

Interventie en interventie-onderzoek

6

6.1 Inleiding

Van interventies nemen we aan dat ze helpen. Maar of dat echt zo is, is voor de meeste interventies niet onderzocht. Dat geldt ook voor de talloze interventies tegen pesten. Nu is het uitvoeren van een degelijk onderzoek naar de werkzaamheid van interventies niet bepaald makkelijk. Het inzetten van interventies die niet goed zijn onderzocht, brengt echter grote risico's met zich mee. In het minst erge geval is er sprake van verspilling van tijd en moeite; in het ergste geval nemen de problemen toe in plaats van af. Dat is zeker niet de bedoeling. Hoe beoordelen we de effectiviteit of werkzaamheid van interventies? Deze vraag staat hier centraal. Voor we die vraag beantwoorden, bespreken we eerst wat interventies zijn en welke verschillende soorten interventies er bestaan. Daarna laten we zien dat niet alle interventies helpen en dat problemen, zoals pesten en slachtofferschap, zelfs kunnen verergeren. Vervolgens gaan we in op effectiviteitsonderzoek, ook wel effectonderzoek of interventieonderzoek genoemd. Effectiviteitsonderzoek stelt de volgende vragen, aangeduid als de vijf W's: Wat Werkt Wanneer, voor Wie en Waarom? Beantwoording van die vragen vereist niet alleen dat interventies goed in elkaar zitten, maar ook dat het effectiviteitsonderzoek goed is opgezet. Is dat niet het geval, dan kun je geen harde conclusies over de effectiviteit trekken. Natuurlijk zijn ook de kosten van een interventie van belang. Daarom bespreken we de zogenoemde kosteneffectiviteit. Op basis van de kwaliteit van een interventie en het interventieonderzoek beoordeelt de Erkenningscommissie Interventies de effectiviteit van interventies. In de laatste paragraaf gaan we in op de niveaus van erkenning van deze commissie.

Dit hoofdstuk heeft tot doel een goede basis te leggen voor de volgende hoofdstukken over specifieke interventies.

6.2 Wat is een interventie?

Interventie betekent letterlijk 'tussenkomen'. Een eerste kenmerk van interventies is dat ze doelgericht zijn. Het doel van interventies is de psychische, sociale

en/of cognitieve ontwikkeling van kinderen en jongeren te bevorderen en om risico's en problemen in die ontwikkeling te voorkomen (Ince & Van den Berg, 2010). Daarmee zijn interventies in feite niet alleen gericht op het bevorderen van het welzijn van kinderen, maar op dat van alle betrokkenen. Dus ook op het welzijn van hun ouders en broertjes en zusjes, de gemeenschap of de buurt waarin ze leven en de bredere maatschappij (Yirmiya, 2010).

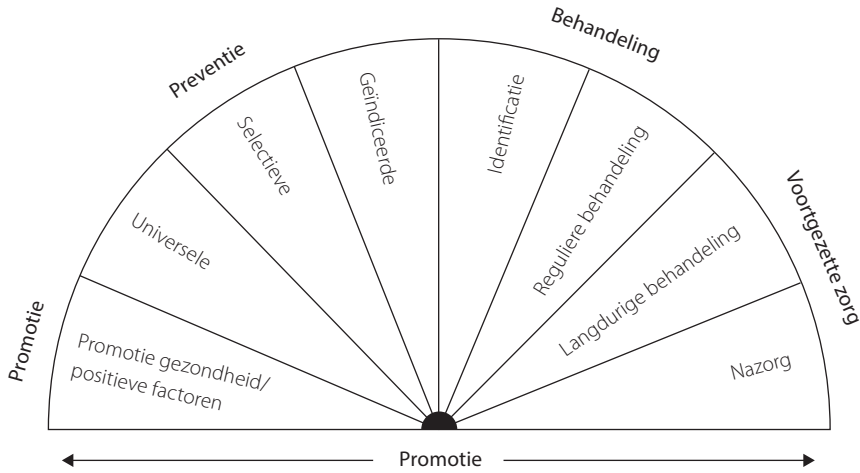
Daarnaast gaat het bij interventies om 'extra' of 'toegevoegde' activiteiten. Dit betekent dat interventies niet behoren tot de activiteiten die kinderen of ouders normaal krijgen. Een lesprogramma of lesmethode voor de basisschool is bijvoorbeeld geen interventie, maar een antipestprogramma voor de basisschool is dat wel (Ince & Van den Berg, 2010).

Tot slot hebben interventies de vorm van een project of programma; ze zijn planmatig. Dit houdt in dat de doelen goed omschreven zijn en geëvalueerd kunnen worden, dat de activiteiten en randvoorwaarden om de doelen te bereiken goed gedefinieerd zijn en dat de activiteiten afgebakend en/of gefaseerd zijn in de tijd (Ince & Van den Berg, 2010; Van Loon, Van der Meulen & Minnaert, 2011).

6.3 Het interventiespectrum

Je kunt soorten interventies onderscheiden naar het *moment* waarop de interventie ingrijpt. Ingrijpen kan als het probleem al zichtbaar aanwezig en vastgesteld is (behandeling en voortgezette zorg), maar ook als er nog geen probleem is of als het probleem zich aan het ontwikkelen is (promotie en preventie). Deze soorten interventies – promotie, preventie, behandeling en voortgezette zorg – zijn aan elkaar verwant. Zo kan een interventie die oorspronkelijk bedoeld is als behandeling, ook preventief werken op de lange termijn. *Minder Boos en Opstandig*, gericht op agressief en oppositioneel gedrag in de kindertijd, blijkt bijvoorbeeld preventief te werken tegen middelengebruik in de adolescentie (zie hoofdstuk 9). Ook zijn er interventies die je voor meerdere doeleinden kunt inzetten. Het antipestprogramma *Prima* (zie hoofdstuk 7) bijvoorbeeld, streeft ernaar pesten zowel te verminderen als te voorkomen. De opvatting is tegenwoordig dan ook dat promotie, preventie, behandeling en voortgezette zorg geen kwalitatief verschillende activiteiten zijn, maar een continuüm van hulpvormen. Figuur 6.1 geeft dat weer.

Het interventiecontinuüm is oorspronkelijk ontwikkeld voor de geestelijke gezondheidszorg (GGZ). Vandaar dat de focus op psychische stoornissen ligt. Pesten en gepest worden, zijn geen stoornissen. Niettemin is het model ook relevant om de soorten interventies op het gebied van pesten te karakteriseren. We bespreken het spectrum hierna.



Figuur 6.1 Het interventiespectrum (gebaseerd op Mrazek & Haggerty, 1994; O'Connell et al., 2009; Weisz et al., 2005)

Promotie

- *Gezondheidspromotie/versterking van positieve factoren* ('health promotion/positive development strategies') is gericht op de algemene bevolking. Het doel is het versterken van positieve factoren en daardoor het verkleinen van het risico op latere problemen en/of het vergroten van de kans op een gunstige ontwikkeling. Promotie is pas naderhand aan het model toegevoegd. Op het continuüm gaat promotie vooraf aan preventie, maar promotie is ook van belang tijdens preventie, behandeling en voortgezette zorg (O'Connell, Boat & Warner, 2009). Voorbeelden zijn het vergroten van sociale vaardigheden en het bevorderen van de kwaliteit van relaties met leeftijdgenoten (Weisz, Sandler, Durlak & Anton, 2005).
Op het terrein van pesten en ander regelovertredend gedrag is het *Taakspel* (zie hoofdstuk 8) een voorbeeld van gezondheidspromotie. Een van de doelen van het *Taakspel* is het scheppen van een positief en veilig klassenklimaat.

Preventie

Preventie omvat universele, selectieve en geïndiceerde preventie.

- *Universele preventie* is gericht op de algemene bevolking of een deel van de populatie zonder een verhoogd risico. De interventie is wenselijk voor iedereen. Voorbeelden zijn campagnes tegen zinloos geweld of tegen roken gericht op de hele bevolking en schoolprogramma's over alcoholgebruik gericht op alle leerlingen.

Een voorbeeld op het gebied van pesten is de campagne *Stop digitaal pesten* van SIRE. Deze is gericht op alle jongeren én hun ouders. De campagne spoort ouders aan om het gedrag van hun kinderen op internet beter te controleren.

- *Selectieve preventie* is gericht op hoogrisicogroepen die nog (vrijwel) geen problemen of klachten hebben. De hoogrisicogroepen hebben een belangrijke risicofactor gemeenschappelijk. De identificatie van risicogroepen is doorgaans niet gebaseerd op screening van de klachten of symptomen van het kind. Selectieve preventie van problematisch alcoholgebruik bij jongeren bijvoorbeeld, richt zich op subgroepen van jongeren die een verhoogd risico hebben op alcoholproblematiek, maar nog niet per se overmatig hoeven te drinken. Dit kunnen kinderen van ouders met een middelenprobleem zijn, kinderen in het speciaal onderwijs of kinderen met risicovolle persoonlijkheidskenmerken (Van Deursen, Saleminck, Lammers & Wiers, 2010).

Als het om pesten gaat, is de zomertraining *Plezier op School* (zie hoofdstuk 9) een voorbeeld van selectieve preventie. Deze training richt zich onder meer op aanstaande brugklassers die angstig of onhandig in de omgang met anderen zijn en daardoor risico lopen gepest te worden.

- *Geïndiceerde preventie* is gericht op individuen die niet voldoen aan de diagnostische criteria voor een psychische stoornis, maar die wel al een aantal symptomen of beginnende klachten hebben. Voorbeelden zijn interventies voor kinderen die al overmatig drinken, maar nog niet voldoen aan de diagnose alcoholmisbruik/alcoholafhankelijkheid.

Een geïndiceerde preventie die ook voor pesten relevant lijkt, is *Alles Kidzzz* (zie hoofdstuk 9). Deze interventie wil voorkomen dat kinderen met gedragsproblemen later een gedragsstoornis ontwikkelen.

De roep om effectieve promotie en preventie in de jeugdzorg en het onderwijs wordt steeds luider (Kutash, Duchnowski & Lynn, 2006; Mrazek & Haggerty, 1994; Weisz et al., 2005). Promotie en preventie kunnen immers niet alleen problemen voorkomen, zij kunnen ook de vraag naar behandeling en voortgezette zorg verminderen.

Als problemen toch tot ontwikkeling zijn gekomen, zijn behandeling en voortgezette zorg van belang.

Behandeling

Behandeling bestaat uit identificatie en reguliere behandeling.

- Bij *identificatie* ('case identification') gaat het om het vaststellen van stoornissen of problemen bij mensen of in groepen via screening en diagnostiek. Het is duidelijk dat identificatie in een zo vroeg mogelijk stadium en zo betrouwbaar mogelijk zou moeten gebeuren.

Het antipestprogramma *Prima* bijvoorbeeld, gebruikt de Pestmeter om daders en slachtoffers in schoolklassen te identificeren (zie hoofdstuk 2 en 7).

- *Reguliere behandeling* ('standard treatment' of 'time-limited therapy') houdt de toepassing van behandeling in van mensen of groepen bij wie een stoornis of probleem is vastgesteld. Het gaat om een beperkte periode van zorg, bijvoorbeeld twaalf sessies of totdat een bepaalde vooruitgang is geboekt. Een voorbeeld van reguliere behandeling op het gebied van pesten en andere gedragsproblemen is *Minder Boos en Opstandig* (zie hoofdstuk 9).

Voortgezette zorg

Onder voortgezette zorg vallen langdurige behandeling en nazorg aan personen bij wie de stoornis of het probleem blijft voortduren (Mrazek & Haggerty, 1994; Weisz et al., 2005). Belangrijke doelen zijn verbetering van de kwaliteit van leven en vermindering van de negatieve consequenties van de stoornis, ook voor mensen in de directe omgeving.

- *Langdurige behandeling* verwijst naar aanvullende behandeling. Het doel is het resultaat van de reguliere behandeling te versterken of uit te breiden bij cliënten die niet volledig hebben geprofiteerd van de behandeling of bij wie de kans op herhaling en terugval groot is. Een voorbeeld is het monitoren van medicatiegebruik (zoals antidepressiva) wanneer medicatie moet worden voortgezet.
- *Nazorg* omvat allerlei diensten waarbij rehabilitatie vooropstaat. Rehabilitatie heeft als doel het sociale en schoolse of beroepsmatige functioneren van cliënten met chronische problemen te bevorderen.

Een illustratie van voortgezette zorg op het terrein van pesten betreft *KiVa* (zie hoofdstuk 7). Als scholen *KiVa* twee jaar lang op de juiste manier inzetten, worden zij gecertificeerde *KiVa*-scholen. Deze scholen ontvangen bijscholing over actuele onderwerpen op het gebied van pesten

Samenvattend

Tegenwoordig ziet men promotie, preventie, behandeling en voortgezette zorg niet meer als kwalitatief verschillende activiteiten, maar als een continuüm van hulp (zie figuur 6.1). Veel interventies combineren elementen van deze categorieën. Het interventiecontinuüm is ook bruikbaar om de soorten interventies op het gebied van pesten te karakteriseren. Deskundigen menen dat in de jeugdzorg en het onderwijs meer nadruk op effectieve preventie (inclusief promotie) moet komen te liggen.

6.4 Noodzaak van effectonderzoek

Stel, de huisarts heeft je vanwege gezondheidsklachten doorgestuurd naar een specialist. Voor welke behandeling kies je?

- a) een behandeling waarvan de werkzaamheid is aangetoond;
- b) een behandeling die nog niet is onderzocht.

De keuze lijkt duidelijk. Toch denken mensen vaak anders wanneer het over interventies in de jeugdzorg gaat. In december 2013 konden mensen in de Nieuwsbrief Jeugd van het NJi antwoord geven op de vraag: *Een commissie gaat de effectiviteit van programma's tegen pesten beoordelen, omdat het kabinet wil dat scholen vanaf 2015 uitsluitend bewezen effectieve antipestprogramma's gebruiken. Een verstandige eis?* Het aantal stemmers was 131. Een verbijsterende 73,3% was tegen. Een van de argumenten was dat 'scholen uitstekend zelf in staat zijn om een passend antipestprogramma te maken' (De Lange, 2013).

In het voorgaande bespraken we dat interventies gericht zijn op het bevorderen van het welzijn van kinderen en andere betrokkenen. Vanwege dat doel gaat men er vaak van uit dat interventies inderdaad helpen. Maar kan men zonder meer uitgaan van een gunstig effect?

Helaas kunnen interventies tot toename van problemen leiden in plaats van afname. Al ruim vijftien jaar geleden stelden Dishion en Andrews (1995) vast dat een groepsinterventie voor tieners die het risico liepen delinquent te worden juist averechts werkte. Tieners die aan de interventie deelnamen, vertoonden na één jaar en ook nog na drie jaar juist meer delinquent gedrag dan de controlegroep. De onderzoekers verklaarden de onverwachte negatieve effecten door 'deviancy training'. Uit opnamen van de groepsbijeenkomsten bleek dat jongeren die over hun delinquente daden vertelden, werden beloond door de aandacht en het lachen van de andere groepsleden (Dishion, McCord & Poulin, 1999). Negatieve effecten zijn echter niet beperkt tot groepsinterventies. Ook individuele interventies kunnen leiden tot een toename van delinquent gedrag (Deković, 2010). Daarnaast worden ook op andere probleemgebieden, zoals angst en trauma, negatieve effecten van interventies gevonden (Barlow, 2010). Dat geldt ook voor programma's tegen pesten. Door sommige antipestprogramma's neemt het pesten of gepest worden juist toe in plaats van af (Ttofi & Farrington, 2011).

Wat nu als door een interventie de problemen weliswaar niet verergeren, maar ook niet verminderen? Geldt dan 'baat het niet, dan schaadt het niet'? Interventies brengen kosten met zich mee. Als een interventie niet helpt, betekent dat verspilling van geld, tijd, moeite en menskracht. In die zin levert toepassing van een ineffectieve interventie wel degelijk schade op. Daarnaast heeft deelname aan een interventie die niet helpt gevolgen voor de cliënten. Deković (2010) noemt demoralisatie, een gevoel van hopeloosheid en afname van de bereidheid om in de toekomst hulp te zoeken en te aanvaarden. Ook in die zin levert een niet-werkzame interventie dus schade op. Kortom, het adagium 'baat het niet, dan schaadt het niet' gaat niet op voor interventies.

Effectonderzoek kan aantonen of een interventie daadwerkelijk helpt of niet. Slechts een klein deel van alle interventies is echter geëvalueerd aan de hand van deugdelijk onderzoek dat conclusies over effectiviteit toelaat. Dit bespreken we in de volgende paragrafen.

6.5 Effectonderzoek: de vijf W's

Effectonderzoek – ook wel effectiviteitsonderzoek of interventieonderzoek genoemd – zou antwoord moeten geven op een aantal vragen, die bij wijze van ezelsbruggetje de vijf W's worden genoemd (Orobio de Castro, 2007; Yirmiya, 2010):

- *Wat*: waaruit bestaat de interventie (op papier en in de praktijk)?
- *Werkt*: welke effecten – gewenst en ongewenst – heeft de interventie? Ten opzichte van welke conditie (niets doen of andere interventie; individuele bijeenkomsten of groepsbijeenkomsten, ouder- of kindcursussen, of een combinatie van typen)?
- *Wanneer*: onder welke omstandigheden – bijvoorbeeld kwaliteit van de uitvoering van de interventie; therapeutische alliantie – treden die effecten het sterkst op?
- Voor *wie*: bij welke deelnemers – bijvoorbeeld van welke leeftijd en sekse – treden de effecten (het sterkst) op?
- *Waarom*: welke mechanismen veranderen door de interventie? Hierbij gaat het om de toetsing van de theorie of het model inzake risicofactoren en beschermende/bevorderlijke factoren achter de interventie.

Om antwoord te geven op die vragen moet zowel de interventie als het effectonderzoek aan bepaalde eisen voldoen (Kazdin, 2003). Hoewel er verschillende criteria bestaan (European Platform for Investing in Children (EPIC), 2014; Trimbos-instituut, 2014; Van Yperen, 2007), gaat het echter steeds om dezelfde aandachtspunten: een interventie moet goed zijn beschreven en onderbouwd (eisen aan de interventie) én het effectonderzoek moet zowel interne als externe validiteit bezitten (eisen aan het effectonderzoek).

6.6 Eisen aan de interventie: duidelijk beschreven, toepasbaar in de praktijk en theoretisch onderbouwd ('Wat?')

Duidelijk beschreven

Effectieve hulp veronderstelt dat volstrekt duidelijk moet zijn voor welke problemen de hulp oplossing biedt. Met andere woorden: het *expliciete doel* van de interventie moet vaststaan. Een van de doelen van de voorloper van de *Prima*-methode is bijvoorbeeld 'het verminderen van het aantal basisschoolleerlingen

dat gepest wordt met minimaal 60%' (Fekkes & Paulussen, 2012). Daarnaast moeten de relevante kenmerken van de *doelgroep* zijn aangegeven (zoals leeftijd en specifieke problemen) en de kenmerken die deelname uitsluiten (bijvoorbeeld een verstandelijke handicap).

Het moet eveneens helder zijn wat de *werkwijze of methodiek* inhoudt. Een goede beschrijving bevat doorgaans een stappenplan of een protocol dat aangeeft in welke volgorde en met welke frequentie, duur en intensiteit de activiteiten plaatsvinden. Daarnaast is een overzicht van de in te zetten middelen nodig. De middelen omvatten mensen (zoals de schoolleiding, coördinator, leerkrachten en ouders) en het benodigde materiaal (zoals speciale werkboekjes).

Toepasbaar in de praktijk

Een interventie moet natuurlijk ook uitvoerbaar zijn. Dit betekent ten eerste dat de interventie *overdraagbaar* moet zijn. Hulpverleners moeten de interventie kunnen leren, zodat zij haar kunnen uitvoeren. Daarvoor is meestal een duidelijke handleiding of een protocol nodig. De kwaliteit van de trainers en de beschikbaarheid van ervaren begeleiders of supervisors zijn eveneens belangrijk.

Ten tweede moeten de *randvoorwaarden* duidelijk zijn. Dit zijn voorwaarden waaraan voldaan moet zijn, wil de interventie het beoogde doel ooit kunnen bereiken. Het gaat hier om kosten in termen van personeel, tijdsinvestering en materiaal. Daarnaast gaat het om de steun en medewerking van de instelling, die de interventie immers moet betalen en organiseren. Vaak moet de hulpverlener aan bepaalde eisen voldoen wat betreft kennis (bijvoorbeeld van leertheoretische principes) en ervaring (bijvoorbeeld met het werken met groepen). Soms moet ook de omgeving aan bepaalde voorwaarden voldoen. De omgeving moet bijvoorbeeld rustig en veilig zijn.

Tot slot moet een interventie *voldoende flexibel* zijn om te kunnen inspelen op specifieke situaties of op nieuwe problemen die aan het licht komen.

Theoretisch onderbouwd

Als een interventie niet is afgestemd op de oorzaken van het probleemgedrag, is de kans groot dat de interventie geen effect heeft. Interventies zouden daarom gebaseerd moeten zijn op duidelijke modellen voor het ontstaan en het in stand blijven van de problemen waarop ze zijn gericht. Zulke 'ontwikkelingsmodellen' zijn, met andere woorden, gebaseerd op een *theoretisch kader* of op een *analyse van risicofactoren en beschermende/bevorderlijke factoren* die het ontstaan of voortbestaan van problemen kunnen verklaren (Matthys & Lochman, 2010). Theoretische onderbouwing maakt aannemelijk dat de interventie zou kunnen werken. De mechanismen die verantwoordelijk zijn voor de problemen worden immers aangepakt.

In de volgende hoofdstukken zien we dat interventies tegen pesten bijvoorbeeld zijn gebaseerd op theoretische kaders als de rolbenadering van pesten (*KiVa* in hoofdstuk 7, *Friendly ATTAC* in hoofdstuk 8) en het sociale-informatieverwerkingsmodel van Crick en Dodge (*Alles Kidzzz, Ik kies voor Zelfcontrole* en *Minder Boos en Opstandig* in hoofdstuk 9).

De eis van theoretische onderbouwing sluit niet uit dat interventies in de praktijk kunnen ontstaan. Geconfronteerd met concrete hulpvragen kan een hulpverlener een zinvolle interventie ontwerpen en zich daarbij baseren op praktijkinzichten. Deze kunnen later theoretische en empirische ondersteuning krijgen. Een voorbeeld hiervan is de weerbaarheidstraining *Sta Sterk* van de Stichting Omgaan met Pesten (www.omgaanmetpesten.nl). Deze training was al geïmplementeerd. De inventarisatie van de theoretische grondslagen ervan gebeurde achteraf met behulp van onderzoekers (zie hoofdstuk 9).

Samenvattend

Er zijn drie basisvoorwaarden waaraan een goede interventie moet voldoen. De eerste is een goede beschrijving van het doel, de doelgroep en de methodiek. De tweede is uitvoerbaarheid in de praktijk: kunnen hulpverleners de interventie leren en wat is er nodig om de interventie in de praktijk te kunnen toepassen? De derde voorwaarde, theoretische onderbouwing, maakt aannemelijk dat de interventie zou kunnen werken, omdat de interventie factoren en processen aanpakt die verantwoordelijk zijn voor de problemen. Een interventie die een kans van slagen heeft, voldoet aan deze drie voorwaarden. Of de interventie écht werkt, moet blijken uit het effectonderzoek.

6.7 Eisen aan effectonderzoek: interne en externe validiteit

Interne validiteit

Bij interne validiteit gaat het om de vraag of de interventie verantwoordelijk is voor de eventuele verandering na de interventie of dat externe factoren ('confounders') een rol spelen. Twee belangrijke onderzoeksdesigns in dit verband zijn het quasi-experimentele design en de 'randomized controlled trial'.

Quasi-experimenteel design. De effectiviteit van een interventie kan enigszins worden aangetoond met een quasi-experimenteel onderzoek waarbij een interventiegroep (experimentele conditie) wordt vergeleken met een controlegroep (controleconditie). De controlegroep kan bestaan uit kinderen die geen interventie krijgen, op de wachtlijst staan, een placebo krijgen of een alternatieve interventie (actieve controlegroep) ontvangen. De alternatieve interventie kan bestaan uit één andere interventie of een keuze van meerdere interventies, zoals reguliere of gebruikelijke zorg ('care as usual') (Gresham & Lochman, 2008). Door de interventiegroep met een controlegroep te vergelijken, is de interne

validiteit hoger dan wanneer er geen controlegroep is. Vergelijking met een controlegroep maakt aannemelijk dat de interventie het effect veroorzaakt. Zonder controlegroep weten we niet zeker of de veranderingen echt toe te schrijven zijn aan de interventie of juist aan andere oorzaken (zoals verloop van tijd).

Een nadeel van quasi-experimenteel onderzoek is dat de verdeling van personen over de experimentele conditie en de controleconditie niet op basis van toeval gebeurt, maar op basis van andere factoren, zoals tijd of beschikbaarheid van therapeuten. Hierdoor kan het zijn dat beide groepen verschillen op relevante kenmerken zoals leeftijd en sekse. Hierdoor is de interne validiteit in gevaar; het onderzoek zal dan niet kunnen aantonen dat het effect van de interventie daadwerkelijk door de interventie komt en niet is toe te schrijven aan de verschillen tussen de groepen. De oplossing voor dit probleem is gebruik te maken van matchingprocedures of een 'matched control group'. Aan de hand van relevante kenmerken van de interventiegroep – zoals ernst van de problemen, aanwezigheid van andere problemen (comorbiditeit), sekse en leeftijd – wordt een controlegroep samengesteld met dezelfde kenmerken ('matching'), zodat beide groepen ongeveer gelijk zijn en de uitkomsten na interventie met elkaar vergeleken kunnen worden.

Een ander voorbeeld van een quasi-experimenteel design, dat vaak wordt gebruikt bij onderzoek naar schoolprogramma's tegen pesten, is het leeftijds-cohortdesign ('age cohort design'). Olweus (1991) bijvoorbeeld, vergeleek leerlingen na het volgen van het Olweus Bullying Prevention Program (de experimentele conditie) met andere leerlingen van dezelfde leeftijd in dezelfde school vóór het volgen van het programma (de controleconditie). Op die manier probeerde hij te bereiken dat de leerlingen uit beide condities vergelijkbaar waren.

Randomized controlled trial. De beste manier om effectiviteit aan te tonen, is echter het uitvoeren van een randomized controlled trial (RCT). Bij een RCT vindt de toewijzing van personen aan een experimentele conditie (interventiegroep) of een controleconditie (zoals reguliere zorg) plaats op basis van toeval (at random). Bij onderzoek naar pestinterventies worden ook wel eens hele scholen of klassen at random aan condities toegewezen. Bij een voldoende grote steekproef zorgt at random toewijzing ervoor dat de groepen gelijk zijn. Hierdoor heeft een RCT de hoogste mate van interne validiteit.

Hoewel de RCT de gouden standaard is wat betreft effectonderzoek (Farrington, 2003), zijn er ook mensen met bezwaren. Veel behandelaren vinden het onethisch of onaanvaardbaar dat op basis van toeval iemand in de ene conditie terecht komt of in de andere. Verder lenen sommige interventies zich minder voor onderzoek met een RCT (Kelly et al., 2010). In het geval van praktische en principiële bezwaren zijn onderzoekers vaak aangewezen op alternatieven, zoals quasi-experimenteel design, herhaalde $N = 1$ studies en andere statistische technieken (Cook, 2003). Ook al worden deze methoden vaak gebruikt, ze hebben minder bewijskracht dan een RCT (Kelly et al., 2010).

Tot slot nog een kanttekening die voor quasi-experimentele designs én RCT's geldt. Cliënten die meedoen aan effectonderzoek kunnen door het onderzoek

sensitiever worden voor problemen. Kinderen krijgen bijvoorbeeld op school lessen en werkboekjes over pesten. Daardoor kunnen ze meer dan voorheen gaan letten op pesten en pesten meer gaan rapporteren dan vóór de interventie. Hierdoor lijken de effecten van de interventie kleiner. Het is dan ook belangrijk om pesten op verschillende manieren, via verschillende informanten en over een langere periode te meten.

Tot de interne validiteitsproblemen rekenen we ook de *behandelintegriteit*. Wanneer de behandeling anders is uitgevoerd dan voorgeschreven, kun je je afvragen of de resultaten uit het onderzoek wel betrouwbaar zijn. Een manier om dit te controleren is het onderzoeken van de behandelintegriteit. Behandelintegriteit (of programma-integriteit) is een controle op de juiste uitvoering van de behandeling: is de uitvoering van de behandeling gebeurd zoals voorgeschreven volgens het protocol, met andere woorden: was de implementatie juist? Het kan bijvoorbeeld gebeuren dat therapeuten of trainers het programma verkorten of verlengen. Als de behandelintegriteit laag is, dan wil dat zeggen dat de uitvoering van de interventie heel variabel gebeurt. Een hoge behandelintegriteit betekent daarentegen dat de behandeling voor het grootste gedeelte op dezelfde manier plaatsvindt. Zonder behandelintegriteit is de interpretatie van resultaten uit effectonderzoek niet goed mogelijk omdat het dan eigenlijk onderzoek naar meerdere vormen van één interventie betreft.

Externe validiteit

Een ander aspect dat een rol speelt bij effectonderzoek is externe validiteit. Een hoge externe validiteit betekent dat het onderzoek zo is uitgevoerd dat het een betrouwbare afspiegeling is van hoe het er in de praktijk aan toe gaat. De onderzoeksresultaten zijn dan goed te generaliseren. Effectonderzoek kan intern valide zijn – omdat het via een goed opgezette RCT is onderzocht –, maar dat hoeft niet te betekenen dat het ook extern valide is. Effectonderzoek gebeurt vaak in een onderzoeksetting (zoals universiteit) waarbij de werving van cliënten via advertenties plaatsvindt. Pas afgestudeerde en getrainde psychologen die voortdurend supervisie krijgen, geven daarbij de interventie (Kazdin, 2003). In de praktijk hebben cliënten echter vaak ernstige klinische problemen en is er sprake van comorbiditeit. Therapeuten hebben dikwijls minder mogelijkheden voor training en supervisie. Om daadwerkelijk uitspraken te doen over het effect van een interventie in de praktijk ('real world effectiveness') is het nodig RCT's uit te voeren in die praktijk (scholen, klinieken, instellingen, et cetera). In de Engelstalige literatuur maakt men in dit verband onderscheid tussen 'efficacy' en 'effectiveness'. *Efficacy* betekent dat het onderzoek uitgevoerd is onder gecontroleerde en geïdealiseerde condities. Dus met randomisatie, geselecteerde cliënten (beperkte comorbiditeit), de beste therapeuten, een goede en intensieve training en supervisie. Er is dus een lage mate van externe validiteit, maar wel een hoge mate van interne validiteit. Als blijkt dat de interventie effectiever is dan de controleconditie, spreekt men over *efficacy*. *Effectiveness* betekent dat

het onderzoek is uitgevoerd in de klinische praktijk. Zowel de cliënten als de therapeuten en de uitvoering van de therapie zijn een goede representatie van hoe het er in de praktijk aan toe gaat. Met andere woorden: een hoge mate van interne én externe validiteit, met de focus op de generaliseerbaarheid van resultaten (externe validiteit). Als uit het onderzoek blijkt dat de interventie beter werkt dan de controleconditie, spreekt men over effectiveness (Weisz, Jensen & McLeod, 2005).

Samenvattend

Het onderzoeksdesign met de hoogste mate van interne validiteit is de RCT. Door een RCT kan bepaald worden of de interventie daadwerkelijk verantwoordelijk is voor verandering van de problemen. In de praktijk komen echter veel quasi-experimentele onderzoeken voor. Het onderzoeksdesign met de hoogste mate van externe validiteit is onderzoek dat een afspiegeling is van de klinische praktijk. Een RCT uitgevoerd in de klinische praktijk verdient dan ook de wereldtitel voor goed effectonderzoek.

6.8 Uitkomsten van interventieonderzoek ('Werkt?')

Het gebruik van meerdere uitkomstmaten

Om het effect van een interventie te meten, gebruiken we meerdere 'uitkomstmaten'. Uitkomstmaten bestaan vaak uit verschillende (typen) instrumenten, zoals vragenlijsten, interviews of observaties en worden idealiter afgenomen bij meerdere informanten. Verschillende informanten (bijvoorbeeld ouder en kind) kunnen immers anders tegen zaken aankijken. In interventieonderzoek worden de instrumenten vlak voor de interventie (voormeting) afgenomen en vervolgens nogmaals na de interventie (nameting) om te zien of het gedrag is veranderd.

Bij het gebruik van meerdere uitkomstmaten kun je onderscheid maken tussen *primaire* en *secundaire* uitkomstmaten. Primaire uitkomstmaten zijn de belangrijkste maten waar je een effect op verwacht. De primaire uitkomstmaten moeten goed corresponderen met het hoofddoel van de interventie, zoals een reductie van pesten of gepest worden of ander probleemgedrag. Secundaire uitkomstmaten corresponderen met de subdoelen van de interventie waar men (indirect) een effect op verwacht. Voorbeelden zijn communicatie, sociale vaardigheden, opvoedvaardigheden en sfeer in de klas.

Naast de duidelijke voordelen van meerdere uitkomstmaten zijn er echter ook nadelen. Wat als de interventie wel verbetering op sommige uitkomstmaten laat zien, maar geen verbetering op andere? Werkt de interventie dan wel of niet? En wat als een interventie alleen werkt op secundaire uitkomstmaten en niet op de primaire uitkomstmaten? Of als uitkomstmaten op basis van ouderrapportage

wel positieve effecten laten zien, maar de uitkomstmaten op basis van kindrapportage niet? Kun je dan wel concluderen dat de interventie wel of niet effectief is? Hier zijn geen richtlijnen voor; de onderzoeker dient zelf kritisch te zijn.

Nameting en follow-up

Een ander aspect dat meespeelt bij de conclusie of een interventie effectief is, is de tijd. Vaak vinden zogeheten follow-upmetingen plaats om te onderzoeken of het effect gedurende een bepaalde tijd (bijvoorbeeld een halfjaar of een jaar) beklijft. Als er wel een effect wordt gevonden vlak na de interventie, maar niet meer na een jaar, dan is er dus alleen een kortetermijneffect en geen langetermijneffect. Een voorbeeld hiervan is de voorloper van het antipest-programma *Prima*. In een effectstudie bleek na het interventiejaar zowel pesten als gepest worden significant te zijn afgenomen op de interventiescholen vergeleken met de controlescholen. Na twee jaar waren deze verschillen echter weer verdwenen (Fekkes, Pijpers & Verloove-Vanhorick, 2006).

Andersom kan het ook: soms vindt men geen effect op de nameting, maar wel op de follow-upmeting. Dit heet 'sleeper effect', omdat het effect even tijd nodig heeft om te ontwaken en zichtbaar te worden. Er is dan alleen een langetermijneffect en geen of slechts een klein kortetermijneffect. Zo bleek het positieve effect van een preventieve interventie gericht op depressie na twee jaar bijna verdubbeld (Gillham, Reivich, Jaycox & Seligman, 1995). Ook negatieve effecten blijken soms pas na een tijdje op te treden. De negatieve effecten van groepsinterventie bij delinquente jongeren, beschreven in paragraaf 6.4, bleken niet meteen na de interventie, maar pas na één en drie jaar op te treden. Het is dus belangrijk om op meerdere tijdstipmomenten een meting te doen om te kijken of de effecten mogelijk veranderen.

Statistische significantie

Het effect van een interventie drukt men meestal uit in een statistisch significant verschil op de nameting tussen de interventiegroep en de controlegroep. Een verschil is statistisch significant als de kans klein is dat het op toeval berust. De statistische analyses¹ geven antwoord op de vraag: werkt de interventie beter dan de controlegroep?

1 Bijvoorbeeld (M)ANOVA, regressieanalyse en multilevel analyse. Vaak wordt hierbij de voormeting voor de zekerheid als zogenaemde 'covariaat' meegenomen. Dit houdt in dat men corrigeert voor eventuele verschillen in de ernst van het probleemgedrag tussen de experimentele groep en de controlegroep op de voormeting.

Effectgrootte en odds ratio

Een statistisch significant verschil dat uitwijst dat de interventie de beoogde effecten sorteert, is mooi. De vraag blijft echter: wat stelt dat verschil dan voor? Deze vraag is vooral van belang bij grote steekproeven, omdat je daar eerder significante resultaten vindt dan bij kleine steekproeven. Vandaar dat onderzoeksresultaten bij voorkeur ook de effectgrootte ('effect size') bevatten. De effectgrootte wordt berekend tussen twee groepen en geeft informatie over de omvang van het verschil tussen die groepen. Het voordeel van de effectgrootte is dat het een gestandaardiseerde score is. Hierdoor is vergelijking op verschillende uitkomstmaten (die vaak gebaseerd zijn op verschillende instrumenten en/of informanten) mogelijk. Dat geldt ook voor de resultaten van verschillende studies. De bekendste index voor effectgrootte is Cohen's d (zie kader 6.1). De odds ratio (OR) is een andere statistische vorm van een effectgrootte, maar dan voor dichotome variabelen (bijvoorbeeld wel of niet gepest worden, het wel of niet hebben van een bepaalde diagnose) (zie kader 6.1).

Bij de interpretatie van de effectgrootte is het belangrijk om rekening te houden met het type controleconditie. De effectgroottes bij een vergelijking van een interventiegroep met een controlegroep die geen interventie krijgt, zijn doorgaans groter dan wanneer de controlegroep een alternatieve interventie aangeboden krijgt. De vergelijking tussen een interventiegroep en een controlegroep die zorg krijgt, is dus strenger (Weisz, Jensen-Doss & Hawley, 2006).



Voorbeeld

Cohen's d en odds ratio

Cohen's d wordt berekend voor continue variabelen. Dit zijn variabelen die iedere mogelijke waarde kunnen aannemen in een bepaald interval (bijvoorbeeld *de mate* van gepest worden). De berekening van Cohen's d gebeurt als volgt: het verschil tussen het gemiddelde van de interventiegroep en het gemiddelde van de controlegroep op de nameting, gedeeld door de gepoolde standaarddeviatie van deze twee groepen: $(M_{\text{interventie}} - M_{\text{controle}}) / SD_{\text{pooled}}$.² Bij d behoort een (95%) betrouwbaarheidsinterval dat een range van mogelijke waarden rond de gevonden d aangeeft.

Doorgaans gelden de volgende interpretaties van Cohen's d : verslechtering of negatief effect ($d < 0$), geen effect ($d = 0$), klein effect ($d = 0,2$), matig effect ($d = 0,5$) en groot effect ($d = 0,8$). Er zijn omzettingen bedacht om de effectgroottes beter te interpreteren. Een voorbeeld hiervan is het omzetten van de effectgrootte naar een percentage niet-behandelde cliënten dat beter of slechter af is dan de gemiddelde behandelde cliënt (= 50%). Hoe groter de effectgrootte, hoe hoger het percentage. Bij een klein effect ($d = 0,2$) is de cliënt in het algemeen beter af dan 58% in de controlegroep; er is dan slechts 8% winst. Bij een matig effect ($d = 0,5$) is de cliënt beter af dan 69% in de controlegroep (19% winst) en bij een groot effect ($d = 0,8$) is de cliënt beter af dan 79% in de controlegroep (29% winst) (Van Yperen, 2010). Durlak (2009) waarschuwt terecht tegen het klakkeloos gebruiken

2 Er zijn correcties of aanpassingen van de formule, bijvoorbeeld voor kleine steekproeven (Durlak, 2009).

van de vuistregels van Cohen. Bij de interpretatie van d moet je bijvoorbeeld rekening houden met de kwaliteit en het design van de studie. Effectgroottes van verschillende studies kun je alleen goed interpreteren als die studies vergelijkbaar zijn, bijvoorbeeld qua controlegroep, uitkomstmaten, informanten en het betrouwbaarheidsinterval van d .³ Ook is het goed om stil te staan bij het feit dat zelfs kleine effectgroottes toch heel belangrijk kunnen zijn, vooral bij ernstige diagnoses of condities. Stel dat een interventie voor depressie ervoor zorgt dat de kans op een suïcidepoging afneemt ($d = 0,3$), dan is een winst van 12% wel degelijk waardevol.

De berekening van de odds ratio (OR) gebeurt voor dichotome variabelen. Dit zijn variabelen die slechts twee waarden kunnen hebben (bijvoorbeeld wel of niet gepest worden). Een *odd* (in het Nederlands 'wedverhouding' of 'kansverhouding') is de verhouding tussen de waarschijnlijkheid dat iets gebeurt (of zo is) en de waarschijnlijkheid dat iets niet gebeurt (of niet zo is). Stel Matty en Albert zijn twee tennissers die in de afgelopen wedstrijden vijf keer tegen elkaar hebben gespeeld. Matty heeft drie keer gewonnen en Albert twee keer. De odds dat Matty de volgende wedstrijd wint, is 3 tegen 2 (3:2). Een *odds ratio* (OR) betekent dat twee odds worden vergeleken, namelijk de odds voor succes voor de interventiegroep en de odds voor succes voor de controlegroep. Het volgende rekenvoorbeeld maakt dit duidelijk. Er is een pestinterventiegroep en een controlegroep. Bij de nameting kijkt men hoeveel kinderen er nog worden gepest:

	Interventiegroep	Controlegroep
Aantal kinderen dat <i>niet</i> wordt gepest	80 (A)	50 (B)
Aantal kinderen dat wordt gepest	20 (C)	50 (D)
$OR = \frac{80/20}{50/50} = \frac{AD}{BC} = \frac{A/B}{C/D} = 4$		

Een OR kleiner dan 1 betekent een ongewenst of negatief effect, een waarde van 1 betekent geen effect en een waarde groter dan 1 is een positief effect.

Ook bij de OR behoort een (95%) betrouwbaarheidsinterval. Wanneer de waarde 1 niet in het betrouwbaarheidsinterval ligt, dan is de odds ratio significant.

De OR uit het rekenvoorbeeld is 4 en blijkt significant te zijn. Dit betekent dat de odds voor niet gepest worden voor de interventiegroep vier keer zo hoog zijn als voor de controlegroep.⁴ Een kind in de interventiegroep heeft dus een duidelijk hogere kans om niet meer te worden gepest dan een kind uit de controlegroep.

Kader 6.1

-
- 3 Als het betrouwbaarheidsinterval groot is, wat vaak bij een kleine steekproef het geval is, dan is de schatting van d niet zo precies. Het is niet juist om precieze en minder precieze d 's te vergelijken en conclusies over het verschil in grootte te trekken.
- 4 Dit betekent *niet* dat een kind in de interventiegroep vier keer minder kans heeft om te worden gepest dan een kind in de controlegroep. Zo'n interpretatie zou gebaseerd moeten zijn op de berekening van relatieve risico ratio's (RR's) en niet op OR's.

Klinische significantie en Reliable Change Index

Statistische significantie geeft dus aan dat de kans klein is dat het verschil tussen de interventiegroep en de controlegroep op toeval berust. De effectgrootte geeft aan hoe groot dat verschil is. Maar wat kunnen we daarmee in de praktijk? Wanneer is een resultaat *klinisch significant* of relevant? Het gaat hier om de vraag of de score van een persoon van problematisch, disfunctioneel of klinisch is veranderd naar die van de 'normale' populatie of normgroep. Een voorbeeld hiervan is de score op de veelgebruikte vragenlijst, de Child Behavior Checklist (CBCL). Na afloop van het onderzoek kijkt de onderzoeker bij hoeveel kinderen (percentage) in de interventie- en controlegroep de score onder het klinische (grens)gebied is gezakt (zie bijvoorbeeld *Alles Kidzzz* in hoofdstuk 9). Ook het wel of niet meer voldoen aan de criteria van een diagnose hoort hierbij. Indien klinische afkappunten ('cut-off scores') ontbreken, kan de onderzoeker ze trouwens zelf berekenen (Asscher, Dekovic, Prinzie, Hermanns & Van den Akker, 2009; Jacobson, Folette & Revenstorf, 1984; Jacobson & Truax, 1991). Statistisch significant betekent niet automatisch dat resultaten ook klinisch significant zijn. Zo is het mogelijk dat de interventiegroep vergeleken met de controlegroep een flinke afname in het probleemgedrag laat zien en dat de effectgrootte (d) groot is, maar dat veel kinderen nog steeds in het klinische gebied scoren. Er is dan nog altijd werk aan de winkel.

Daarmee zijn we er nog niet. Een proefpersoon vult een vragenlijst vrijwel nooit twee keer hetzelfde in. De score varieert bijvoorbeeld door leereffecten. Om na te gaan of de verandering op een vragenlijst echt een verandering in gedrag, gevoel of denken weergeeft en niet aan de (test-hertest)betrouwbaarheid van het instrument ligt, raadt men aan om voor elke persoon de betrouwbare verandering of Reliable Change Index (RCI; Jacobson & Truax, 1991) te berekenen. De RCI geeft de statistisch significante *verandering* aan. Dat wil zeggen of een kind *vooruitgang* heeft geboekt, of er sprake is van *achteruitgang* of dat een kind bij benadering *gelijk is gebleven*.⁵ In dat laatste geval is de verandering niet groot genoeg om betekenisvol te zijn. Vandaar dat men ook wel zegt dat de RCI betekenisvolle verandering aangeeft.

Het berekenen van statistische significantie en effectgrootte is gericht op groepsgegevens en verschaft waardevolle informatie op groepsniveau. De klinische significantie en de RCI geven aanvullende informatie over de betekenis

5 Een betrouwbare verandering betekent dat een verandering uitgedrukt in de verschillen scores gecorrigeerd wordt voor de standaardmeetfout. Meer formeel: de RCI is de score op de voormeting minus de score op de nameting, gedeeld door de standaardfout van het verschil tussen de twee uitkomstmaten (S_{diff}). S_{diff} wordt berekend op basis van de standaardfout van de meting ('standard error of measurement', SEM), met behulp van de volgende formule: $S_{diff} = \sqrt{2(SEM)^2}$. Als de berekende RCI groter is dan 1,96 of kleiner is dan -1,96, dan is de verandering groot genoeg om betekenisvol te zijn. Vervolgens kun je de percentages berekenen van het aantal kinderen dat vooruitgang toonde ($RCI > 1,96$), achteruitgang liet zien ($RCI < -1,96$) of bij benadering hetzelfde bleef ($-1,96 > RCI > 1,96$).

van de veranderingen binnen de groep. Klinische significantie geeft aan of de veranderingen die zijn opgetreden ook reële verandering in het dagelijks functioneren betekenen (Kendall, 1999). De RCI geeft aan of de veranderingen groot genoeg zijn om betekenisvol te zijn en of ze dan positief of negatief zijn (Hawley, 1995). In het ideale geval vind je zowel betrouwbare positieve verandering als klinische significantie.

Conclusies uit meerdere effectstudies

Uitkomsten uit effectonderzoek kunnen worden gecombineerd in een review of een meta-analyse. Dit maakt algemene uitspraken mogelijk over de effectiviteit van een bepaald programma of een bepaald type programma (bijvoorbeeld schoolbrede programma's tegen pesten). Er moeten dan wel al meerdere onderzoeken zijn geweest die dezelfde interventie of doelgroep hebben onderzocht. Een *review* is een samenvatting van meerdere onderzoeksresultaten op basis van een literatuurstudie. In een *meta-analyse* worden meerdere onderzoeken samengenomen en opnieuw, maar nu gelijktijdig geanalyseerd, zodat een uitspraak kan worden gedaan over het gemiddelde behandelingseffect. De analyse berekent een globale gemiddelde effectgrootte, gebaseerd op de verschillende effectgroottes uit meerdere studies. Een meta-analyse komt vaak met een strenger oordeel dan een review. Auteurs van reviewartikelen gaan vaak mee met de conclusies van auteurs van interventiestudies en die zijn vaak gunstiger dan uit de statistische resultaten blijkt.

In het algemeen lijken interventieprogramma's voor jeugdigen een klein effect te hebben (Weisz et al., 2006). Meta-analyses naar gebruikelijke zorg bij jeugdigen met diverse diagnoses laten zelfs gemiddeld helemaal geen effect zien (Weiss, Catron, Harris & Phung, 1999; Weisz, 2004; Weisz, Donenberg, Han & Weiss, 1995). Uit meta-analyses van schoolbrede programma's tegen pesten blijken bescheiden positieve uitkomsten (zie paragraaf 7.6).

Samenvattend

Het heeft de voorkeur om in effectonderzoek te kijken naar veranderingen op meerdere meetmomenten (voor de interventie, na de interventie en follow-ups), op meerdere uitkomstmaten (primair en secundair) en afgenomen bij meerdere informanten (kind, ouder, leerkracht, observaties). Statistische significantie geeft aan dat het verschil op de nameting tussen de interventiegroep en de controlegroep niet op toeval berust. Voor interpretatie van de omvang van dit verschil is het zinvol om een effectgrootte te berekenen. Naast deze klassieke aanpak van effectiviteitsonderzoek die naar groepsgemiddelden kijkt, is het nuttig om te bepalen of de kinderen zo zijn veranderd dat hun scores niet meer in het klinisch gebied vallen (klinische significantie). Bovendien is het aan te raden om aan de hand van de RCI te bepalen of een kind betrouwbaar vooruit

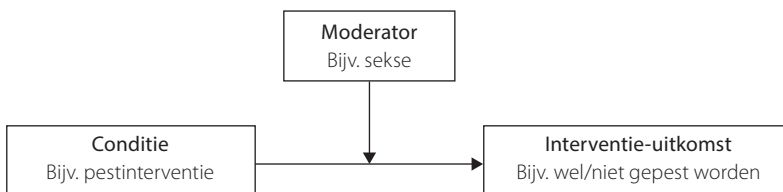
of achteruit is gegaan. Meta-analyses zijn belangrijk, omdat deze de uitkomsten van meerdere studies combineren. Dit biedt de mogelijkheid voor het doen van één globale uitspraak over de effectiviteit van een aantal programma's of een bepaald type programma.

6.9 Moderatoren ('Wanneer' en 'Voor wie')

Zoals te verwachten valt, werkt een interventie niet bij iedereen hetzelfde. De een heeft er meer profijt van dan de ander. Vandaar dat het ook belangrijk is om te onderzoeken wanneer een interventie effectief is en voor wie, de zogenoemde *moderatoren*.

Moderatoren zijn kenmerken of omstandigheden die de richting (positief of negatief) of de sterkte van de relatie tussen interventie en de behandeluitkomst beïnvloeden (zie figuur 6.2). Met andere woorden: wanneer/onder welke omstandigheden en voor wie werkt de interventie het best? Voorbeelden van moderatoren zijn kenmerken van het kind (zoals sekse, leeftijd, etniciteit en aard en ernst van psychische klachten), kenmerken van de ouders (zoals sociaaleconomische status, etniciteit, burgerlijke staat en psychische klachten) en kenmerken van de behandeling (zoals duur van de behandeling). Dus diverse kenmerken die mogelijk het succes van de behandeling zouden kunnen beïnvloeden. Bijvoorbeeld: werkt de interventie beter voor meisjes dan voor jongens, voor kinderen of voor adolescenten, voor allochtone of autochtone kinderen (Kazdin & Nock, 2003; Kraemer, Wilson, Fairburn & Agras, 2002)?

Een voorbeeld van een moderatoranalyse voor sekse en ernst van de problemen is het effectonderzoek naar het programma *Plezier op School* (zie hoofdstuk 9). Jongens rapporteerden dat *Plezier op School* effectief was in het verminderen van gepest worden, zeker wanneer de ernst van de problematiek op de voormeting bovengemiddeld was. Voor meisjes werd dit effect niet gevonden.



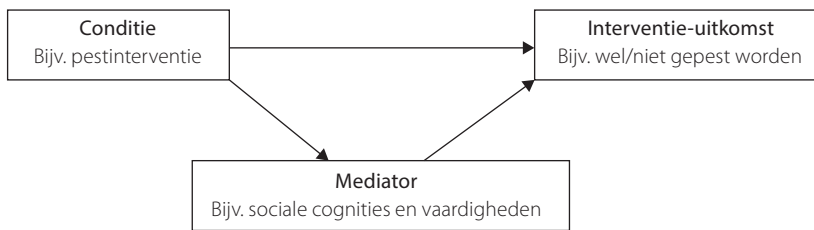
Figuur 6.2 Moderatoreffect

6.10 Mediatoren ('Waarom')

Naast het inzicht in voor wie de interventie werkt en wanneer de interventie het best werkt, is het ook belangrijk om inzicht te krijgen in het waarom. Waarom werkt deze specifieke interventie? Mediatoren geven inzicht in hoe of door welk

proces de interventie effect heeft. Ze vormen de verklarende schakel tussen de interventie en de behandeluitkomst. Vaak hangt de mediator samen met de onderliggende theorie van een interventie (zie paragraaf 6.6). Stel, een interventie voor pesten is gebaseerd op de sociale-informatieverwerkingstheorie (SIP; zie hoofdstuk 3). Dan zou de mediator kunnen zijn het veranderen van sociale cognities en het versterken van sociale vaardigheden (die op hun beurt weer het gepest worden beïnvloeden). Figuur 6.3 geeft dit weer. Om mediatoren te kunnen onderzoeken, zijn wel meerdere meetmomenten nodig. Indien de mediatoren goed in kaart zijn gebracht, is het mogelijk om de interventie zo aan te passen dat alleen de werkzame elementen overblijven.

In hoofdstuk 9 staat een voorbeeld van een analyse naar mediatoren voor het programma *Plezier op School*. Het verminderen van negatieve gedachten (zoals 'Ik denk dat ik iets verkeerd zal doen') blijkt belangrijk voor de afname van gepest worden. Helaas is nog weinig bekend over mediatoren in antipest-programma's. Espelage (2015) spreekt in dit verband van de noodzaak tot 'opening the black box'.



Figuur 6.3 Mediatoreffect

Algemeen en specifiek werkzame factoren

Als interventies werken, wil dat niet zeggen dat dit altijd komt door factoren ontleend aan de onderliggende theorie en de veronderstelde mediators. De behandelingen van de beroemde magnetiseur Mesmer zijn hier een voorbeeld van. Tegenwoordig vinden we Mesmers theorie onzin, maar zijn behandelingen bleken vaak wel te werken (zie kader 6.2). Hoe komt dat? Naast de interventie zelf (specifiek werkzame factoren) zijn er andere processen die van invloed kunnen zijn op de behandeluitkomst. Zo berekende Lambert (1992) dat de specifieke interventie slechts 15% van de effectiviteit verklaart. De andere processen zijn placebo-effect (zoals hoop en verwachting van de cliënt, 15%), cliënt- en omgevingsfactoren (levensgebeurtenissen, zoals verhuizing enzovoort) en algemeen werkzame factoren (30%). Er is echter veel kritiek op de berekening van Lambert (Van Yperen, 2010). Barnhoorn en collega's (2013) hebben in een recent literatuuronderzoek bekeken voor welke algemene werkzame factoren er empirisch bewijs is. Hierbij maakten ze onderscheid tussen cliëntfactoren, professionalfactoren en therapeutische alliantie. Van een aantal

cliëntfactoren is bekend dat ze zorgen voor een minder gunstig effect van de behandeling, zoals meer problemen, zwaardere problemen, een lage sociaal-economische status, afkomstig uit een groot gezin, en hogere leeftijd van het kind. Ook het behoren tot een etnische minderheid is een negatieve factor: het hulpaanbod is vaak minder geschikt. Vertrouwen en motivatie hebben juist een gunstig effect. Professionalfactoren die de behandeluitkomst positief beïnvloeden, zijn vermogen om ouders bij de interventie te betrekken, tijdige probleemherkenning, vertrouwen en geloof in de geschiktheid van de gekozen interventie, en eigenschappen zoals empathie, warmte en openheid. Ook een sterkere therapeutische alliantie (een affectieve band tussen hulpverlener en cliënt, en overeenstemming over de aanpak en de te behalen doelen) heeft een gunstig effect. Over het relatieve gewicht van de verschillende werkzame factoren konden de auteurs geen conclusie trekken; hierover bestaat nog veel discussie.



Voorbeeld

De magnetiseerbehandelingen van Mesmer

De Duitse arts Franz Anton Mesmer leefde in de tijd van de Verlichting en gaf een natuurwetenschappelijke verklaring voor zijn behandeling. Hij zag gezondheid als de vrije stroom van magnetische vloeistof – wat hij ‘dierlijk magnetisme’ noemde – door duizenden kanalen in het lichaam. Ziekte was het gevolg van belemmeringen in het dierlijk magnetisme. Door die belemmeringen op te heffen, werd de gezondheid hersteld. Als dat niet vanzelf gebeurde, kon de therapeut als geleider van dierlijk magnetisme optreden. Bij een individuele behandeling ging de therapeut tegenover de cliënt zitten en maakte hij contact via knieën en handen, terwijl hij de patiënt strak in de ogen keek. Bij een groepsbehandeling zaten de patiënten in een groot bad, gevuld met water, ijzervijsel en stukjes glas. Uit het bad staken ijzeren staven die het dierlijk magnetisme doorgaven. Soms waren alleen al oog- of handbewegingen van Mesmer voldoende om het dierlijk magnetisme over te brengen en de energiestroom weer op gang te brengen. Ook in Nederland kreeg het dierlijk magnetisme aanhangers. Dat kwam vooral door een demonstratie van de behandeling in Rotterdam. Proefpersoon was de 33-jarige Batje, die symptomen van hysterie vertoonde, zoals heftige stuiptrekkingen, krampen, flauwtes en rare lichaamshoudingen. De behandeling werd bijgewoond door een groot publiek, waaronder doktoren en hoge ambtenaren. Batje werd twee keer per dag gemagnetiseerd. In de loop van ongeveer een maand ging haar toestand aanmerkelijk vooruit. Ze slook haar dikke buik, die ‘als van een vijf maanden zwangere vrouw’ was geweest. Ze reageerde niet langer met tics, tremoren en lichte flauwtes op onverwachte prikkels. Als bewijs van haar volledige herstel loste zij zelf een pistoolschot, terwijl de therapeut boven haar hoofd op een trom roffelde. Ze doorstond deze beproeving lachend (Vijselaar, 2001, p. 312-313).

Samenvattend

Het onderzoek naar moderatoren geeft inzicht in voor wie en onder welke omstandigheid een interventie het best werkt. Mediatoren geven inzicht in hoe en waarom de interventie werkt. Naast de werking van de specifieke interventie zelf hebben placebo-effect, algemeen werkzame factoren en cliënt-/omgevingsfactoren invloed op behandeluitkomsten. De meningen zijn verdeeld over welk ingrediënt nu de behandeluitkomst het meest beïnvloedt. In effectiviteitsstudies is het daarom belangrijk om ook algemeen werkzame factoren te onderzoeken.

6.11 Kosteneffectiviteit

Ook geld kan een rol spelen bij de evaluatie van interventies. In een kosteneffectiviteitsstudie worden de financiële aspecten van een interventie gecombineerd met de effectiviteit van een interventie. Zo is bijvoorbeeld een interventie ten opzichte van een andere interventie kosteneffectief wanneer de interventie effectiever is en minder kost. Maar ook als een interventie ten opzichte van een andere interventie duurder is, kan deze toch kosteneffectief zijn, omdat de betere effectiviteit opweegt tegen de hogere kosten. Een studie naar kosteneffectiviteit evalueert en vergelijkt uitkomsten en kosten (investeringen en besparingen) van twee of meer interventies, waarbij de kosten zowel investeringen betreffen als mogelijke (toekomstige) besparingen. Voor pestprogramma's zijn kosteneffectiviteitsstudies nog erg weinig uitgevoerd (Tofsi & Farrington, 2011). Dit geldt trouwens voor de hele GGZ en jeugdzorg.

Wat betreft de kosten kan men denken aan de kosten van de interventie zelf, zoals de kosten van een hulpverlener. Daarnaast zijn er kosten voor het kind en zijn/haar ouders, zoals huishoudelijke hulp, (speciale) kinderopvang of eigen bijdragen. Maatschappelijke kosten als gevolg van werkverzuim of schoolverzuim kunnen ook worden meegenomen. Het berekenen van de effecten kan op twee manieren. De eerste manier is het gebruiken van dezelfde uitkomstmaten als in effectonderzoek (bijvoorbeeld wel/niet gepest worden). Een tweede, veelgebruikte manier is het berekenen van een specifieke effectmaat die rekening houdt met de kwaliteit van leven (Quality Adjusted Life Years; QALY). Hierin worden de gewonnen of verwachte levensjaren gecorrigeerd voor de kwaliteit van deze levensjaren. Door de totale kosten en effecten van verschillende interventies met elkaar te vergelijken, kunnen we bepalen of een interventie kosteneffectief is (Bodden, Dirksen & Bögels, 2009).

Het doel van kosteneffectiviteitsstudies is het wetenschappelijk ondersteunen van beleidskeuzes door de overheid, beleidsmedewerkers en verzekeraars over de besteding van financiële middelen. Stel, er is een bepaald budget van de overheid om pesten tegen te gaan. Dan is het zaak goed na te gaan welk programma hier het meest voor in aanmerking komt op basis van de kosten (zodat zo veel mogelijk scholen kunnen deelnemen) én op basis van effecten (zodat pesten effectief gereduceerd wordt). Een programma dat effectiever maar ook veel

duurder is, kan niet op alle scholen worden gegeven. Een wat minder effectieve interventie die goedkoper is, kan wel op alle scholen worden gegeven. Welke interventie verdient dan de voorkeur? Onderzoek naar kosteneffectiviteit kan hierop een antwoord geven.

Het enige programma in dit boek waarvoor de kosteneffectiviteit is berekend, is *Minder Boos en Opstandig* (zie hoofdstuk 9). Dit programma kwam kosteneffectiever uit de bus dan gebruikelijke zorg aan kinderen met gedragsstoornissen.

6.12 Erkenning van interventies door de Erkenningscommissie Interventies

In de voorgaande paragrafen hebben we beschreven waaraan goede interventies en goed effectonderzoek moeten voldoen. Deze aspecten zie je grotendeels (maar niet helemaal) terug in de criteria die de landelijke Erkenningscommissie Interventies hanteert bij de beoordeling van interventies. De commissie onderscheidt vier niveaus van erkenning, die oplopen in bewijskracht, namelijk 'goed onderbouwd', 'eerste aanwijzingen voor effectiviteit', 'goede aanwijzingen voor effectiviteit' en 'sterke aanwijzingen voor effectiviteit' (MOVISIE, NCJ, NISB, NJi & CGL, 2013).

Goed onderbouwd geeft aan dat de interventie zou kunnen werken. Er is echter nog geen effectonderzoek uitgevoerd. De interventie is duidelijk beschreven, toepasbaar in de praktijk en theoretisch onderbouwd (zie paragraaf 6.6). Er zijn drie niveaus van effectiviteit. Daarbij gaat het om de overtuigingskracht van het interventieonderzoek. *Eerste aanwijzingen voor effectiviteit* geeft aan dat de opzet van het empirisch onderzoek minstens in een zwakke causale bewijskracht voorziet. Er is een voormeting en een nameting, maar geen controleconditie. Bij *goede aanwijzingen voor effectiviteit* voorziet de opzet van het empirisch onderzoek in minstens een matige causale bewijskracht. Het onderzoek heeft een (quasi-)experimentele of een andere opzet (bijvoorbeeld herhaalde casestudies). De studies zijn niet noodzakelijk in de praktijk uitgevoerd of hebben nog geen follow-up. Op het niveau van *sterke aanwijzingen voor effectiviteit* is er voldoende onderzoek van hoge kwaliteit om aan te mogen nemen dat de interventie bij de doelgroep effectief is en beter werkt dan de gebruikelijke zorg én dat dit toe te schrijven is aan de interventie (sterke bewijskracht). Meestal gaat het hier om gecontroleerd onderzoek, zoals RCT's en quasi-experimentele studies. Op het moment van dit schrijven zijn slechts enkele interventies voor kinderen en jeugdigen als *effectief volgens sterke aanwijzingen* beoordeeld. Indien de beoordeling rond is, wordt de interventie opgenomen in de databank Effectieve Jeugdinterventies (www.nji.nl/nl/Databanken/Databank-Effectieve-Jeugdinterventies).

Het hoogste niveau, *kosteneffectief*, werd eerder wel genoemd in de zogenoemde effectiviteitsladder van Van Yperen en Veerman (2008), maar maakt niet langer deel uit van de erkenningsniveaus.

6.13 Samenvatting en advies

Interventies hebben als doel probleemgedrag te verlagen en/of competenties te verhogen. Interventies onderscheiden zich in het moment waarop ze ingrijpen: van promotie tot voortgezette zorg. Opvallend is dat we slechts van een klein deel van alle interventies weten of ze werken. Dat komt doordat slechts weinig interventies zijn onderzocht op effectiviteit. Een beperkt aantal daarvan is geëvalueerd aan de hand van een onderzoeksdesign dat conclusies over effectiviteit toelaat. Hiervoor bestaan uiteenlopende redenen, zoals gebrek aan geld, tijd en expertise.

Dit hoofdstuk is een inleiding op de volgende hoofdstukken over specifieke interventies. We hopen dat we u als lezer hebben gestimuleerd om kritisch naar interventies te kijken en te letten op de kwaliteit van interventies. Dit kan door bij een interventie na te gaan of die voldoet aan vastgestelde eisen: is de interventie duidelijk beschreven, is zij toepasbaar in de praktijk en is zij theoretisch goed onderbouwd? Dit kan ook door de (onderzoeks)literatuur te lezen over interventies. Misschien is de interventie al beoordeeld door de Erkenningscommissie en opgenomen in de databank. De databank is echter niet volledig en up-to-date. Daarom is het goed deze vuistregel te hanteren: hoe hoger de interne en externe validiteit van het onderzoek, des te groter de mogelijke bewijskracht (van potentie tot werkzaam). Een randomized controlled trial (RCT) uit de klinische praktijk is de gouden standaard. Het uitvoeren van een RCT kost echter veel tijd en moeite en kan daardoor niet zomaar door iedereen worden uitgevoerd. Indien u zelf cliënten behandelt, zou u kunnen onderzoeken of de interventie die u toepast veranderingen bij uw cliënten laat zien. In het algemeen is het aan te raden om meerdere uitkomstmaten, meetmomenten en informanten te gebruiken binnen effectonderzoek. Tot slot is het belangrijk aandacht te besteden aan *klinisch* significante en betrouwbare veranderingen (RCI) en aan moderatoren, mediators en kosteneffectiviteit. Dankzij effectonderzoek kunnen we uiteindelijk alle cliënten een interventie aanbieden waarvan we weten dat die écht werkt!

6.14 Literatuur

- Asscher, J.J., Deković, M., Prinzie, P., Hermanns, J.M.A., & Akker, A.L. van den (2009). De betekenis van veranderingen in gezinnen die hebben deelgenomen aan het Home-start programma en voorspellers van deze veranderingen. *Pedagogiek*, 29, 247-269.
- Barlow, D.H. (2010). Negative effects from psychological treatments. *American Psychologist*, 65, 13-19.
- Barnhoorn, J., Broeren, S., Distelbrink, M., Greef, M. de, Grieken, A. van, Jansen, W., Pels, T., Pijnenburg, H., & Raat, H. (2013). *Cliënt-, professional- en alliantiefactoren: hun relatie met het effect van zorg voor jeugd*. Den Haag: ZonMW.
- Bodden, D.H.M., Dirksen, C.D., & Bögels, S.M. (2009). Effectiviteit en kosteneffectiviteit bij klinisch angstige kinderen. In T.A.M. Graas, T. Liefwaard, C. Schuengel, W. Slot &

- H. Stegge (Red.), *Praktijkboek Jeugdzorg: Van de Wet op de Jeugdzorg naar de praktijk van nu*. Houten: Bohn Stafleu Van Loghum.
- Cook, T.D. (2003). Why have educational evaluators chosen not to do randomized experiments? *The ANNALS of the American Academy of Political and Social Science*, 589, 114-149.
- Deković, M. (2010). Effecten van interventies: Baat het niet, dan schaadt het niet? *Kind en Adolescent*, 31, 98-101.
- Deursen, D.S. van, Saleminck, E., Lammers, J., & Wiers, R.W. (2010). Selectieve en geïndiceerde preventie van problematisch middelengebruik bij jongeren. *Kind en Adolescent*, 31, 234-246.
- Dishion, T.J., & Andrews, D.W. (1995). Preventing escalations in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology*, 63, 538-548.
- Dishion, T.J., McCord, J., & Poulin, F. (1999). When interventions harm. Peer groups and problem behavior. *American Psychologist*, 54, 755-764.
- Durlak, J.A. (2009). How to select, calculate and interpret effect sizes. *Journal of Pediatric Psychology*, 34, 917-928.
- Espelage, D. L. (2015). Taking peer victimization to the next level: Complex interactions among genes, teacher attitudes/behaviors, peer ecologies, & classroom characteristics. *Journal of Abnormal Child Psychology*, 43, 77-80.
- European Platform for Investing in Children (EPIC) (2014, 22 september). Verkregen via http://techcrunch.comhttp://europa.eu/epic/about/index_en.htm.
- Farrington, D.P. (2003). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Sciences*, 587, 49-69.
- Fekkes, M., & Paulussen, T. (2012). PRIMA – een schoolbreed antipestprogramma voor basisscholen. In: F. Goossens, M. Vermande & M. van der Meulen (Red.), *Pesten op school. Achtergronden en interventies* (pp. 146-151).
- Fekkes, M., Pijpers, F.I., & Verloove-Vanhorick, S.P. (2006). Effects of an anti-bullying school policy on bullying and health symptoms. *Archives of Pediatrics and Adolescent Medicine*, 160, 638-644.
- Gillham, J.E., Reivich, K.J., Jaycox, L.H., & Seligman, M.E.P. (1995). Preventing depressive symptoms in school children: Two year follow-up. *Psychological Science*, 6, 343-351.
- Gresham, F.M., & Lochman, J.E. (2008). Methodological issues in research using cognitive-behavioral interventions. In M.J. Mayer, R. van Acker, J.E. Lochman & F.M. Gresham (Eds.), *Cognitive-behavioral interventions for emotional and behavioral disorders: School-based practice* (pp. 82-108). New York: Guilford Press.
- Hawley, D.R. (1995). Assessing change with preventive interventions: The reliable change index. *Family Relations*, 44, 278-284.
- Ince, D., & Berg, G. van den (2010). *Overzichtsstudie interventies voor migrantenjeugd: Ontwikkelingsstimulering, preventie en vroeghulp*. Utrecht: Nederlands Jeugdinstituut.
- Jacobson, N.S., Follette, W.C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336-352.
- Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- Kazdin, A.E. (2003). Psychotherapy for children and adolescents. *Annual Review of Psychology*, 54, 253-276.
- Kazdin, A.E., & Nock, M.K. (2003). Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44, 1116-1129.

- Kelly, M., Morgan, A., Ellis, S., Younger, T., Huntley, J., & Swann, C. (2010). Evidence based public health: A review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Social Science and Medicine*, 71, 1056-1062.
- Kendall, P.C. (1999). Clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 283-284.
- Kraemer, H.C., Wilson, G.T., Fairburn, C.G., & Agras, W.S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59, 877-883.
- Kutash, K., Duchnowski, A.J., & Lynn, N. (2006). *School-based mental health: An empirical guide for decision-makers*. Tampa, FL: University of South Florida.
- Lambert, M.J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists: Handbook of psychotherapy integration. In J.C. Norcross & M.R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 94-129). New York, NY: Basic Books.
- Lange, M. de (2013, 16 december). Evidence werken een verstandige eis? [Web log post]. Verkregen via <http://kennisnetjeugd.nl/blog/65-evidence-based-werken-een-verstandige-eis>
- Loon, D. van, Meulen, B.F. van der, & Minnaert, A.E.M.G. (2011). *Effectonderzoek in de gedragswetenschappen. Methodologische moeilijkheden en mogelijkheden*. Den Haag: Boom Lemma.
- Matthys, W., & Lochman, J.E. (2010). *Oppositional defiant disorder and conduct disorder in childhood*. Hoboken, NJ: Wiley.
- MOVISIE, Nederlands Centrum Jeugdgezondheid (NCJ), Nederlands Instituut voor Sport en Beweging (NISB), Nederlands Jeugdinstituut (Nji), & RIVM Centrum Gezond Leven (CGL) (2013). *Erkenningstraject interventies. Criteria voor gezamenlijke kwaliteitsbeoordeling 2013-2018*. Bilthoven/Ede/Utrecht: Auteurs.
- Mrazek P.J., & Haggerty, R.J. (Eds.) (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academies Press.
- O'Connell, M.E., Boat, T.F., & Warner, K.E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.
- Olweus, D. (1991). Bully/victim problems among school children: Basic facts and effects of a school-based intervention program. In D.J. Pepler & K.H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 411-448). Hillsdale, NJ: Erlbaum.
- Orobio de Castro, B. (2007). *Woede, Wraak & Leedvermaak: Op zoek naar drijvende krachten achter de ontwikkeling van gedragsproblemen*. Rede verkort uitgesproken bij de aanvaarding van het ambt van hoogleraar op het gebied van de Experimentele Ontwikkelingspsychopathologie aan de faculteit Sociale Wetenschappen van de Universiteit Utrecht.
- Trimbos-instituut (2014). Onlinehulp stempel. Verkregen via www.onlinehulpstempel.nl.
- Ttofi, M.M., & Farrington, D.P. (2011). Effectiveness of school-based programs to reduce bullying: a systematic and meta-analytic review. *Journal of Experimental Criminology*, 7, 27-56.
- Vijselaar, J. (2001). *De magnetische geest. Het dierlijk magnetisme 1770-1830*. Nijmegen: SUN.
- Weiss, B., Catron, T., Harris, V., & Phung, T.M. (1999). The effectiveness of traditional child psychotherapy. *Journal of Consulting and Clinical Psychology*, 67, 82-94.
- Weisz, J.R. (2004). *Psychotherapy for children and adolescents: Evidence-based treatments and case examples*. New York, NY: Cambridge University Press.

- Weisz, J.R., Donenberg, G.R., Han, S.S., & Weiss, B. (1995). Bridging the gap between laboratory and clinic in child and adolescent psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 688-701.
- Weisz, J.R., Jensen-Doss, A., & Hawley, K.M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61*, 671-689.
- Weisz, J.R., Jensen, A.L., & McLeod, B.D. (2005). Development and dissemination of child and adolescent psychotherapies: Milestones, methods, and a new deployment-focused model. In E.D. Hibbs & P.S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically-based approaches* (2nd ed.) (pp. 9-39). Washington, DC: American Psychological Association.
- Weisz, J.R., Sandler, I.N., Durlak, J.A., & Anton, B.A. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist, 60*, 628-648.
- Yirmiya, N. (2010). Editorial: Early prevention and intervention – the five W (and one H) questions. *Journal of Child Psychology and Psychiatry, 51*, 1297-1299.
- Yperen, T. van (2007). *Integraal erkend. Naar een afstemming erkenning jeugdinterventies*. Utrecht: Nederlands Jeugdinstituut.
- Yperen, T. van (Red.) (2010). *55 Vragen over effectiviteit. Antwoorden voor de jeugdzorg*. Utrecht: Nederlands Jeugdinstituut.
- Yperen, T.A. van, & Veerman, J.W. (2008). *Zicht op Effectiviteit. Handboek voor praktijk-gestuurd effectonderzoek in de jeugdzorg*. Delft: Eburon.