

COLUMN

De schrijver van de 'column' kan een wetenschappelijk expert of een klinisch deskundige zijn, maar kan ook een gedragswetenschapper, een groepsleider, een manager en/of een ouder-begeleider zijn. Wat zij gemeen hebben is hun ervaring met (onderzoek bij) jeugdigen met een lichte verstandelijke beperking. In deze Onderzoek & Praktijk geven wij het woord aan Robert Didden, verbonden aan de Radboud Universiteit Nijmegen.

EFFECTONDERZOEK: OP WEG NAAR EVIDENCE-BASED PRACTICE

*Robert Didden*¹¹

Verleden jaar hebben de VOBC en De Borg een symposium georganiseerd waarmee het onderwerp *Behandeling van gedrags- en psychiatrische stoornissen bij jeugdigen en volwassenen met een lichte verstandelijke beperking* vol in de aandacht van klinici en onderzoekers werd geplaatst (zie Didden & Moonen, 2007). Inmiddels is met behulp van ZONMW subsidie en eigen middelen een voor Nederlandse begrippen uniek en grootschalig onderzoek ingezet. Onder meer zal er onderzocht worden óf, en zo ja, hoe effectief de behandeling van dergelijke stoornissen bij deze doelgroep is. Niet alleen wordt het effect van kortdurende therapievormen onderzocht, waaronder cognitief gedragstherapeutisch programma's gericht op agressief gedrag bij jeugdigen en volwassenen. Ook zal het effect van een meer langdurende klinische behandeling bij volwassenen worden onderzocht.

Voor Nederlandse begrippen is dit een uniek onderzoeksprogramma, dat voor een deel wordt uitgevoerd binnen een onlangs opgericht consortium. De verwachtingen zijn hoog gespannen. Het doen van effectonderzoek is echter lastig en kent vele valkuilen en beperkingen. In deze bijdrage worden enkele van deze beperkingen kort samengevat. Effectonderzoek mag zich verheugen in een toenemende belangstelling in verschillende sectoren van de hulpverlening (jeugdzorg, gehandicaptenzorg, etc.). Het aantal studies in de wetenschappelijke literatuur bij jeugdigen en volwassenen met een lichte verstandelijke beperking is de laatste 15 jaar aanzienlijk toegenomen, vooral gepubliceerd in buitenlandse (Amerika, UK) tijdschriften. Oppervlakkige lezing van deze studies en meta-analyses doet vermoeden dat gedragsproblemen succesvol verminderd kunnen worden en dat interventies evidence-based zijn. Maar een meer nauwkeurige en kritische bestudering van deze studies moet ons meer bescheiden stemmen. Veel studies zijn kwalitatief onvoldoende. Ik zal mij in deze bijdrage beperken tot niet-farmacologische studies. De kritiek op farmacologische studies mag als bekend worden verondersteld. Zeer recent hebben de Britse psychiater Deb en zijn collega's wetenschappelijke studies naar het effect van verschillende soorten psychofarmaca op gedragsproblemen bij volwassenen over de knie gelegd. De meeste studies op dit gebied zijn onvoldoende gecontroleerd (i.c. gebrekkige interne validiteit) en de 'evidence-base' van verschillende en veel gebruikte typen psychofarmaca, waaronder

¹¹ Dr. Robert Didden is als universitair hoofddocent werkzaam bij de vakgroep Orthopedagogiek van de Radboud Universiteit Nijmegen, is tevens hoofd van het Trajectum Kenniscentrum en is als GZ-psycholoog werkzaam bij Trajectum-Hanzeborg, een SGLVG-centrum te Zutphen. Correspondentieadres: Vakgroep Orthopedagogiek, Radboud Universiteit Nijmegen, Postbus 9104, 6500 HE Nijmegen, email: r.didden@pwo.ru.nl.

bijvoorbeeld antipsychotica (Deb et al., 2007) en antidepressiva (Sohanpal et al., 2007), is nog steeds mager.

De term therapie kan men breed opvatten. Meestal wordt hieronder verstaan een in de tijd en plaats (therapieruimte) afgebakende en systematisch uitgevoerde en op een theorie gebaseerde professionele inspanning (hulpverlening), die erop is gericht symptomen van een psychiatrische stoornis en/of gedragsprobleem te verminderen. Maar ook binnen zogenaamde sociotherapeutische of therapeutische milieus en woongroepen binnen instellingen kan een behandeling worden uitgevoerd. Ik zal me hier beperken tot eerstgenoemde vormen van therapie.

‘Evidence-based practice’ veronderstelt dat de kwaliteit van het onderzoek goed is en dat een studie aan een aantal belangrijke wetenschappelijke eisen voldoet. Echter, als men de studies op dit terrein nauwkeuriger analyseert moet men vaststellen dat de meeste ervan serieuze tekortkomingen hebben. Neem bijvoorbeeld een recente studie van McCabe et al. (2006), gepubliceerd in *Journal of Intellectual Disability Research*, een Engelstalig gerenommeerd tijdschrift. In hun studie kreeg een groep volwassenen met depressieve symptomen een therapie die bestond uit cognitieve en gedragstherapeutische elementen. Er was een (wachtlíjst?) controlegroep die geen therapie kreeg. De therapie leidde bij mensen uit de experimentele groep tot een vermindering van depressieve klachten, tezamen met een verhoging van de eigenwaarde en een vermindering van automatische negatieve gedachten. De auteurs concluderen dat deze therapie effectief is bij de behandeling van depressieve symptomen bij volwassenen met een lichte verstandelijke beperking.

Dit is mijns inziens onterecht. Er zijn een aantal tekortkomingen van deze studie aan te wijzen die de validiteit van de conclusies beperken. Deze tekortkomingen zijn min of meer exemplarisch voor veel interventiestudies (inclusief de farmacologische) die bij mensen met een lichte verstandelijke beperking kunnen worden gevonden. Een van de belangrijkste tekortkomingen is wel het ontbreken van een placeboconditie. Dat wil zeggen dat er een controlegroep is die niet de therapie kreeg maar een netherapie. Hiermee kan het effect van een niet te onderschatten buitenexperimentele factor worden onderzocht respectievelijk buitengesloten, namelijk het effect van aandacht of een andere generieke factor (bijvoorbeeld een therapeutgerelateerde en/of andere stressverlagende factor) die verantwoordelijk is voor een verbetering in het welzijn van de deelnemers. Het opnemen van een netherapie in zo’n studie is waardevoller dan een controlegroep die geen therapie krijgt. Bovendien vermelden McCabe et al. dat in de experimentele groep maar werd gestreefd om een situatie te creëren ‘where group members felt comfortable sharing their personal concerns with others and where they were able to learn skills to improve their quality of life’ (p. 241). Dus zowel generieke factoren binnen de experimentele groep alsmede het ontbreken van dergelijke factoren binnen de controlegroep, zorgen ervoor dat McCabe et al. niet zonder meer kunnen stellen dat de therapie derhalve effectief was.

In de klinische praktijk hebben generieke factoren niet zelden een belangrijke invloed op het verloop van klachten. Dit geldt vooral bij internaliserende problematiek, zoals beschreven in de McCabe et al. studie. Of dit ook geldt voor externaliserende gedrag, zoals bijvoorbeeld agressief gedrag, is de vraag. Hier lijkt de invloed van buitenexperimentele en generieke factoren kleiner te zijn. Maar op voorhand mag niet worden uitgesloten dat ook externaliserend gedrag gevoelig is voor generieke factoren. Het opnemen van een netherapie in een experimentele studie is uiteraard niet vrij van praktische en ethische problemen. Hiermee zijn voorts ook extra manuren en kosten gemoeid. Maar gelet op het

belang van effectonderzoek en de implicaties van een gunstige uitkomst ervan, verdient het aanbeveling netherapieën in wetenschappelijk onderzoek op te nemen.

Een tweede opmerking naar aanleiding van het bestuderen van de McCabe et al. studie heeft betrekking op het design waarin de gegevens werden verzameld. Therapieën die dicht in de buurt komen van wat we evidence-based practice noemen, zijn toegepaste gedragsanalyse en cognitieve gedragstherapieën. Toegepaste gedragsanalyse slaat op het proces waarin door nauwkeurige observatie de uitlokkende en instandhoudende factoren van probleemgedrag worden geïdentificeerd, waarna een interventie wordt ontworpen die bestaat uit (a) het wijzigen van deze factoren en (b) het aanleren van alternatief sociaal competent gedrag bij de cliënt. Eigen meta-analytisch onderzoek laat zien dat deze benadering effectief is, vooral bij jeugdigen die in een institutionele setting verblijven (Didden et al., 2006). Maar ook bij deze studies kunnen kritische opmerkingen worden geplaatst. Uit deze meta-analyse bleek onder andere dat de meeste studies onvoldoende gecontroleerde $N=1$ designs gebruikten, zoals het AB design, waarbij na een periode van 'niets doen' (basislijn) en periode van interventie volgt. Meer intern valide $N=1$ designs, zoals bijvoorbeeld het meervoudige basislijn design, het 'alternating treatments design' en het 'reversal of ABAB design', waren in de minderheid. Deze teleurstellende bevinding is opvallend aangezien dergelijke designs goed zijn toe te passen in de klinische praktijk. Bij het 'reversal design' bijvoorbeeld verzamelt de onderzoeker/clinicus gegevens gedurende een langere periode (bv. registratie van het aantal agressieve incidenten) en past hij na een periode van 'niets doen' een interventie toe, waarna hij na verloop van tijd de interventie weer stopt en gegevens verzamelt gedurende een periode van 'niets doen'. De cliënt is dan als het ware zijn eigen controle en verschillen in agressieve incidenten tussen perioden van 'niets doen' en interventie bij deze cliënt duiden dan op een effect van behandeling dat moeilijk kan worden toegeschreven aan andere, toevallige factoren. Uiteraard is de generaliseerbaarheid van de uitkomsten van $N=1$ studies beperkt, maar dit is dan ook niet het oogmerk van $N=1$ studies. Generaliseerbaarheid staat bij deze designs niet voorop. In het licht van de heterogeniteit van de doelgroep is dit misschien ook niet te realiseren. Een voordeel van een $N=1$ design is dat de behandeling en dataverzameling is afgestemd op de individuele cliënt en diens unieke situatie. Een goede bespreking van $N=1$ designs voor de klinische praktijk wordt gegeven door Kazdin (2002).

Helaas worden $N=1$ designs door sommige onderzoekers (zie o.a. Van Nieuwenhuijzen et al., 2006) nog steeds beoordeeld als zijnde inferieur aan het pretest posttest controlegroep design. Deze onderzoekers verzuimen te vermelden dat ook groepsdesigns met controlegroepen tekortkomingen hebben. Naast het fenomeen 'regressie naar het gemiddelde' (tendens van extreme scores om naar het gemiddelde te kruipen bij herhaalde metingen) hebben we het over het ontbreken van 'random sampling' (willekeurige steekproef uit de populatie), 'random assignment' (deelnemers willekeurig toewijzen aan experimentele- en controlegroep), voldoende statistische power (ten gevolge van kleine onderzoeksgroepen) en 'procedurele betrouwbaarheid', waarbij getoetst wordt of de therapeut zich ook houdt aan het protocol of gemaakte afspraken, wordt in verreweg de meeste effectstudies geen gegevens verzameld. Dit is belangrijk als men streeft naar een geprotocolleerde en gestandaardiseerde therapie en wanneer men erachter wil komen welke elementen van een therapie een gunstig resultaat voorspellen.

Groepsdesigns worden vaak toegepast bij cognitieve gedragstherapieën. Vooral de zogenaamde '*anger management*' therapieën zijn populair. Uitkomsten van gecontroleerde studies met deze therapieën zijn doorgaans gunstig. (Overigens, de kans dat studies met ongunstige uitkomsten van effectonderzoek worden gepubliceerd, is zeer klein.) Voorts zijn

deze therapieën redelijk goed onderbouwd en steunen zij op onafhankelijk theoretisch onderzoek. Tot op heden zijn internationaal enkele tientallen studies gepubliceerd, waarbij gebruikt wordt gemaakt van pretest posttest controlegroep designs. Ofschoon een kwantitatieve meta-analyse van dergelijke therapieën nog niet is gepubliceerd, mag men op basis van bestudering van de afzonderlijke studies concluderen dat zo'n benadering effectief is, vooral ten behoeve van de behandeling van agressief gedrag. Echter, net als bij de studie van McGabe et al. voor internaliserende problemen zijn er belangrijke kanttekeningen te plaatsen bij dergelijke studies (zie ook vorige alinea). Dit betreft onder andere de problematiek van de statistische significantie op basis waarvan wordt geconcludeerd dat een therapie effectief is. Probleem is, dat de conclusie is gebaseerd op verschillen tussen groepen in gemiddelde scores en niet zelden zien we dat de verschillen relatief klein (ofschoon significant) zijn. Dit impliceert dat bij sommige cliënten de therapie heeft geleid tot een verergering van de problematiek en/of dat er bij sommige cliënten niets is veranderd, terwijl de onderzoekers op basis van significantie desalniettemin concludeert dat de therapie effectief is. Bovendien is dit verschil niet zelden zó klein dat men 'in het dagelijks leven' geen noemenswaardige verbetering in het gedrag van de cliënt kan waarnemen. Dit is de reeds lang bestaande controverse aangaande het verschil tussen statistische en klinische significantie. Doorgaans wordt, ten onrechte, de conclusie over het gunstig effect van een therapie gebaseerd op het eerste. En als er al een behoorlijk gemiddeld gunstig effect wordt gevonden, blijft de vraag bestaan welke factoren hebben bijgedragen tot dit resultaat, meestal onbeantwoord. Niet zozeer óf een therapie effectief is, maar welke cliënt- of therapiekenmerken die effectiviteit voorspellen, is een interessante vraag. Het zal nog vele jaren van gespecialiseerd en goed gecontroleerd onderzoek vergen om deze vraag te kunnen beantwoorden.

Een laatste belangrijke tekortkoming die we in de McGabe et al. studie menen vast te stellen, is een 'bias' of vertekening bij het invullen van de vragenlijsten waarmee de psychische klachten worden gemeten. Het belangrijkste bezwaar is dat alle personen wisten in welke conditie zij zaten en dit heeft zonder twijfel een belangrijke invloed op de manier waarop de antwoorden op de vragenlijsten worden gescoord. Het lijkt me dat dit voor mensen met een lichte verstandelijke beperking zelfs sterker geldt dan voor mensen zonder dergelijke beperkingen. Er zijn vraagstekens te zetten bij het veelvuldig en louter gebruik van vragenlijsten bij effectonderzoek. Er is altijd wel iets mis met vragenlijsten. Doorgaans is de interne consistentie ervan wel in orde, maar als we kijken naar de interrater betrouwbaarheid zien we in veel gevallen dat deze onvoldoende is. Met andere woorden: twee of meer mensen zijn het er niet over eens hoe ernstig een probleem is of hoe vaak een probleem zich voordoet en óf een probleem zich wel voordoet. Een goede convergente validiteit (overeenstemming in scores tussen twee schalen die hetzelfde pretenderen te meten) heft dit probleem niet op. Of een schaal ook meet wat het wil meten, blijft vaak onduidelijk. Een alternatief voor vragenlijsten, namelijk meer 'objectieve' maten, is moeilijk te realiseren. Zeker als er veel participanten zijn, ligt het alternatief niet voor het oprapen. Ook meer 'objectieve' methoden, zoals het via de computer registreren van incidenten, zijn niet vrij van problemen.

Nieuw ontwikkelde methoden als de *Impliciete Associatie Test (IAT)*, die de laatste jaren worden gebruikt bij het theoretisch onderzoek naar aandachtsprocessen bij normaal begaafde personen, vormen in de toekomst mogelijk een aanvulling op het arsenaal van meetinstrumenten bij effectonderzoek. Kort gezegd komt het bij een IAT erop neer dat onbewuste attitudes en processen bij een cliënt gemeten kunnen worden door middel van eenvoudige computertaakjes (voor een beschrijving van de stand van zaken van het wetenschappelijk onderzoek op dit terrein, zie Wiers, 2007; zie ook www.implicit.eu).

Uitkomsten van een dergelijke test kunnen mogelijk duiden op verstoringen of afwijkingen in attitudes die mogelijk samenhangen met psychologische problematiek, bijvoorbeeld bij alcoholmisbruik en verslaving. Het is denkbaar dat dergelijke technieken kunnen worden ingezet bij de meting van attitudes ten aanzien van seksueel delictgedrag en agressief gedrag. Recente ontwikkelingen in het effectonderzoek zijn voorts het bepalen van cortisolspiegels en hartslagvariabiliteit ten behoeve van stressmetingen en het meten van endogene melatoninespiegels bij het opsporen van verstoringen in het circadiane slaapwaakritme. Deze ontwikkelingen kunnen in de toekomst bijdragen aan het verbeteren van de kwaliteit van effectonderzoek bij mensen met een (lichte) verstandelijke beperking. Eerste stappen op deze terreinen zijn reeds gezet.

We concluderen dat onze huidige praktijk van hulpverlening nog ver verwijderd is van wat we 'evidence-based practice' noemen. Maar met het door de VOBC en De Borg ingezette onderzoek zal een grote stap in deze richting worden gedaan ten behoeve van een kwalitatief goede behandeling en begeleiding van jeugdigen en volwassenen met een lichte verstandelijke beperking.

Literatuur

- Deb, S., Sohanpal, S., Soni, R., Lenotre, L., & Unwin, G. (2007). The effectiveness of antipsychotic medication in the management of behaviour problems in adults with intellectual disabilities. *Journal of Intellectual Disability Research, 51*, 766-777.
- Diden, R., Korzilius, H., Van Oorsouw, W., & Sturmey, P. (2006). Behavioral treatment of challenging behavior in individuals with mild mental retardation: A meta-analysis of single subject research. *American Journal on Mental Retardation, 111*, 290-298.
- Diden, R., & Moonen, X. (2007). *Met het oog op behandeling. Effectieve behandeling van gedragsstoornissen bij mensen met een lichte verstandelijke beperking*. Utrecht: VOBC en De Borg.
- Kazdin, A. (2002). *Research designs in clinical psychology* (4e druk). London: Allyn & Bacon.
- McCabe, M., McGillivray, J., & Newton, D. (2006). Effectiveness of treatment programmes for depression among adults with mild/moderate intellectual disability. *Journal of Intellectual Disability Research, 50*, 239-247.
- Sohanpal, S., Deb, S., Thomas, C., Soni, R., Lenotre, L., & Unwin, G. (2007). The effectiveness of antidepressant medication in the management of behaviour problems in adults with intellectual disabilities: A systematic review. *Journal of Intellectual Disability Research, 51*, 750-765.
- Van Nieuwenhuijzen, M., Orobio de Castro, B., & Matthys, W. (2006). Problematiek en behandeling van LVG jeugdigen: Een literatuurreview. *Nederlands Tijdschrift voor de Zorg aan mensen met verstandelijke beperkingen, 32*, 211-229.
- Wiers, R. (2007). *Slaaf van het onbewuste. Over emotie, bewustzijn en verslaving*. Amsterdam: Uitgeverij Bert Bakker.