen.

Een derde punt betreft *de programma's* die worden gebruikt. Die zijn er vele tientallen en in uiteenlopende varianten, met als hoofdindeling programmagestuurd, dan wel ontwikkelingsgericht (ofwel meer vanuit het programma of de leerstof, resp. vanuit het kind). Tot recent werd ervan uitgegaan dat een programma 'bewezen effectief' zou moeten zijn, momenteel is die eis losgelaten en dient het te gaan om een programma dat de brede ontwikkeling stimuleert. Dit is een aanzienlijk soepelere eis, die veel meer ruimte laat. Deze versoepeling is niet in lijn met het advies van de Commissie Parlementair Onderzoek Onderwijsvernieuwingen (2008) die adviseerde dat beleidsmaatregelen, interventies en programma's pas grootschalig ingevoerd zouden mogen worden, wanneer de werking daarvan via onderzoek empirisch is aangetoond, c.q. bewezen effectief is (*evidence-based*). Er wordt in VVE-verband vaak gesproken over 'erkende' programma's, dat wil zeggen programma's die voorkomen in de databank Effectieve Jeugdinterventies van het Nederlands Jeugdinstituut (NJI, 2016a). Van alle VVE-interventies zijn er inmiddels zeven erkend, maar op één uitzondering na komt er geen verder dan het allerlaagste erkenningsniveau 'goed onderbouwd'. Eén programma krijgt het een-na-laagste oordeel 'effectief volgens eerste aanwijzingen', maar die aanwijzingen gaan alle kanten op (positief – negatief; zwak – sterk; verschillend per domein). Veel programma's zijn überhaupt nog niet op hun effectiviteit onderzocht en voor drie programma's zijn alleen een paar kleine lokale studies verricht.

### 3. De effectiviteit van VVE

#### 3.1 Effecten

Hiervoor is aangegeven dat er geen enkel Nederlands VVE-programma bestaat dat bewezen effectief is. Maar waar komt dan toch dat geloof in de positieve werking van VVE vandaan? Hierna worden de resultaten van onderzoek dat vooral is verricht in de VS, de bakermat van VVE, kort beschreven. Er zijn de afgelopen decennia, sinds begin jaren zestig, vele honderden studies verricht naar de effecten van dergelijke interventieprogramma's. De resultaten van al die studies zijn op hun beurt weer samengevat in een groot aantal literatuuroverzichten, reviewstudies en (statistische) meta-analyses (bv. Andrews & Slate, 2001; Barnett, 2001, 2008; Baumüller, 2013; Burger, 2010; Camilli, Vargas, Ryan & Barnett, 2010; Chambers, Cheung & Slavin, 2015; Diamond, Justice, Siegler & Snyder 2013; Karoly, Kilburn & Cannon, 2005; Lowenstein, 2011; Marope & Kagae, 2015; Melhuish, Eezy-Stevens, Petrogiannis, Ariescu, Penederi, Rentzou, Tawell, Leseman & Broekhuisen, 2015; Slot, 2014; Yoshikawa, Weiland, Brooks-Gunn, Burchinal, Espinosa, Gorley, Ludwig, Magnuson, Philips & Zaslow, 2013). De conclusie die doorgaans wordt getrokken luidt dat programma's *van hoge kwaliteit* een positief effect *kunnen* hebben op de schoolloopbaan van kinderen, en vervolgens ook op hun arbeidsmarktpositie en maatschappelijk functioneren.<sup>2</sup> In het algemeen is het zo dat het effect op cognitieve aspecten van de schoolloopbaan sterker is dan dat op niet-cognitieve aspecten, maar dat beide na verloop van tijd in sterkte afzwakken of helemaal verdwijnen (*fade-out*). Om een idee te geven van de sterkte van dergelijke effecten zijn de resultaten van de meta-analyse van Camilli et al. (2010) van een totaal van 123 afzonderlijke studies wellicht illustratief. Voor het cognitieve domein bedroeg de effectgrootte 0.23, voor de schoolse domein 0.14 en voor het sociale domein 0.16. Bij de interpretatie van dergelijke effecten (*effect sizes*) wordt vaak de vuistregel van Cohen (1988; zie ook NJi, 2016b) aangehouden: <0.20 verwaarloosbaar; 0.20 – 0.50 klein; 0.50 – 0.80 middelgroot; >0.80 groot. Het gaat zo bezien dus gemiddeld genomen om verwaarloosbaar kleine effecten.

Bij deze bevinding is het echter belangrijk een onderscheid te maken naar type onderzoek; grofweg zijn dat er twee. Allereerst studies die zijn uitgevoerd volgens de 'gouden standaard', waarbij kinderen *at random* zijn toegewezen aan een experimentele dan wel controlegroep en waarbij verschillen in achtergronden die het effect (mede) zouden kunnen bepalen zo op voorhand worden geëlimineerd. Dit betreft meestal (zeer) kleinschalige experimenten en hun aantal is ook beperkt. Verreweg de meeste studies zijn echter verricht volgens een quasi-experimenteel model, waarbij geprobeerd wordt de invloed van eventueel verstorende factoren (*selection bias*) zo goed mogelijk te elimineren via statistische controle achteraf. Het betreft in het algemeen studies naar de effecten van interventies die op grote schaal worden toegepast en waarbij veel kinderen zijn betrokken.

#### 3.2 Experimenten

In de discussie rond voor- en vroegschoolse educatie zijn het vaak de resultaten van de in de VS uitgevoerde experimenten die worden opgevoerd als toonbeeld van succesvolle interventies (bv. door Nobel-laureaat James Heckman). Deze projecten, van zeer hoge kwaliteit, zijn in de zestiger en zeventiger jaren uitgevoerd en de deelnemers zijn nog lang daarna gevolgd. Vanwege de aandacht die deze projecten in de discussie rond VVE steeds krijgen wordt er hier dieper op ingegaan.

Een eerste voorbeeld van zo'n modelprogramma betreft het Abecedarian Project (Campbell, Ramey, Pungello, Sparling & Miller-Johnson, 2002; Coalition for Evidence-Based Policy, 2016a). Aan dit programma namen 57 kinderen deel, waarbij hun ontwikkeling werd vergeleken met die van 54 kinderen in een controlegroep. (Bij de toewijzing was er overigens een substantiële afwijking van de nagestreefde *randomness* opgetreden.) De typische moeder was 20 jaar, had 10 jaar onderwijs gevolgd, was ongehuwd, leefde in een meer-generatie huishouden, en had geen inkomen. Alle moeders leefden onder de armoedegrens en waren (op 1 na) *African American*. Het project startte in 1972 en kinderen (en hun moeders) namen vanaf de geboorte gedurende 8 jaar deel, de hele dag, 5 dagen in de week en 52 weken per jaar. Het aanbod was kindgericht en besloeg de domeinen kennis, taal en gedrag. Tot 3 jaar was er 1 leidsters per 3 kinderen, daarna 1 per 6 kinderen. Naast activiteiten op het kinderdagverblijf en school, waren er in de schoolfase ook thuisactiviteiten. Elke week bezorgde een speciale leerkracht een pakketje met geïndividualiseerde oefeningen die de ouders elke dag gedurende ten minste een kwartier met het kind moesten doen. Het project kostte \$18000 per kind per jaar (prijspeil 2013).

Een tweede vaak aanhaalde project is het Perry Preschool project, hetgeen in 1962 startte met 58 3- en 4-jarige kinderen die opgroeiden in een zware achterstandssituatie (Barnett, 2001; Campbell et al., 2002; Coalition for Evidence-Based Policy, 2016b; Heckman Pinto & Savelyev, 2013). Zij gingen 2 jaar lang 5 halve dagen per week naar de pre-school; een controlegroep van 65 kinderen deed dat niet. De doelgroep bestond uit kinderen van *African American* ouders die in armoede opgroeiden. Bijna de helft van de kinderen leefde in een eenoudergezin. Een bijzonder en uitzonderlijk kenmerk is dat er ook op intelligentie is geselecteerd: alleen kinderen met een IQ tussen de 70 en 85 punten werden toegelaten. De ervaringen door het kind opgedaan in de groep werden via wekelijkse huisbezoeken van 1.5 uur gelinkt aan de leerervaringen thuis. De leidsters waren allen zeer hoog opgeleid en hadden een diploma om les te mogen geven op een voorschool, een basisschool, en ook in het speciaal onderwijs. Gedurende het project was er 1 leidster op elke 5 à 6 kinderen beschikbaar. Het curriculum ging uit van het principe van actief participerend leren, waarbij kinderen en volwassenen als gelijkwaardige partners in het leerproces werden behandeld. Het project kostte \$11300 per kind per jaar (prijspeil 2007). (Dit programma werd later in Nederland bewerkt tot het VVE-programma Kaleidoscoop.)

Beide experimenten lieten op een aantal dimensies positieve effecten gezien, voor sommige dimensies niet alleen op de korte, maar ook op de (zeer) lange termijn. Omdat deze effecten werden vastgesteld via een gerandomiseerd experiment, werden ze het visitekaartje voor voor- en vroegschoolse educatieve programma's. Volgens verschillende onderzoekers is dit echter niet terecht. Lowenstein (2011) bijvoorbeeld, waarschuwt voor het extrapoleren, c.q. generaliseren van de positieve effecten van modelprogramma's (zie ook Bauchmüller, 2013; Hanushek & Lindseth, 2009). Op de eerste plaats in verband met demografische ontwikkelingen en veranderingen in het aanbod van voor- en vroegschoolse voorzieningen (bv. de toegenomen arbeidsparticipatie van moeders en het grotere en gevarieerdere aanbod aan voorzieningen). Op de tweede plaats door de projecten zelf (kleine, homogene steekproeven van *African Americans* levend onder de armoedegrens; slechts per project één locatie; bovendien bleek uit recente heranalyses dat de positieve effecten alleen voor de meisjes golden). Op de derde plaats in verband met de unieke kwaliteit van de projecten die volstrekt onvergelijkbaar is met die van voorzieningen die normaliter door kinderen en ouders worden gebruikt (zeer hoge kwaliteit, zeer kostbaar; bovendien was er in ieder geval van het Perry Project een zeer sterke betrokkenheid van de onderzoeker bij de uitvoering).

Slavin en Smith (2009) tonen aan dat in dit type experimenten met kleine steekproeven de effecten dramatisch worden opgeblazen; bij een steekproef van minder dan 50 kinderen bedraagt het gemiddelde effect 0.44, bij een steekproef van meer dan 2000 kinderen niet meer dan 0.09. Volgens Burger (2010) is rechtstreekse vergelijking van Amerikaanse met Europese interventies niet alleen problematisch vanwege kwaliteitsverschillen in de uitvoering, maar ook omdat de Amerikaanse kinderen zich doorgaans in een veel zwaardere achterstandspositie bevinden dan hun Europese (en zeker Nederlandse) leeftijdsgenoten. Voor het Perry Preschool project komt daar nog bij dat de doelgroep behalve op gezinskenmerken ook nog geselecteerd is op zeer lage intelligentie. In Nederland zouden de betreffende kinderen naar het speciaal onderwijs zijn gegaan. De leidsters in de VS hadden daarom ook een onderwijsbevoegdheid voor het speciaal onderwijs. Onderzoekers van de Inspectie van het Onderwijs (Van de Kuilen en Van Dongen; Wikipedia, 2016) vermoeden bovendien dat de gevonden effecten geen direct resultaat zijn van het gevolgde programma, maar een indirect gevolg van het feit dat, doordat de kinderen hele of halve dagen werden opgevangen, de zeer laaggeschoolde moeders een opleiding konden afmaken en daardoor niet alleen een hoger ontwikkelingsniveau kregen, maar ook betere kansen hadden op de arbeidsmarkt. Heckman, Pinto en Savelyev (2013) stellen tegenover deze kritiek dat het bij zeer kleine steekproeven toch mogelijk is met geavanceerde statistische technieken op correcte wijze de mate van significantie te schatten. Ook volgens Schweinhart, Montie, Xiang, Barnett, Belfield en Nores (2005) valt het met de kritiek wel mee. Hoe het ook zij, tegenover het voordeel van de sterke interne validiteit van de experimenten staat hun zwakke externe validiteit. En volgens Yoshikawa, Weiland, Brooks-Gunn, Burchinal, Espinosa, Gorley, Ludwig, Magnuson, Philips en Zaslow (2013) zijn er ook meerdere alternatieven om in quasi-experimentele designs toch redelijk adequaat te kunnen controleren voor selectie-effecten. Met andere woorden: selectiebias hoeft volgens hen niet zo'n probleem te zijn als vaak wordt voorgesteld. De kritiek op quasi-experimentele studies is dan ook niet per se terecht. Fukking et al. (2015) zijn dan ook van mening dat beide onderzoekstypen voor- en nadelen kennen en elkaar daarmee goed kunnen aanvullen.

### **3.3 Kwaliteit**

In veel studies naar de effectiviteit van VVE worden ook pogingen gedaan te achterhalen waardoor die effectiviteit (in het geval ze wordt aangetoond) bepaald wordt. Als centrale voorwaarde voor succes wordt vooral gewezen op de kwaliteit van de interventie. Wat daaronder precies moet worden verstaan varieert nogal per bron en land. Wat dit laatste valt uit het overzicht van UNICEF (2008) af te leiden dat de voorzieningen in Europa aan veel meer kwaliteitsstandaarden voldoen dan die in de VS. In hun overzicht presenteren Andrews en Slate (2001) een keur aan structurele en proceskenmerken als criteria voor hoge kwaliteit: kleine groepen, lage leidster/kind-ratio's, ouderbetrokkenheid, een geïntegreerd aanbod (niet alleen op het gebied van onderwijs en opvoeding, maar ook op dat van gezondheid, voeding en maatschappelijke ondersteuning), een aanbod afgestemd op de ontwikkeling van het (individuele) kind, regelmatige toetsing en monitoring, hooggekwalificeerde leidsters, vroeg beginnen en ook doorgaan (binnen het gezin, buurt en school) nadat de interventie zelf is gestopt, een hoge intensiteit (qua aantal dagen en aantal uren per dag), het opdoen van dagelijkse leerervaringen, een gevarieerd aanbod gedurende de dag, en afwisselende groepssamenstellingen. Er bestaat echter lang niet altijd overeenstemming over deze kwaliteitsindicatoren en het ontbreekt ook vaak aan voldoende empirische ondersteuning. Zo rapporteren Camilli et al. (2010) positieve effecten van directe instructie, dat wil zeggen van leraar-geïnitieerde schoolse activiteiten, in tegenstelling tot de frequent gepropageerde door de kinderen zelf geïnitieerde activiteiten. Ook vonden zij een

sterk negatief effect van een geïntegreerd aanbod. Barnett (2001) wijst er in zijn overzicht op dat het inzetten van ouders, om zo indirect de ontwikkeling van kinderen te stimuleren, nauwelijks of geen effect heeft, en bovendien zeer kostbaar is. Volgens hem laten daarnaast programma's met een geïntegreerd aanbod eveneens teleurstellende resultaten zien. Melhuish et al. (2015) daarentegen zijn van mening dat juist de combinatie van kinderopvang van hoge kwaliteit en huisbezoeken effectief is. Burger (2010) betrok startleeftijd, intensiteit en duur in zijn analyses, maar vond nauwelijks bewijs voor de invloed van deze programmamakenmerken. Lowenstein (2011) rapporteert zelfs over onderzoek waaruit blijkt dat naarmate kinderen meer tijd weg zijn bij hun moeder ze meer probleemgedrag vertonen. In een recentelijk, grootschalig Nederlands onderzoek (Karssen, Van der Veen, Veen, Van Daalen & Roeleveld, 2013) is voor een reeks van structurele randvoorwaarden (zoals aantal leidsters en opleiding leidsters) en de kwaliteit van het aanbod in kinderdagverblijven en peuterspeelzalen nagegaan of deze van invloed zijn op een aantal cognitieve en sociaal-emotionele uitkomstmaten bij instroom in de basisschool. Dat bleek niet het geval te zijn. Datzelfde gold voor de kwaliteit van het aanbod in de kleutergroepen op de uitkomstmaten in groep 2. Het voorgaande samenvattend: tegenover uitgebreide overzichten met bewijs voor kwaliteitsindicatoren staan volop studies die dat bewijs weer ontkrachten.

#### **4. Recente Nederlandse publicaties**

Zeer recentelijk zijn in een kort tijdsbestek drie belangwekkende, Nederlandse publicaties verschenen over VVE, één eind 2015 en twee in het voorjaar van 2016. Deze hebben veel media-, beleids- en wetenschappelijke aandacht getrokken, met name omdat ze ingaan op de kern van VVE, te weten de effectiviteit ervan en de mogelijke bijdrage aan het voorkomen en verminderen van onderwijsachterstanden. Vanwege hun impact zijn deze publicaties kritisch tegen het licht gehouden; hierna volgt een beknopte bespreking.

##### **4.1 Een meta-analyse van Nederlands studies**

In Nederland is de afgelopen decennia een beperkt aantal studies verricht naar effecten van VVE op de ontwikkeling van kinderen, soms kleinschalig, soms grootschalig. Fukkink, Jilink en Oostdam (2015) zijn via een statistische meta-analyse nagegaan wat dergelijke studies bij elkaar genomen hebben opgeleverd. Ze concentreerden zich daarbij op grootschalige retrospectieve en kleinschalige quasi-experimentele studies die zijn verricht in de periode 2000 – 2015, dat wil zeggen sinds de inwerkingtreding van het landelijke VVE-beleid. Hun inventarisatie leverde 11 afzonderlijke studies<sup>3</sup> met 21 verschillende deelstudies op, waarin in totaal 165 uitkomsten werden gerapporteerd gelegen in de vier inhoudelijke domeinen taal, rekenen, algemene intelligentie en sociaal-emotioneel. Deze uitkomsten zijn omgezet in de gestandaardiseerde effectgrootte  $d$  (Cohen, 1988; zie ook hierboven). De analyses laten zien dat het geaggregeerde effect voor geen van de vier onderscheiden domeinen significant afwijkt van nul; het is 0.03, ofwel – in de woorden van de onderzoekers – ‘kleiner dan klein’. Dit resultaat heeft alles bij elkaar betrekking op ruim 50000 kinderen en meer dan 60 miljoen uren VVE. Analyse laat ook zien dat er sprake is van een licht dalende trend in de tijd, hetgeen in tegenspraak is met de verwachting dat ten gevolge van de sinds 2000 doorgevoerde beleidsaanpassingen en investeringen het niveau gestegen zou zijn. De bevindingen wijken ook sterk af van positieve *success stories* uit de Verenigde Staten. Mogelijke oorzaken voor deze discrepantie liggen volgens Fukkink c.s. in verschillen in met name tijd, doelpopulatie en aanpak. Zo zijn bijvoorbeeld de doelgroepen van zeker de eerste programma's in de VS qua sociaal herkomstmilieu onvergelykbaar met de Nederlandse doelgroep en valt er daardoor in de VS ook



veel meer winst te behalen. Bovendien zijn die programma's daar ook veel intensiever in tijd en is er in Nederland een alternatief, regulier aanbod in kinderdagverblijven, peuterspeelzalen en kleutergroepen van de basisschool, dat in termen van pedagogische kwaliteit nauwelijks verschilt van VVE. Daarnaast is er in de VS vaak ook sprake van een aanbod in een instelling in combinatie met een sterke oudercomponent thuis.

## 4.2 Pre-COOL

Onlangs kopte NRC (Vasterman, 2016) met 'Nu blijkt ineens: taalles voor jonge kinderen werkt dus wél', en vervolgens: 'Eindelijk is er duidelijkheid: voor- en vroegschoolse educatie (VVE) heeft zin'. Dit zou blijken uit de Pre-COOL cohortstudie die is gericht op het verkrijgen van zicht op effecten van opvang en educatie op de ontwikkeling van alle kinderen, niet alleen van VVE in het kader van achterstandsbestrijding. Het onderzoek is opgezet om aan de methodologische kritiek op de quasi-experimentele studies tegemoet te komen; het slaat via de opname van een vroege beginmeting in het design een brug tussen een streng-experimentele opzet en een quasi-experimentele opzet met onvoldoende controle voor mogelijke selectie-effecten. Volgens het interview in NRC zou de achterstand van VVE-kinderen worden gehalveerd en daarmee zijn aangetoond dat VVE zin heeft. Er zou echter meer winst worden behaald wanneer kinderen eerder naar de opvang of peuterspeelzaal zouden gaan en het aanbod daar zou worden geïntensiveerd, dus veel meer uren per week. Daarbij zou alles staan of vallen met de professionaliteit van de leidsters, zo luidde het eindoordeel. Alle aanleiding om het onderliggende rapport te bestuderen (Leseman & Veen, 2016).

In de inleiding van dit artikel is aangegeven dat er in de praktijk veel variatie bestaat in de doelgroepdefinitie van gemeenten. Dat kan een probleem zijn bij een grootschalige evaluatie over gemeenten heen. In Pre-COOL is dit probleem omzeild door drie mogelijke doelgroepen te definiëren, op basis van opleiding (moeder maximaal lbo vs. hoger), etniciteit (een van de ouders in niet-Westers land geboren vs. elders) en thuistaal (een buitenlandse taal en Fries vs. Nederlands en Nederlandse streektaalen en dialecten). Vervolgens is gekeken hoe deze kinderen zich op enkele gebieden hebben ontwikkeld tussen – in principe – meetmoment 1 (op 2-jarige leeftijd) en meetmoment 4 (op 6-jarige leeftijd). Of een kind ook écht tot de VVE-doelgroep behoort (in de betreffende gemeente) is niet meegenomen. Evenmin of het kind ook feitelijk VVE heeft gevolgd. Er is alleen gekeken of het kind in de *voorschoolse* fase op een VVE-instelling heeft gezeten die VVE aanbod (vergelijkbare informatie over de *vroegschoolse* fase is niet meegenomen). Probleem is dat veel VVE-instellingen gemengde groepen kennen, met zowel doelgroep- als niet-doelgroep-kinderen; wie van hen VVE heeft gevolgd blijft zo onduidelijk. Bovendien wisselen kinderen regelmatig van instelling. Evenmin is meegenomen hoe lang en met welke intensiteit het kind VVE heeft gehad. Sommige kinderen volgen het gedurende een paar maanden enkele ochtenden per week, anderen anderhalf jaar elke ochtend. Vaak bezoeken niet-doelgroepkinderen twee dagdelen de instelling en doelgroepkinderen vier dagdelen.<sup>4</sup> In het geval een kind VVE heeft gevolgd is ook niet meegenomen met welk programma of welke programma's is gewerkt. Kortom, omdat er in de analyses geen een-op-een relatie bestaat tussen feitelijke doelgroep en feitelijk aanbod loopt alles door elkaar en is het niet mogelijk eenduidige conclusies te trekken over de effectiviteit van VVE.

Als wordt afgezien van al deze onoverkomelijke voorbehouden, wat zijn dan de uitkomsten? Bij de uitkomstmaat 'selectieve aandacht' werd voor de doelgroep in termen van opleiding een effect van 0.42 gevonden (ofwel klein tot middelgroot), voor doelgroep in termen van etniciteit een

effect van 0.38 en voor doelgroep in termen van thuistaal van 0.24 (niet-significant). Deze positieve effecten willen zeggen dat de doelgroep en niet-doelgroep in de loop der jaren naar elkaar toegroeien, of anders gezegd: het verschil tussen beide groepen wordt kleiner. Er bleek geen significant effect van VVE op te treden. Voor de uitkomstmaat 'woordenschat' was het effect voor opleiding 0.16 (niet-significant), voor etniciteit 0.46 en voor thuistaal 0.32 (ofwel klein). Hier werd alleen een significant VVE-effect gevonden voor etniciteit (en niet opleiding en thuistaal), dat was 1.17 (ofwel zeer sterk).<sup>6</sup> Voor de uitkomstmaat 'speel-werkhouding' was het effect voor opleiding 0.00, voor etniciteit -0.44 en voor thuistaal -0.37. Deze beide negatieve effecten houden in dat het verschil tussen doelgroep en niet-doelgroep groter wordt; de niet-doelgroep gaat het vergeleken met de doelgroep steeds beter doen. Ook hier bleek geen significant effect van VVE op te treden. Voor de uitkomstmaat 'ontluikende rekenvaardigheid' was het effect voor alle drie de soorten doelgroepen 0.00, met andere woorden het aanvangsverschil bleef over alle metingen heen even groot. Ook hier geen effect van VVE. Voor de uitkomstmaat 'taalvaardigheid' was het effect voor opleiding 0.25, voor etniciteit 0.37 en voor thuistaal 0.52. En ook hier geen effect van VVE. Samenvattend: voor vijf uitkomstmaten is nagegaan of er verschillen bestaan in de ontwikkeling van kinderen in relatie tot drie gezinskenmerken ('doelgroepen'). De effectgroottes varieerden van -0.44 tot +0.46, met andere woorden soms liepen de doelgroepkinderen in op de niet-doelgroepkinderen, en soms werd het verschil juist groter. Opvallend was dat voor speel-werkhouding de achterstand van de doelgroep groter werd; voor rekenvaardigheid bleef het aanvangsverschil bestaan. In slechts één geval bleek er een effect van VVE op te treden, namelijk bij woordenschat voor de doelgroep op basis van etniciteit.

Kortom, Pre-COOL is een uniek onderzoek qua opzet en data, en kan ongetwijfeld zeer waardevol zijn als het gaat om een goed beeld te krijgen van voor- en vroegschoolse voorzieningen en de effecten daarvan op de ontwikkeling van jonge kinderen, maar de suggestie zoals ze in NRC wordt gewekt dat nu is aangetoond dat VVE werkt is een zeer voorbarige aangezien ze in het geheel niet wordt ondersteund door de bevindingen. Op een totaal van 15 analyses werd slechts in één geval een positief effect gerapporteerd, dus ongeveer op kansniveau.

### **4.3 Kansrijk onderwijs**

Onlangs heeft het CPB (2016) de studie Kansrijk onderwijsbeleid gepubliceerd. Op basis van de literatuur is nagegaan welke beleidsmaatregelen op het gebied van onderwijs wél en welke níet effectief zijn. Bij de keuze van de literatuur is het CPB zeer selectief geweest, al neemt het daarbij naar eigen zeggen qua strengheid een middenpositie in. Alleen studies in de vorm van gerandomiseerde experimenten of quasi-experimentele studies met controle voor zelfselectie werden goed genoeg bevonden. Het merendeel daarvan was onderwijs-economisch van aard. Het resultaat is ronduit ontluisterend. Wereldwijd, maar zeker voor Nederland, is er volgens het CPB slechts een zeer beperkt aantal 'goede' studies verricht. De basis voor de beleidsmaatregelen op het gebied van onderwijs is dus nog veel smaller dan vaak werd verondersteld. Dat geldt zeker voor het beleid rond VVE. Volgens het CPB voldoet geen enkel Nederlands effectonderzoek naar VVE aan haar criteria.<sup>5</sup> Maar ook buiten Nederland valt het aantal goede studies op een paar handen te tellen. Het CPB gaat op basis van die studies voor negen mogelijke VVE-maatregelen na hoeveel leerwinst die zouden opleveren. Omdat het (noodgedwongen) om buitenlands onderzoek gaat en de omstandigheden daardoor veelal onvergelykbaar zijn, is een correctie toegepast. Als referentiepunt opteerde het CPB daarbij voor de toetsscores taal en rekenen uit het internationale PISA-onderzoek.<sup>7</sup> Voor de kwantitatieve onderbouwing van de negen maatregelen

is gebruik gemaakt van zes unieke studies. Volgens het CPB zijn er nog enkele studies die niet zijn gebruikt maar die wel tot vergelijkbare resultaten zouden leiden; de data voldeden echter niet helemaal of er was sprake van doublures. Wat bleek nu? Het effect van één maatregel kon niet worden onderzocht en twee maatregelen leverden geen leerwinst op. Voor de overige zes maatregelen wordt, gebaseerd op in totaal drie studies, een geringe leerwinst gerapporteerd. De meeste winst verwacht het CPB wanneer de intensiteit van (de huidige modale) 10 uur voorschools VVE per week wordt opgeschroefd naar 24 uur, dus *full-time* voor een peuter. Dat zou jaarlijks een extra investering vergen van circa €100 per kind, bovenop de ruim €600 die het nu kost; in totaal zo'n €10 miljoen.

Het CPB baseert zich bij dit alles voornamelijk op het in 1965 in Amerika gestarte Head Start. Dit betreft een mengelmoes aan lokale programma's voor de voorschoolse periode, met een sterke oudercomponent en bedoeld voor gezinnen levend onder de armoedegrens. Head Start is in de meeste opzichten *niet* vergelijkbaar met het Nederlandse VVE, niet alleen niet qua financiering en budget, maar ook niet qua doelgroepen, organisatie, duur en intensiteit, vormgeving en inhoud (vgl. USC, 2007). Vijf CPB-maatregelen op het gebied van VVE steunen echter vrijwel volledig op één evaluatie van Head Start, waarbij gebruik is gemaakt van een gerandomiseerde controlegroep. Als echter de betreffende rapportage wordt geraadpleegd (Puma, Bell, Cook, Heid, Broene, Jenkins, Mashburn & Downer, 2012), dan blijken de resultaten allesbehalve eenduidig en hard te zijn. In het rapport wordt een enorm groot aantal effecten gerapporteerd, gelegen op het cognitieve en sociaal-emotionele vlak, maar ook met betrekking tot gezondheid, opvoeding en onderwijs; in totaal betreft het liefst 414 effecten. Anders dan gangbaar wordt bij het toetsen van deze effecten niet uitgegaan van het normaal voor deze situatie geldende significantieniveau van  $p < 0.01$ , maar van het uitzonderlijk tolerante niveau van  $p \leq 0.10$ , wat de onderzoekers kwalificeren als *suggestive evidence*. Van deze 414 effecten zijn er slechts 60 significant op dit *relaxed* significantieniveau, waarvan er bovendien 4 niet te interpreteren zijn en 8 ook nog eens negatief uitpakken. Blijven er 48 positieve effecten over (ofwel 12.5%). Hetgeen in lijn met het gehanteerde  $p$ -niveau maar net boven kansniveau uitstijgt. Het maximale gestandaardiseerde, positieve effect bedraagt 0.35, maar er zijn ook negatieve effecten van maximaal -0.24. Puma et al. (2016: xvii) vatten de resultaten als volgt samen '...there were initial positive impacts from having access to Head Start, but by the end of 3rd grade there were very few impacts found ... in any of the four domains of cognitive, social-emotional, health and parenting services. The few impacts that were found did not show a clear pattern of favorable or unfavorable impacts for children.' Met andere woorden: er lijkt bijzonder weinig empirische basis voor het CPB-advies.

## 5. Conclusies

VVE is erop gericht peuters en kleuters die thuis onvoldoende schoolse stimulering ontvangen daarvoor in het kinderdagverblijf, op de peuterspeelzaal en de basisschool te compenseren via een gericht educatief programma. Het doel is te voorkomen dat jonge kinderen bij hun start met het 'echte' onderwijs in de basisschool al een achterstand hebben. Het huidige VVE bestaat inmiddels ruim 15 jaar en er naar schatting ruim 4 miljard euro in omgegaan. Enige tijd geleden stelde de Commissie Parlementair Onderzoek Onderwijsvernieuwingen (2008) dat beleidsmaatregelen, instrumenten, aanpakken, interventies en programma's pas grootschalig

ingevoerd zouden mogen worden, wanneer de werking daarvan via onderzoek empirisch was aangetoond (*evidence based*). Hoe staat het dan met de empirische grondslag van VVE?

In wetenschappelijk onderzoek kunnen in methodologisch opzicht ‘strengen’ en ‘rekkelijken’ worden onderscheiden. Zo bezien, zou van de hierboven besproken studies die van Fukking et al. (2015) onder de rekkelijken geschaard kunnen worden, die van Leseman en Veen (2016) zou een tussenpositie innemen, en die van het CPB (2016) zou tot de strengen gerekend kunnen worden. Opmerkelijk is dat het toonbeeld van zelfverklaarde strengheid, het CPB, voor de onderbouwing van haar adviezen uitgaat van een uiterst arbitrair tolerant significantieniveau, zich eenzijdig baseert op een aandeel aan positieve effecten dat, bij het voorgaande punt aansluitend, amper boven raadkans uitkomt, en waarvan de sterkte van de effecten zelf twijfelachtig is. Objectief gezien zou de conclusie van het CPB eerder moeten zijn dat er geen effecten kunnen worden aangetoond. En dat sluit dat aan bij de resultaten van de analyse van Fukking et al. (2015). Want wat methodologische strengheid betreft ontlopen beide studies zich uiteindelijk helemaal niet zo. Wat Pre-COOL betreft is duidelijk geworden dat de suggestie van het NRC-interview dat nu ineens zou blijken dat VVE wél werkt een zeer voorbarige is en in het geheel niet ondersteund wordt vanuit de gerapporteerde analyses.

Bij dit alles dient bovendien nog het volgende in het achterhoofd te worden gehouden. Een eerste punt is dat ten gevolge van de eenzijdige focus op experimentele studies en quasi-experimentele studies met strenge controle voor selectiebias er in meta-analyses uiteindelijk vaak maar een zeer beperkt aantal studies overblijft voor inclusie. Het aantal resterende studies op basis waarvan dan uitspraken worden gedaan is zeer gering. Een ander punt betreft de tendens studies die geen (of zelfs negatieve) effecten laten zien niet te publiceren (het *file-drawer* probleem), waardoor het totaalbeeld gunstiger wordt voorgesteld dan het in werkelijkheid is (bv. Aleman & Van Tuijl, 2000; Burger, 2010). In een recente, omvangrijke analyse van onderwijskundig onderzoek kwantificeren Cheung en Slavin (2016) deze beide punten. Zij laten allereerst zien dat het gemiddelde effect in quasi-experimenteel onderzoek 44 procent hoger ligt dan in gerandomiseerd-experimenteel onderzoek. Daarenboven volgt uit hun analyses dat het effect in gepubliceerd onderzoek 88 procent hoger uitkomt dan dat in ongepubliceerd onderzoek. Alhoewel dat eigenlijk niet mogelijk is, zou dit impliceren dat de door Fukking et al. (2015) gerapporteerde effecten nog een overschatting inhouden.

Verder valt op dat het in het merendeel van de studies de gewoonte is louter te focussen positieve effecten; meestal rapporteert men slechts gemiddelden, zonder de range te geven of het aandeel nul- en negatieve effecten te noemen. Er lijkt daarmee evident sprake van eenzijdige gerichtheid op het vinden en rapporteren van gunstige resultaten. Het eerder genoemde *file-drawer* probleem doet zich dus niet alleen voor over de studies heen, maar evenzeer binnen de (uiteindelijk wel) geselecteerde studies. Dit selectief *shoppen* heeft als bijkomend nadeel dat er weliswaar regelmatig over enkele positieve effecten wordt gerapporteerd, maar dat het op basis van deze gefragmenteerde informatie nauwelijks mogelijk is een effectieve interventie te ontwikkelen bestaande uit samenhangende componenten en gericht op verschillende niveaus (vgl. Camilli et al., 2010). Fukking et al. (2015) stellen vast dat uit de discussieparagraaf van enkele VVE-rapporten een sterke *bias* spreekt richting het rapporteren van (louter) positieve effecten. De empirisch vastgestelde, tegenvallende effecten lijken voor de boodschappers een *inconvenient truth*. En daarbij is er wellicht ook een link met hun opdrachtgevers, de ministeries en politieke partijen, die vanuit sociaal-politieke overwegingen, met het oog op de gedane investeringen en

het feit dat het lastig is op een eenmaal ingeslagen beleidstraject terug te komen, er alle baat bij hebben met positieve resultaten te komen. Dat dit speelt blijkt bijvoorbeeld uit het relaas van Eldering (2002), die in de negentiger jaren het gezinsprogramma Opstap, de Nederlandse bewerking van het Israëliëse HIPPY, evalueerde. Dat liet geen effecten zien, niet bij de moeders, noch bij de kinderen, maar zij werd door haar opdrachtgever, het ministerie van WVC, onder grote druk gezet met een positief verhaal te komen. Ruim voordat haar negatieve conclusies werden gepubliceerd, had tot haar verbijstering het ministerie al besloten het betreffende project op grote schaal in te voeren.

De afgelopen decennia is gebleken dat maatregelen die worden genomen op een later tijdstip in de schoolloopbaan van achterstandskinderen geen of maar een zeer beperkt effect sorteren. Onder het motto ‘voorkomen is beter dan genezen’ is sindsdien de beleidsaandacht vrijwel volledig verschoven naar de voor- en vroegschoolse fase en is er vol ingezet op stimuleringsprogramma’s voor peuters en kleuters. Het doel is niet zozeer achterstanden te bestrijden, dan wel ze te voorkomen. De empirische onderbouwing van dit beleid, de *evidence base*, is echter vooralsnog bijzonder mager en heeft nog een lange weg te gaan.

## Noten

- 1 Illustratief is het onderzoek van Keegstra (2010) naar het effect van VVE-taalstimuleringsprogramma’s. Bij de start in groep 1 bleek dat van 44% van de allochtone en 5% van de autochtone kinderen het taalbegrip onvoldoende was. Na een jaar VVE had nog slechts 5% van de allochtone kinderen een onvoldoende. Bij de meeste autochtone kinderen was de taalscore echter stabiel gebleven, ook van de 5% met een onvoldoende bij de eerste meting. Van de totale groep autochtonen was er zelfs 12% op achteruit gegaan. Keegstra geeft als mogelijke verklaring dat het bij de allochtone kinderen om een (te remediëren) blootstellingsachterstand gaat (dus het achterstandsperspectief) en bij de 5% allochtone en autochtone kinderen met een (blijvende) onvoldoende qua taalbegrip om een specifieke taalstoornis, bepaald door een aangeboren zwak taalleervermogen (dus het zorgperspectief). Voor deze laatste groep lijkt taalstimulering via een VVE-programma volgens haar dus niet de juiste oplossing.
- 2 Overigens is niet altijd even duidelijk wat nu precies het verschil is tussen het aanbod in reguliere voorzieningen en dat in voorzieningen met een specifiek, extra aanbod voor achterstandskinderen. Ook in reguliere voorzieningen ontvangen kinderen die dat nodig hebben vaak een extra (individueel) aanbod. Bovendien zitten achterstandskinderen vaak in gemengde groepen met niet-achterstandskinderen die meeprofiteren van het extra aanbod aan hun groepsgenootjes. Illustratief is het onderzoek van De Haan, Leseman en Elbers (2011) waaruit bleek dat VVE geen effect had. Een verklaring daarvoor was volgens de onderzoekers dat leidsters die werkten zónder VVE-programma er even goed of zelfs beter in slaagden om ontwikkelingsstimulerende activiteiten te initiëren en begeleiden als leidsters die wél met een VVE-programma werkten.
- 3 Hierbij ontbraken overigens enkele studies, alle met géén effect. Zoals de dissertatie van Bauchmüller (2013), die, gebruikmakend van grootschalige, longitudinale landelijke en regionale data, geen significante effecten kon aantonen. Uit de experimentele studie (*randomized controlled trial*) van Van Druten-Frietman, Gijssels, Denessen en Verhoeven (2014) bleek dat Taallijn, dat is ontwikkeld als versterking van bestaande VVE-programma’s,

geen toegevoegde waarde had. Er was geen effect van Taallijn op de groei van de woordenschat.

- 4 In het NRC-artikel wordt juist benadrukt dat startleeftijd en aanbodintensiteit zeer bepalend zijn. Maar dat is in deze studie niet onderzocht.
- 5 Hetgeen dus impliceert dat ook de meta-analyse van Fukkink et al. (2015) volgens het CPB geen wetenschappelijk verantwoorde empirische basis biedt. Overigens heeft het CPB enkele effectevaluaties over het hoofd gezien die wél voldoen aan haar criteria, maar geen positief effect aantoonde. Zoals de studies van Bauchmüller (2013) en Van Druten-Frietman, Gijssels, Denessen en Verhoeven (2014), zie noot 3, en de studie van Van Schooten en Slegers (2008) waarin op basis van onderzoek in enkele gemeenten werd geconcludeerd dat VVE daar niet werkt en waarbij ook werd gewezen op negatieve effecten van VVE.
- 6 Opmerkelijk is dat er ook een effect van VVE werd gevonden voor de niet-doelgroep, i.c. de Westerse kinderen. Dit bedroeg 0.30, maar was niet significant. Hoe een dergelijk effect geïnterpreteerd zou moeten worden is niet duidelijk. Waarschijnlijk heeft het te maken met het feit dat etniciteit (landelijk gezien) geen formeel doelgroepcriterium is en niet duidelijk is of de betreffende kinderen ook feitelijk wel VVE hebben gevolgd (dit laatste geldt overigens evenzeer voor de doelgroepkinderen, voor wie wél een VVE-effect wordt gerapporteerd).
- 7 Er bestaat overigens redelijk wat kritiek juist op de vergelijkbaarheid van deze scores van uiteenlopende landen (bv. Feniger & Lefstein, 2014; Goldstein, 2004). En ook gaat het bij PISA om 15-jarigen, en bij VVE om 2.5- – 6-jarigen, maar dat alles terzijde.

### **Geraadpleegde literatuur**

- Aleman, A., & Tuijl, C. van (2000). Meta-analyse, heterogeniteit, en de effecten van voorschoolse educatieve programma's: Een kritische beschouwing. *Pedagogiek*, 20(2), 162-166.
- Andrews, S., & Slate, J. (2001). Prekindergarten programs: A review of the literature. *Current Issues in Education*, 4(5). Op 08/11/10 gedownload van <http://cie.ed.asu.edu/~volume4/~number5/>
- Barnett, W. (2001). *School reform proposals: The research evidence*. Early childhood education. Temple, AZ: Arizona State University.
- Barnett, W. (2008). *Preschool education and its lasting effects: Research and policy implications*. New Brunswick, NJ: Rutgers, The State University of New Jersey.
- Bauchmüller, R. (2013). *Investing in early childhood care and education: The impact of quality on inequality*. Dissertatie Universiteit Maastricht.
- Burger, K. (2009). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25(2), 140-165.
- Camilli, G., Vargas, S., Ryan, S., & Barnett, W. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teacher College Record*, 112(3), 579-620.
- Campbell, F., Ramey, C., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6(1), 42-57.
- Chambers, B., Cheung, A., & Slavin, R. (2015). *Literacy and Language Outcomes of Balanced and Developmental-Constructivist Approaches to Early Childhood Education: A Systematic*

- Review. September 21, 2015. Op 06/08/16 gedownload van [http://www.bestevidence.org/word/early\\_child\\_ed\\_Sept\\_21\\_2015.pdf](http://www.bestevidence.org/word/early_child_ed_Sept_21_2015.pdf)
- Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283-292.
- Coalition for Evidence-Based Policy (2016a). *Abecedarian Project*. Op 05/08/16 geraadpleegd op <http://evidencebasedprograms.org/1366-2/abecedarian-project>
- Coalition for Evidence-Based Policy (2016b). *Perry Preschool Project*. Op 05/08/16 geraadpleegd op <http://evidencebasedprograms.org/1366-2/65-2>
- Commissie Parlementair Onderzoek Onderwijsvernieuwingen (2008). *Parlementair Onderzoek Onderwijsvernieuwingen*. 's-Gravenhage: Sdu Uitgevers.
- CPB (2016). *Kansrijk onderwijsbeleid*. Den Haag. CPB.
- Crane, J., & Barg, M. (2003). *Do early childhood intervention programs really work?* Coalition for Evidence-Based Policy. Op 05/08/16 gedownload van <http://evidencebasedprograms.org/wp-content/uploads/2012/12/Do-Early-Intervention-Programs-Really-Work7.pdf>
- Diamond, K., Justice, L., Siegler, R., & Snyder, P. (2013). *Synthesis of IES research on early intervention and early childhood education*. Washington, DC: USDOE.
- Driessen, G. (2012). *Variatie in Voor- en Vroegschoolse Educatie. Een onderzoek naar de uiteenlopende wijzen waarop in gemeenten vorm wordt gegeven aan VVE*. Nijmegen: ITS.
- Driessen, G. (2013). *De bestrijding van onderwijsachterstanden Een review van opbrengsten en effectieve aanpakken*. Nijmegen: ITS.
- Driessen, G. (2015). *VVE als panacée?* Presentatie CPB Policy seminar, Centraal Planbureau, Den Haag, 17 september 2015..
- Driessen, G. (2016). *VVE: Investeren of integreren?* Presentatie Thematische ministerstaf, Ministerie van Financiën, Den Haag, 16 februari 2016.
- Driessen, G., Veen, A., & Daalen, M. van (2015). *VVE-doelgroepkinderen in de voorschoolse fase. Indicering en aanbod*. Nijmegen: ITS.
- Druten-Frietman, L. van, Gijssel, M., Denessen, E., & Verhoeven, L. (2014). Het effect van 'Taallijn' op de woordenschatontwikkeling van peuters. *Pedagogische Studiën*, 91(6), 383-396.
- Eldering, L. (2002). *Enkele episodes uit het leven van Marokkaanse en Turkse gezinnen*. Leiden: Universiteit Leiden.
- Feniger, Y., & Lefstein, A. (2014). How not to reason with PISA data: An ironic investigation. *Journal of Education Policy*, 29(6), 845-855.
- Goldstein, H. (2004). International comparisons of student attainment: Some issues arising from the PISA study. *Assessment in Education: Principles, Policy, and Practice*, 11(3), 319-330.
- Fukkink, R., Jilink, L., & Oostdam, R. (2015). *Met een blik op de toekomst. Een meta-analyse van de effecten van VVE op de ontwikkeling van kinderen in Nederland*. Amsterdam: Hogeschool van Amsterdam.
- Haan, A., Leseman, P., & Elbers, E. (2011). *Pilot gemengde groepen 2007-2010*. Utrecht: Universiteit Utrecht.
- Hanushek, E., & Lindseth, A. (2009). *Schoolhouses, courthouses, and statehouses. Solving the funding-achievement puzzle of America's public schools*. Princeton, NJ: Princeton University Press.
- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *The American Economic Review*, 103(6), 2052-2086.

- Karoly, L. A., Kilburn, M., & Cannon, J. (2005). *Early childhood interventions: Proven results, future promises*. Op 04/06/15 gedownload van [http://www.rand.org/pubs/monographs/2005/RAND\\_MG341.pdf](http://www.rand.org/pubs/monographs/2005/RAND_MG341.pdf)
- Karssen, M., Veen, I. van der, Veen, A., Daalen, M. van, & Roeleveld, J. (2013). *Effecten van deelname aan en kwaliteit van voor- en vroegschoolse educatie op de ontwikkeling van kinderen*. Amsterdam: Kohnstamm Instituut.
- Keegstra, A. (2010). *Language problems in young children. General assumptions investigated*. Dissertatie Rijksuniversiteit Groningen.
- Leseman, P., & Veen, A. (Eds.) (2016). *Ontwikkeling van kinderen in relatie met kwaliteit van voorschoolse instellingen. Resultaten uit het Pre-Cool cohortonderzoek*. Amsterdam: Kohnstamm Instituut.
- Lowenstein, A., (2011). Early Care and Education as educational panacea: What do we really know about its effectiveness? *Educational Policy*, 25(1), 92-114.
- Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penederi, E., Rentzou, K., Tawell, A., Leseman, P., & Broekhuisen, M. (2015). *A review of research on the effects of Early Childhood Education and Care (ECEC) on child development*. Op 15/09/15 gedownload van [ecec-care.org](http://ecec-care.org)
- NJi (2016a). *Databank Effectieve Jeugdinterventies*. Op 20/08/16 geraadpleegd via <http://www.nji.nl/nl/Databank/Databank-Effectieve-Jeugdinterventies/Erkende-interventiesNJi>
- (2016b). *Effectgrootte*. Op 20/08/16 geraadpleegd via <http://www.nji.nl/Effectgrootte>
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study Final Report, OPRE Report # 2012-45*. Washington, DC: U.S. Department of Health and Human Services.
- Schweinhart, L., Montie, J., Xiang, Z., Barnett, W., Belfield, C., & Nores, M. (2005). *The High/Scope Perry Preschool Study through age 40. Summary, conclusions, and frequently asked questions*. Ypsilanti, MI: High/Scope Press.
- Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic review in education. *Educational Evaluation and Policy Analysis*, 31(4), 500-506.
- Schooten, E. van, & Slegers, P. (2008). *Onderzoek naar de effectiviteit van VVE en peuterspeelzalen in Oosterhout en Den Bosch*. Amsterdam: SCO-Kohnstamm Instituut.
- Slot, P. (2014). *Early childhood education and care in the Netherlands. Quality, curriculum, and relations with child development*. Dissertatie Universiteit Utrecht.
- UNICEF (2008). *The child care transition*. Florence: UNICEF Innocenti Research Centre.
- USC (2007). *Head Start Act*. Washington, DC: United States of America in Congress.
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M., Espinosa, L., Gorley, W., Ludwig, J., Magnuson, K., Philips, D., & Zaslow, M. (2013). *Investing in our future: The evidence base on preschool education*. Ann Arbor, MI: Society for Research in Child Development.
- Vasterman, J. (2016). Nu blijkt ineens: taalles voor jonge kinderen werkt dus wél. *NRC*, 7 juni 2016.
- Wikipedia (2016). *Voor- en Vroegschoolse Educatie*. Op 13/08/16 geraadpleegd op [https://nl.wikipedia.org/wiki/Voor-\\_en\\_vroegschoolse\\_educatie?oldid=47128613](https://nl.wikipedia.org/wiki/Voor-_en_vroegschoolse_educatie?oldid=47128613)