

# Het *evidence*beest heeft kuren

## Gebruik en misbruik van EBP, RCT- en EST-methodologie<sup>i</sup>

*Ik kan het niet vaak genoeg herhalen: ons doel is niet om de resultaten van RCT's te verguizen, maar om erop te wijzen dat ze wel eens niet zo onherroepelijk zouden kunnen zijn als hun pleitbezorgers claimen.*  
Westen, Novotny en Thompson-Brenner (2005, p. 430)

*Ik ben het er helemaal mee eens dat we behoefte hebben aan méér wetenschap, méér grondig en nauwkeurig gewogen bewijsmateriaal – meer, niet minder. (...) Maar we moeten wél onderscheid maken (...) tussen wetenschap en propaganda pro domo.*  
Paul Wachtel (SEPI-Forum, 2004)

### Aan- en inleiding

Eind vorig jaar ontving het Tijdschrift voor Psychotherapie kort na elkaar twee keer post van de Vereniging voor Gedragstherapie en Cognitieve therapie (VGCT). Eerst een mapje met folders over 'Gedragstherapie en Cognitieve Therapie' en vervolgens een persbericht. In beide gevallen werd nogal opgescheept over cognitieve gedragstherapie. De brief bij de folders (VGCT, 2005a) bevatte ronkende zinnen als 'De leden van de VGCT zijn namelijk gespecialiseerd in de meest effectieve vorm van psychotherapie ter wereld: de cognitieve gedragstherapie' en 'De cognitieve gedragstherapie is het enige psychotherapiespecialisme waarvan al vaak is – en telkens weer – wordt (*sic*<sup>ii</sup>) aangetoond dat het werkt'. In het persbericht (VGCT, 2005b) werd de lezer na een paar zakelijke mededelingen ook weer – en hier nog minder relevant – getraakteerd op de volgende propaganda: 'Bij de VGCT zijn beoefenaren van de gedragstherapie en de cognitieve therapie aangesloten. Deze specialismen binnen de psychotherapie zijn *evidence-based* en worden algemeen, ook internationaal, beschouwd als de meest effectieve vormen van behandeling binnen de geestelijke gezondheidszorg.'<sup>iii</sup>

Ergernis hierover bracht me ertoe me eens in *evidence-based practice* (hierna: EBP) te verdiepen, in het bijzonder in de zin en onzin van *randomized clinical trials* (hierna: RCT's). Voor alle zekerheid: ik ben een voorstander van EBP. Dat alle hulpverleners overtuigd zijn van de validiteit en effectiviteit van hun eigen werkwijze (of het nu homeopathie, astrologie, iriscopie of een van de vele psychotherapieën betreft) is voor mij al voldoende reden te menen dat een wetenschappelijke toets van al die behandelvormen bepaald geen overbodige luxe is.

Dit stuk zal soms een kritiek op de cognitieve gedragstherapie (hierna: CGT) lijken. Dat is het niet. Wat hier gekritiseerd wordt, is het geloof in RCT's (juister: de selectieprocedure van *empirically supported therapies* oftewel de 'EST-methodologie') als de onfeilbare scherperechter in de brandende kwestie 'to be or not to be empirically supported'. Want wat de doorslag blijkt te geven zijn niet de 'objectieve' onderzoeksgegevens, maar de menselijke zwakheid om de onderzoeksresultaten naar eigen hand te zetten.

De psychotherapie heeft wel wat te danken aan de (al dan niet cognitieve) gedragstherapie. Deze laatste stroming heeft er welhaast in haar eentje voor gezorgd dat psychotherapie zich heeft bewezen in de concurrentiestrijd met psychofarmaca. Het zou dan ook flauw zijn de CGT en haar vertegenwoordigers nu te verwijten dat zij zo ijverig zijn geweest de effectiviteit van de eigen therapie te onderzoeken voordat de andere scholen daar zin in hadden. Toch is dit maar één kant van het verhaal: aan 'bewezen effectiviteit' zitten allerlei mitsen & maren en haken & ogen, voordat zij ook *echt* bewezen is.

Ik begin met een korte geschiedenis van RCT's aan de hand van Healy (1997 en 2002). Daarna volgt een relaas over het gebruik en vooral het misbruik van dit onderzoeksdesign. Daar is in dit tijdschrift al eerder over geschreven, maar altijd vanuit het perspectief van de bovenliggende (zie Jaspers, 2000; Ten Broeke, Korrelboom & Meijer, 2000; Albersnagel, 2004; Van Balkom, Van Vliet & Spinhoven, 2004) of de onderliggende partij (Lietaer, 1995; Gomperts, 1999; Takens, 2004a, 2004b). Zelf voel ik me op geen enkele manier gebonden aan een van de psychotherapeutische scholen, wat uiteraard niet betekent dat ik een 'neutraal' of 'objectief' standpunt inneem.

Ik zal me vooral concentreren op de recente literatuur en discussie in het Angelsaksische taalgebied, want het debat in de Verenigde Staten bepaalt meestal wat er bij ons gaat gebeuren (dit geldt trouwens

in veel mindere mate voor Vlaanderen). Voor deze ene keer heb ik niet zo veel bezwaren tegen de voorttrekkersrol van de VS: er bestaat onder toonaangevende psychologen aldaar zeer veel onvrede over de preoccupatie met EST's en de uitwassen van de EST-methodologie. Als ik de hoogstaande en hoog oplopende discussies in kringen van de *American Psychological Association* goed inschat – zie Levant (2005) en Norcross, Beutler en Levant (2005) –, dan is er een koerswijziging op til. Hopelijk deelt u deze conclusie na lezing van het onderstaande.

### Een korte terugblik

Enig inzicht in de geschiedenis relativeert het heden en is niet alleen voor mijzelf maar misschien ook voor veel lezers nuttig. Niemand zal vandaag de dag nog volhouden dat aderlaten een goede therapie was; maar toen – *toen* geloofde men er heilig in. We kunnen daar natuurlijk lacherig over doen, maar over honderd jaar zal er even hard gelachen worden om de dingen waar wij nu onvoorwaardelijk in geloven. David Healy schaarst dit soort zaken onder de *therapeutic zeal*, de gedrevenheid van artsen en psychiaters (en psychotherapeuten), die behalve tot excessen als aderlaten, purgeren, lobotomie en allerhande psychotherapeutische gekkigheden ook tot belangwekkende ontdekkingen kan leiden en heeft geleid. 'Eén motief om gecontroleerde proeven in te voeren was de beteugeling van de furor therapeuticus' (Healy, 1997, p. 225).

Vlak na de Tweede Wereldoorlog werd de neppil (*placebo*) voor het eerst door Harry Beecher in een (voorloper van de) RCT toegepast. Aan het begin van de jaren '60 was men het er in grote meerderheid over eens dat placebo-gecontroleerd onderzoek het beste was. 'Er weerklonken nog altijd stemmen die klaagden over de ontmenselijkende aspecten van RCT's, maar tegenover elk van die geluiden stonden andere, die inhielden dat het onethisch zou zijn om niet-werkzame [farmaco]therapieën op de markt te brengen of te handhaven – en dat het grote publiek moest worden beschermd tegen zijn eigen hardnekkige neiging om zichzelf te vergiftigen' (p. 101).

Hoe verging het de psychotherapie? In de jaren '70 hadden meta-analyses van Smith en Glass de gunstige effecten van psychotherapie aangetoond. Maar als puntje bij paaltje kwam, was men in grote delen van de gezondheidszorg de mening toegedaan dat psychologische therapieën voor bepaalde stoornissen ofwel domweg niet werkten of inferieur waren aan farmacotherapie. In 1995 probeerde de *Division 12 (Clinical psychology) task force on promotion and dissemination of psychological procedures* van de *American Psychological Association* behandelingen door psychologen te promoten door criteria op te stellen waarmee 'empirisch gevalideerde behandelingen' (later herbenoemd tot 'empirisch ondersteunde therapieën', kortweg: EST's<sup>iv</sup>) voor bepaalde stoornissen kunnen worden vastgesteld. De *task force* kwam met 18 behandelingen waarvoor de empirische steun in RCT's goed genoeg werd geacht. Het gevolg was niet alleen een herwaardering van en geestdrift voor psychologische therapieën, maar ook en vooral heel veel gekissebis. Enerzijds vergrootte het rapport van de *task force* de erkenning van aantoonbaar effectieve psychotherapieën onder het grote publiek, bij beleidsmakers en in opleidingsprogramma's. Maar tegelijkertijd wonden veel psychotherapeuten zich op over het feit dat de *task force* zijn aandacht exclusief had gericht op korte, geprotocolleerde therapieën en op *specifieke* therapie-effecten in plaats van op de *algemene* factoren, die aantoonbaar een groot deel van de variantie in de therapieresultaten verklaren bij uiteenlopende stoornissen. In het rapport waaraan ik deze wijsheden ontleen, volgt dan ook de volgende waarschuwing ter bezwering van de wijdverbreide onvrede:

*De doelstellingen van EBP om de kwaliteit en kosteneffectiviteit te verbeteren en om de verantwoording (accountability) te vergroten zijn lofwaardig en vinden onder psychologen brede steun, al is het empirische bewijs voor vooruitgang over de hele linie door de implementatie van deze doelstellingen nog steeds karig [want de totale psychische-ziektelast van de bevolking neemt nog steeds toe, ml.]. Maar de gemeenschap van psychologen – en die omvat zowel wetenschappers als clinici – maakt zich zorgen dat het streven naar EBP wordt of kan worden misbruikt als rechtvaardiging om de toegankelijkheid van de zorg en de keuze uit behandelvormen ten onrechte in te perken. (Levant, 2005, p. 3)*

Er zijn ook psychologen die zich wat minder diplomatiek uitdrukken. Wampold schrijft in een recente bundel onder auspiciën van dezelfde APA een verhandeling onder de nietsverhullende kop: er is 'Geen spatje bewijs om de stelling te schragen dat empirisch ondersteunde behandelingen voor specifieke stoornissen betere resultaten zouden boeken dan andere [niet-empirisch ondersteunde] therapieën' (Wampold, 2005, pp. 299-308).

## RCT's en psychotherapie

Maar laten we eerst de roze bril nog even op houden en beginnen met de positieve resultaten van effectonderzoek naar psychotherapie. Van den Burg heeft het in zijn kruistocht tegen de zwaar overschatte effectiviteit van psychofarmaca (Van den Burg, 1994, 1995, 1998a, 1998b) alleen over CGT. De cognitieve gedragstherapieën komen er in vergelijking met psychofarmaca zeer genadig af. Maar dan volgt de niet onbelangrijke nuancering 'dat voor *geen*' van de angststoornissen [gegeneraliseerde angststoornis in ruime zin, idem in enge zin, paniekstoornis/agorafobie en sociale fobie] CGT, een farmacotherapie of een combinatie een panacee is: voor alle benaderingen geldt dat veel patiënten niet optimaal verbeteren en regelmatig blijken mensen die niet goed op benadering A reageren dat vervolgens wel te doen op benadering B. Uit de onderzoeksbevindingen tot nu toe kan dan ook hoogstens een benadering van eerste keuze gedestilleerd worden bij een "modale" patiënt, bij wie geen zwaarwegende argumenten (bijvoorbeeld de voorkeur van de patiënt) een rol spelen om een andere weg in te slaan. [Maar ondanks] deze relativisering valt er veel voor te zeggen bij alle vier angststoornissen, net zoals bij depressies, CGT zonder medicatie als benadering van eerste keuze te beschouwen' (p. 45).

Je kunt je afvragen of Van den Burg in zijn kritiek op psychofarmaca niet te positief is over CGT. Westen, onderzoeker en psychodynamisch psychotherapeut, heeft een andere mening over CGT en EST (het overgrote deel van de EST's is CGT<sup>vi</sup>):

*De RCT-resultaten van therapieën die het afgelopen decennium wijd en zijd worden betiteld als 'empirisch ondersteunde behandelingen' (EST's) voor depressie, paniek, gegeneraliseerde angststoornis (GAD), boulimia nervosa en obsessief-compulsieve stoornis (OCD) suggereren dat deze therapieën aanvankelijk inderdaad leiden tot substantiële verbeteringen in het geval van pijnlijke stemmingsbeelden en pathologische gedragingen, die diep ingrijpen in de levensvreugde en het adaptieve functioneren van mensen. De therapieën die wij meta-analytisch hebben onderzocht, zijn even effectief gebleken als of effectiever dan farmacotherapieën, die het voordeel hadden dat de farmaceutische industrie miljarden dollars kon steken in onderzoek en marketing. Tegelijkertijd steunen de beschikbare onderzoeksgegevens een genuanceerdere en empirisch evenwichtiger kijk op de effectiviteit van therapieën dan die welke wordt gesuggereerd door het wijdverbreide gebruik van termen als 'empirisch ondersteund', 'empirisch gevalideerd' of 'eerste-keuzebehandeling'. (...) [D]e effecten van korte, geprotocolleerde psychotherapieën variëren enorm al naargelang de stoornis, waarbij sommige onderzoeksresultaten de claim 'empirisch ondersteund' legitimeren en andere er ons aan herinneren dat empirische ondersteuning in de wetenschap gewoonlijk geen dichotome variabele is. Met uitzondering van CGT voor paniek herstelde de meerderheid van de patiënten in de therapieën voor alle stoornissen die we hebben onderzocht [depressie, paniekstoornis, GAD, boulimia nervosa en OCD] niet. Ze bleven last houden van symptomen, ook al bleek die last nu aanzienlijk verlicht en werd niet langer voldaan aan de DSM-criteria: ze verzochten om voortzetting van de behandeling of ze vielen binnen de periode van een of twee jaar na beëindiging van de EST terug (...). (Westen, Novotny & Thompson-Brenner, 2004, p. 650)*

Er valt niet aan te ontkomen: we hebben het goede nieuws nu wel zo'n beetje gehad. In de volgende paragrafen zet ik aan de hand van een aantal stellingen de belangrijkste problemen met RCT's en de EST-methodologie op een rij. Ik baseer me hierbij in de eerste plaats op Westen en anderen (2004) – als alleen een (bladzij)nummer is vermeld, wordt verwezen naar dat artikel.

### ***Wat als een (relatief arbitraire) operationalisatie begint, dwingt de clinicus vervolgens in een keurslijf***

Een protocol ontstaat als een noodzakelijke operationalisatie van een variabele (behandeling X) en wordt vervolgens tot een prescriptief keurslijf, want alleen de geprotocolleerde behandelinterventies mogen worden toegepast. 'De clinicus kan niet in vol empirisch vertrouwen elementen uit het ene of het andere therapiepakket halen, want alleen het pakket *als geheel* is gevalideerd, en *niet* de specifieke onderdelen of mechanismen van de behandeling. Iedere overtreding van het protocol impliceert een ongefundeerd geloof in de validiteit van het eigen klinische oordeel – waarover de clinicus heeft geleerd dat het hoogst waarschijnlijk gemiddeld slechtere resultaten oplevert.' (638) Ergo: de enige empirisch verantwoord werkende behandelaar is hij die geen enkel klinisch oordeel velt en geen enkele beslissing neemt, maar slechts domweg het protocol uitvoert waarin al die beslissingen al voor

hem of haar zijn genomen. Hoeveel therapeuten die EST's uitvoeren, zouden zich daartoe beperken? Anders gezegd: hoeveel therapeuten werken *werkelijk* empirisch verantwoord? Ik ben geen therapeut, maar ik kan het antwoord wel raden. Maar wanneer de strak geprotocolleerde therapie de toekomst heeft, zijn de therapeuten nog slechts een soort protocolslaven.

De stap van interventies als *exemplarisch* voor de bedoelde therapie naar *prototypische* interventies lijkt klein maar is ingrijpend: hij staat voor de verschuiving van de validering van behandelprincipes naar de validering van kant-en-klare behandelpakketten.

### ***Effectonderzoek is zuiver wetenschappelijk van weinig waarde, want het is theorieeloos onderzoek*<sup>vii</sup>**

Voor psychotherapieën waarvan de inhoud heilig wordt verklaard op grond van hun effectiviteit in RCT's geldt: zolang we niet op z'n minst een theorie hebben *waarom* iets werkt (of niet werkt), vergaren we geen (wetenschappelijke) kennis en worden we dus *niets wijzer*. Met alleen RCT's – of beter: RCT's zoals ze nu meestal worden toegepast – leren we niets over wat nu precies de werkzame bestanddelen van de therapie zijn. 'De empirische validering van therapiepakketten is nog altijd hardnekkig gefocust op de behandelresultaten en ontbeert essentiële informatie over wat van een therapie werkzaam is en wat niet.' (Ablon & Marci, 2004, p. 664) De momenteel razend populaire behandeling EMDR (*Eye movement desensitization and reprocessing*) voor posttraumatische stressstoornis is een goed voorbeeld: zelfs de fanatiekste aanhangers van deze methode geven toe dat ze geen idee hebben hoe en waarom EMDR werkt, zeker nu gebleken is dat die oogbewegingen er feitelijk niet toe doen. Door maar te blijven hameren op de vraag of EMDR nu wel of niet werkzaam is (zie Veeninga & Hafkenscheid, 2005), komen we nooit te weten *waarom* deze therapie werkt – of lijkt te werken –, wat de werkzame respectievelijk de niet-werkzame bestanddelen zijn en door welke oorzaken en mechanismen de patiënt verandert. Zodoende kunnen we een therapie ook nooit *verbeteren*; daarvoor is heel ander onderzoek nodig (zie Ablon & Marci, 2004; Westen en anderen, 2004).

### ***Bij wie werken EST's?***

RCT's en EST's verhullen meer dan ze onthullen. Niet alleen is na een succesvolle RCT niet bekend welke interventies genezend werken, maar evenmin wat nu precies het verschil is tussen de patiënten bij wie het effect wél optreedt en de patiënten voor wie de therapie niet heeft gewerkt. 'Steeds meer bewijsmateriaal suggereert dat de (...) aanname die zegt dat as-I-symptomen of -syndromen los van persoonlijkheidsprocessen kunnen worden begrepen, voor de meeste stoornissen onjuist is.' (636) Westen komt met veel voorbeelden van comorbiditeit die goed (beter) te verklaren is uit persoonlijkheidsdisposities (*personality diatheses*). Dit is natuurlijk een onderwerp dat vooral psychoanalytisch georiënteerde therapeuten zal aanspreken, maar misschien niet alléén psychoanalytici.

Volgens Healy bestaat er psychofarmacologisch onderzoek 'dat laat zien dat persoonlijkheidstrekken van patiënten voor 50 procent kunnen voorspellen hoe groot de kans is dat iemand reageert op antidepressiva (...) en ook op antipsychotica.' (Healy, 2002, p. 314) Maar voor dat onderzoek heeft de farmaceutische industrie geen belangstelling, want die wil het toepassingsdomein van een middel niet specificeren (= inperken) maar liefst zo omvangrijk mogelijk maken. Dat veel patiënten te veel psychofarmaca krijgen toegediend, is al een ramp, maar dat wetenschappelijk onderzoek naar het verband tussen stoornissen, geneesmiddelen (of psychotherapie) en persoonlijkheidsdisposities niet van de grond komt, is misschien nog erger: 'RCT's kunnen zo worden ingericht dat ze laten zien welk percentage patiënten naar alle waarschijnlijkheid een bepaalde [farmaco]therapie heilzaam en aanvaardbaar zal vinden. Maar zulk onderzoek wordt niet gedaan. Als wetenschap wil zeggen dat we *alle* mogelijke gegevens over een verschijnsel verzamelen, dan is de moderne psychiatrie ondanks haar retoriek allesbehalve wetenschappelijk. Maar is dit alleen een probleem bij farmacotherapie?' (Healy, 2002, p. 315) Het antwoord (ook Healy is die mening toegedaan) is natuurlijk: Nee, het geldt ook voor psychotherapie. Ook daar staan grote financiële belangen op het spel, zij het niet van een 'industrie' als de farmaceutische, maar wel van specialistische verenigingen en beroepsgroepen.

### ***Bieden EST's de beste interventies?***

Veel effectonderzoek gaat over kortdurende behandelingen die zijn vastgelegd in protocollen. Omdat een EST-protocol niet op grond van theoretische overwegingen is opgesteld, maar een meer of minder toevallige samenstelling kent, bevat het vrijwel zeker niet de best mogelijke interventies. '[V]eel van de procedurele details van een therapiepakket worden tijdens het schrijven van het behandelprotocol noodgedwongen beslist op niet-empirische gronden. Moeten sessies één of twee uur duren? Wekelijks of tweewekelijks? Individueel of groep? Totaal 20 sessies, of 12 of 16? Wanneer de beroepsgroep besluit dat een protocol "empirisch ondersteund" is, volgt daar geenszins logisch uit dat de specificaties die in dat protocol zijn vastgelegd optimaal zijn.' (Haaga, 2004, p. 674) Eenmaal effectief betoond gaat het protocol een eigen leven leiden: niemand twijfelt meer aan (de effectiviteit van) de therapie-ingrediënten.

Het zou best eens zo kunnen zijn dat een behandelmethode op een andere manier werkt dan de therapeut zelf denkt. Zo kwam Castonguay in een procesonderzoek naar CGT voor depressie tot de ontdekking dat focussen (door de therapeut) op disfunctionele cognities negatief correleerde met therapie succes. 'Nadere analyse toonde aan dat deze negatieve correlatie te verklaren was uit problemen in de therapeutische alliantie en uit een inhoudsanalyse kwam naar voren dat de therapeut in sessies met een slechte alliantie volhardde in zijn pogingen om sceptische patiënten te overtuigen van de relevantie van de cognitieve interventie.' (Goldfried & Eubanks-Carter, 2004, p. 671)

In sommige, helaas tamelijk uitzonderlijke gevallen is wél onderzocht *wat precies* het best werkt: '[G]edragstherapeuten hebben veel experimenteel onderzoek uitgevoerd naar procedurele details die te maken hebben met de implementatie van exposuretherapieën voor angststoornissen (...). (...) Moet het exposure in vivo zijn of imaginair of beide? Hoog- of laagfrequent? Wel of niet in aanwezigheid van de therapeut? Met of zonder deelname van de partner? Reëel of virtueel? Al deze vragen zijn experimenteel onderzocht (...) en de resultaten van dat onderzoek zijn te raadplegen door al degenen die gedragstherapeutische protocollen schrijven voor de behandeling van angststoornissen.' (Haaga, 2004, p. 675) De conclusie lijkt voor de hand te liggen: RCT's moeten specifieke interventies of therapeutische *principes* op hun effect onderzoeken, *niet* kant-en-klare behandelpakketten.

### ***De steekproef in een RCT is zelden of nooit klinisch representatief***

EST's en de onderzoeksmethoden om ze te valideren vertonen een aantal karakteristieke eigenschappen: deze '[t]herapieën zijn steevast ontworpen voor een enkelvoudige as-I-stoornis en de selectie van patiënten is erop gericht de homogeniteit van de diagnose te maximaliseren en bijkomende factoren die de variabiliteit van de respons op de therapie zouden kunnen vergroten, te minimaliseren. Therapieën worden geprotocolleerd en zijn van korte en vaste duur om de variabiliteit binnen de condities (*within-group variability*) minimaal te houden.' (632)

Niet de problemen van patiënten zijn de selectiecriteria voor opname in de RCT-steekproef, maar de DSM-diagnose. Gestreefd wordt naar een homogene steekproef van patiënten die dezelfde DSM-diagnose hebben. Omdat de meeste patiënten in de dagelijkse psychotherapeutische praktijk niet aan de DSM-criteria voldoen en blijf geven van comorbiditeit is de onderzochte steekproef voor hen nauwelijks relevant. Als het om comorbiditeit gaat, blijkt uit onderzoek 'consequent dat de meeste as-I-condities comorbide zijn met andere as-I- of as-II-stoornissen, variërend van 50% tot 90% van de gevallen (...).' (635)

Patiënten met de 'bedoelde' as-I-symptomen én met een persoonlijkheidsstoornis maken in onderzoek de kans op een significant resultaat kleiner, dus het is zaak deze patiënten er vooraf uit te zeven ('*outscreening*' wordt dat genoemd). Maar patiënten dropen, voor of tijdens het onderzoek, vermindert de generaliseerbaarheid van de behandeling naar de doelpopulatie, want die is 'zelden symptomatisch zuiver'. (637) Op het punt van de uitsluiting van patiënten in RCT's komt Westen op grond van meta-analyses van het beschikbare onderzoek tot de volgende slotsom:

*In de gemiddelde RCT van therapieën voor de meeste stoornissen, die op dit moment als 'empirisch ondersteund' worden aangemerkt, wordt tussen eenderde en tweederde deel van de patiënten die zich aanmelden van behandeling uitgesloten en de categorieën patiënten die worden uitgesloten lijken vaak niet alleen representatiever maar zijn in naturalistisch onderzoek ook therapieresistenter gebleken.* (658)

Westen en Morrison (2001) voerden een meta-analyse uit op alle onderzoek dat in de jaren '90 in de belangrijkste vaktijdschriften was gepubliceerd over EST's voor depressie, paniekstoornis en gegeneraliseerde angststoornis. De uitsluitingspercentages (*exclusion rates*) bleken te variëren van

65% voor GAD tot 68% voor depressie. Hierbij werden zowel legitieme (psychotische, bipolaire en organische stoornis; medische omstandigheden die de interpretatie van de resultaten zouden kunnen beïnvloeden; patiënten met een hoog suïciderisico) als twijfelachtige criteria gehanteerd.<sup>viii</sup> Andere meta-analyses, voor boulimia nervosa (Thompson-Brenner, Glass & Westen, 2003) en obsessief-compulsieve stoornis (Eddy, Dutra & Westen, 2004), leverden uitsluitingspercentages op van gemiddeld 40% van de *gescreende* patiënten voor boulimia en van 62% voor OCD. Dat is waarschijnlijk een *onderschatting* van het werkelijke cijfer, omdat er maar weinig onderzoeken naar OCD waren die het percentage weggeselecteerde patiënten prijsgaven. Het zal duidelijk zijn dat een RCT waarin zó veel patiënten om uiteenlopende redenen niet in het behandelresultaat worden meegeteld een vertekening is van de klinische werkelijkheid, zodat het met de generaliseerbaarheid en de externe validiteit van zulke onderzoeksgegevens (dat wil zeggen: wat zeggen de RCT-resultaten voor de gemiddelde behandeling in de psychotherapeutische *praktijk* van alledag) niet anders dan droevig gesteld kan zijn.

Een andere misleidende truc: de succescijfers uitdrukken als percentage van het aantal patiënten dat de behandeling voltooit en niet van dat aantal *plus* het aantal drop-outs, laat staan van het totaal aantal patiënten dat als potentiële deelnemer voor het onderzoek is gescreend. Dit alles flatteert de behandelresultaten van RCT's enorm. Als de herkomst van dat soort cijfers al wordt vermeld (wat vaak niet het geval is), dan nog vallen de meeste lezers (hoe goed bestudeert u onderzoeksartikelen?) er niet over.

### ***RCT's omvatten zelden follow-uponderzoek over een voldoende lange termijn***

Westen en Morrison (2001) concludeerden in hun meta-analyse van EST's voor depressie, paniekstoornis en GAD dat voor alle drie stoornissen *follow-up*gegevens voor de lange termijn ontbreken (> 2 jaar), en waar ze er wel zijn (= 2 jaar), kunnen alleen de therapieën voor paniekstoornis als succesvol worden beschouwd. Maar ook daar ontdekten de onderzoekers nog ongerechtigheden: veel patiënten die na 2 jaar als symptomvrij werden gescoord, bleken dat niet na 1 jaar en vice versa. Dit is, vooral bij episodische stoornissen, een belangrijk nadeel van louter kwantitatief onderzoek (zoals RCT's): patiënten die op het moment van de *follow-up* een voldoende scoren, worden meegeteld voor het resultaat, ook als ze *niet* de *hele* voorafgaande *follow-up*periode symptomvrij zijn gebleven. Een ander probleem in dit verband is dat de controlegroep lang niet altijd in de *follow-up*periode wordt gevolgd, omdat het nu eenmaal ethisch niet acceptabel is patiënten al te lang (op een wachtlijst ten behoeve van onderzoek) *onbehandeld* te laten. Relativering en correctie van toevalstreffers met behulp van controledata blijft dan achterwege.

Wat er op langere termijn overblijft van de behandelresultaten van CGT's bij deze stoornissen weten we niet. Voor boulimia geldt: '[S]lechts *eenderde* van de patiënten over de verschillende behandelvormen of in individuele CGT (...) gaf blijk van blijvend herstel (dat wil zeggen: is hersteld bij de beëindiging van de therapie en hersteld *gebleven* tot 1 jaar follow-up).' (646) Eddy en anderen (2004) constateerden dat slechts twee onderzoeken over OCD *follow-up*gegevens presenteerden langer dan een jaar. Maar ook die onderzoeken gingen niet vrijuit: *beide* onderzoeken gebruikten voor patiënten die na een jaar niet meer waren te traceren dezelfde gegevens voor 12 weken als voor 12 maanden, simpelweg door de eerste score 'vooruit te schuiven'. Ik durf er wel wat om te verdedden dat die truc alleen wordt toegepast bij patiënten die symptomvrij of hersteld leken na die eerste twaalf weken.

Dus: de gerapporteerde resultaten betreffen de korte termijn en het is allesbehalve zeker dat ze gedurende langere perioden na beëindiging van de therapie overeind blijven.

### ***Wat de parel zou moeten zijn in de RCT-methodologie – de controleconditie – blijkt de rotte plek***

In RCT's wordt veel te weinig gebruikgemaakt van relevante controlegroepen. Meestal bestaat de controleconditie uit een wachtlijst of geen behandeling. Een relatief 'actieve' controleconditie wordt meestal aangeduid met de term 'treatment as usual' (TAU), wat in de (Amerikaanse) praktijk zoveel betekent als de gebruikelijke therapie 'voor mensen die de middelen ontberen voor een redelijke behandeling.' (Westen en anderen, 2005, p. 429) 'Ongelukkigerwijs bestaan vrijwel alle TAU-controlegroepen uit behandelingen met een laag budget en een lage frequentie [en worden zij uitgevoerd] door minimaal getrainde paraprofessionals die een enorme patiëntenlast het hoofd moeten zien te bieden.' (641/2) Óf het gaat om 'behandelingen' die de naam niet verdienen en die eigenlijk

zijn ‘bedoeld-om-te-mislukken’: ze worden uitgevoerd door dezelfde doctoraalstudenten die de experimentele (bedoeld-om-te-slagen) behandeling geven en die donders goed weten dat deze behandeling voor haar welslagen gebaat is bij het mislukken van de controlebehandeling. ‘In tegenstelling tot placebo-pillen weten placebo-therapeuten dat ze worden verondersteld inert te zijn, waardoor controlecondities in psychotherapieonderzoek heel anders zijn dan placebo’s in medicijnonderzoek.’ (Westen en anderen, 2005, p. 429)

*[A]ls onderzoekers moeten we enorm voorzichtig zijn met het gebruik van termen als ‘therapie-zoals-gebruikelijk’, ‘traditionele therapie’ of ‘behandeling zoals gepraktiseerd in de GGZ’<sup>x</sup> (...), als wat we eigenlijk bedoelen, is: therapie zoals uitgevoerd door pas afgestudeerde, onervaren therapeuten in openbare gezondheidscentra voor niet-bemiddelde patiënten, waar notoir moeilijke patiëntenpopulaties samengaan met een notoir beperkt zorgaanbod. (642)*

Als je de effectiviteit van psychotherapie wilt bewijzen, dan kies je voor een zwakke controle: een wachtlijst of geen behandeling. Maar als je de werkzaamheid en de superioriteit van een specifieke psychotherapeutische aanpak wilt aantonen, dan is een wachtlijst of ‘geen behandeling’ of zelfs *treatment as usual* eigenlijk geen optie. Je krijgt weliswaar mooie hoge effectgroottes, maar je weet niks meer dan je voor het onderzoek al wist. Want het is inmiddels ten overvloede aangetoond: *[E]igenlijk leidt alles wat onderzoekers gedurende tien of twintig sessies met patiënten doen en waarvan ze geloven dat het zal werken tot betere resultaten dan experimentele condities die zijn bedoeld om niet te werken (...). Het enige wat uit zulk onderzoek kan worden geconcludeerd, is dat iets wat bedoeld is om effectief te zijn superieur is aan iets wat bedoeld is niet effectief te zijn. (Westen en anderen, 2005, p. 428).*

Alleen door twee of meer serieuze therapieën die de naam verdienen in een RCT tegen elkaar af te zetten valt uit te maken of een (of meer) ervan de bijdrage van de ‘algemene factoren’<sup>x</sup> overstijgt en zich dus kan beroepen op een specifiek effect. Westen maant onderzoekers te stoppen met iedere vorm van placebocontrole die bedoeld is niet te werken.

*Als het doel is EBP de weg te wijzen zouden onderzoekers nog slechts moeten vertrouwen op twee typen controlecondities: [a] therapieën die worden uitgevoerd door ervaren klinici die proberen een zo goed mogelijk behandelresultaat te bereiken en die niet worden beperkt in de duur van de therapie door ofwel de onderzoekers ofwel subsidiënten; en [b] verschillende duren van dezelfde experimentele behandeling, waardoor niet alleen gecontroleerd wordt voor ‘therapievoorkeur’ (allegiance<sup>xi</sup>) van de therapeut(en) en de onderzoeker(s), maar waardoor tevens empirische gegevens beschikbaar komen over de optimale therapieduur – [een afhankelijke variabele] die nog niet of nauwelijks onderwerp van EST-onderzoek is geweest. (Westen en anderen, 2005, p. 430)*

### ***Iedere impliciete of expliciete suggestie dat niet-onderzocht gelijkstaat aan niet-effectief is een vorm van misleiding***

Er bestaat in de literatuur – soms expliciet en soms impliciet – veel verwarring over ‘empirisch niet-onderzochte’ versus ‘empirisch weerlegde’ en ‘empirisch ongevalideerde’ versus ‘empirisch geïnvaleerde’ psychotherapieën. Omdat sommige – in het bijzonder de langerdurende – behandelvormen niet zo gemakkelijk in de RCT-mal te persen zijn (bijvoorbeeld omdat standaardisering van langerdurende therapieën vrijwel ondoenlijk is), concluderen de fans van EST’s maar al te gretig dat die langerdurende therapieën *niet-valide* zijn. Maar dat is een conclusie die je logisch gesproken helemaal niet mag trekken:

*Wanneer we als vakgebied verkiezen alleen bepaalde therapietypen te onderzoeken kunnen we geen conclusies trekken over een eerste-keuzebehandeling, tenzij we ons beperken tot het (kleine) universum therapieën dat empirische aandacht heeft gekregen. Door de eisen van korte duur en experimentele controle heeft de EST-methodologie verhinderd dat behandelmethoden zijn onderzocht die in de dagelijkse praktijk breed worden toegepast, wat heeft geleid tot de conclusie dat zulke therapieën empirisch geen ondersteuning zouden hebben. Maar deze conclusie vloeit logisch voort uit de methode, zij is niet op empirische gronden getrokken. (640)*

### ***Hoe korter de therapie, des te beter (en/of goedkoper)? Vergeet het maar!***

De meeste EST’s zijn eenvoudigweg te kort voor een goed resultaat<sup>xii</sup>, want de meeste stoornissen laten zich niet in drie maanden genezen. ‘De onwil om langerdurende therapieën te onderzoeken staat in schril contrast met de beschikbare naturalistische onderzoeksresultaten (in elk geval voor

volwassenen), met inbegrip van de gegevens van cognitief-gedragstherapeutische klinici in de dagelijkse praktijk: zij verrichten zelden behandelingen van drie maanden, tenzij ze daartoe worden gedwongen door verzekeraars die schermen met overgegeneraliseerde of verkeerd geïnterpreteerde gegevens uit RCT-onderzoek. Zij [= die onwil] contrasteert ook met de RCT's zelf, die laten zien dat de meeste patiënten na behandeling met de meeste EST's aan het eind van de therapie niet "genezen" zijn en dat veel patiënten die aanvankelijk blijk gaven van verbetering binnen twee jaar terugvallen.' (Westen en anderen, 2005, p. 429) '[E]en aanzienlijke hoeveelheid onderzoeksgegevens demonstreert dat de terugvalcijfers – met of zonder behandeling – hoog zijn voor alle stoornissen op een enkele uitzondering na (voornamelijk angststoornissen met zeer specifieke foci). (...) [D]e modale patiënt die behandeld wordt met kortdurende therapie valt bij de meeste stoornissen (andere dan die waarin een stimulus of voorstelling gekoppeld is aan een *zeer specifieke* cognitieve, affectieve of gedragsmatige respons) terug of vraagt binnen 12 tot 24 maanden om voortgezette behandeling.' (633) Dat de maatschappelijke kosten van behandeling in deze situatie wel eens veel hoger zouden kunnen uitvallen dan de voorstanders van kortdurende therapie beweren, hoef ik u natuurlijk niet voor te rekenen.

### ***De resultaten van RCT's laten zich tot op grote hoogte voorspellen uit de therapievoorkeur van de onderzoekers***

We worden *bedrogen*, of dat nu bewust of onbewust gebeurt. Ik doel hier op de veelbesproken, maar zelden verdisconteerde loyaliteitseffecten (*allegiance effects*). In de eerste plaats: welke behandelingen worden *onderzocht*, wordt sterk bepaald door de ideologische voorkeur van de onderzoekers. Omdat de meeste onderzoekers CGT-*mind*ed zijn, is CGT de meest onderzochte therapie. In de tweede plaats worden de onderzoeksresultaten in hoge mate bepaald door de therapievoorkeur van de onderzoekers. Luborsky en zijn medewerkers kwamen in 1999 tot de volgende schokkende bevinding: 'De gemiddelde loyaliteitsschattingen correleerden 0,85 met het onderzoeksresultaat van 29 therapievergelijkingen. Dat betekent dat *bijna tweederde* van de variantie in de onderzoeksresultaten was toe te schrijven aan de theoretische voorkeur.' (Luborsky & Barrett, 2005, p. 264) Ondanks alle discussies over loyaliteitseffecten en alle correctiemogelijkheden in onderzoek sinds 1999 zagen deze onderzoekers vijf jaar later na bestudering van het recentste onderzoek nog steeds geen enkele aanleiding om hun conclusie te herroepen (Luborsky & Barrett, 2005, p. 266). Zoiets stelt al die harde-featurespretenties (niet alleen van effectonderzoek trouwens) toch in een heel ander daglicht!

### **Wat nu?**

We mogen met een (on)gerust hart concluderen dat RCT's *niet* de ultieme scherprechter zijn waarvoor zij worden aangezien. Er is in veel RCT's gesjoemeld, dat is zeker: de invloed van loyaliteitseffecten op onderzoeksresultaten is verontrustend groot gebleken. Westen stelt daarom voor in ieder onderzoeksteam ten minste één potentiële dwarsligger op te nemen (iemand van een andere oriëntatie), maar het zou natuurlijk nog mooier zijn als de onderzoekssubsidiënten dit als voorwaarde voor financiële steun zouden stellen. Onderzoeksresultaten worden niet alleen te veel naar eigen hand gezet, maar ook door belanghebbenden genegeerd.<sup>xiii</sup>

Kwalijk is ook dat veel onderzoek dat geen significante uitkomsten oplevert, niet wordt gepubliceerd. Dat komt niet alleen omdat onderzoekers liever met positieve en spectaculaire resultaten op de proppen komen, maar ook omdat tijdschriften vaak niet bereid zijn de resultaten (die wetenschappelijk net zo belangrijk zijn als significante uitkomsten) te publiceren. Daardoor schetsen de uitkomsten van meta-analyses, voor zover de onderzoekers er niet in zijn geslaagd ook het ongepubliceerde onderzoek in hun analyse te betrekken, *systematisch* een te rooskleurig beeld van de werkelijke situatie.

Toch moet de conclusie *niet* luiden 'Weg met alle RCT's', want de ene RCT is de andere niet. Westen pleit er bijvoorbeeld voor om in RCT's niet hele therapiepakketten te onderzoeken, maar afzonderlijke interventie(principe)s. Dus lang leve de RCT's, mits ze niet als de *enige* vorm van onderzoek en misschien ook niet als *hoogste* vorm ('gouden standaard') worden opgevat. Westen bepleit parallelle vormen van onderzoek, die elkaar kunnen aanvullen:

*Naast die van RCT's moet een aantal andere soorten onderzoeksgegevens bijdragen aan EBP (...).*

*Geen van alle zijn ze definitief, maar allemaal zijn ze relevant voor het uiteindelijke doel:*

*behandelingen ontwerpen en implementeren die werken. Het hele scala omvat gegevens van naturalistisch onderzoek, ('correlationeel') onderzoek naar het verband tussen therapieresultaten en -variabelen die er positief mee correleren (...), onderzoek naar therapeut- en patiëntvariabelen die*



*samenhangen met het therapieresultaat, fundamenteel wetenschappelijk onderzoek, klinische observatie en antropologische onderzoeksgegevens die van belang zijn voor culturele verschillen in persoonlijkheid, psychopathologie en therapierespons.* (Westen en anderen, 2005, p. 430)<sup>xiv</sup>

Westen (650-655) presenteert een waslijst van aanbevelingen om te komen tot kwalitatief betere RCT's. Ik kan die aanbevelingen voor onderzoekers hier niet opsommen. Overigens kunnen die suggesties ook kritische *volgers* van onderzoek de ogen (verder) openen. Onlangs las ik aan het eind van een van de bijdragen aan dit tijdschrift de dooddoener: 'Ondanks de beperkingen van dit onderzoek mogen we concluderen dat...' terwijl *nergens* in het artikel ook maar met één woord werd gerept van die 'beperkingen'. Zoiets is in mijn ogen onacceptabel. Onderzoek moet niet alleen zorgvuldiger worden opgezet, maar het moet ook zorgvuldiger en gewetensvoller worden gerapporteerd. Hier ligt een mooie taak voor tijdschriftredacties. Het zal duidelijk zijn dat de lezer recht heeft om te weten wat voor soort controlegroep is gebruikt; of er *follow-up* onderzoek is uitgevoerd, en zo ja, over welke posttherapeutische periode; hoe de *follow-up* scores precies tot stand zijn gekomen, in het bijzonder hoe de onderzoekers de scores hebben bepaald van patiënten die bij de *follow-up* niet meer konden worden bereikt; hoeveel patiënten vóór of tijdens het onderzoek zijn uitgesloten van behandeling en op wat voor gronden (plus de eventueel nog wél van hen bekende scores), enzovoort enzovoort. Redacties zullen de verslagen van onderzoek én succesverhalen over de resultaten van onderzoek kritisch moeten volgen en bevragen.

### **Het soortelijk gewicht van de therapeut**

RCT's en de EST-methodologie zijn omarmd uit een diepgeworteld wantrouwen tegen het klinische oordeel: de therapeut zou de werkzaamheid van zijn therapie overschatten en een blind geloof hebben in de zegeningen van zijn eigen therapeutische bekwaamheid. Maar het wordt hoog tijd voor enige herwaardering van het klinische handwerk: 'Meta-analytische en andere onderzoeksgegevens hebben afdoende aangetoond dat de kwaliteit van de therapeutische relatie een groot deel van de variantie in de therapieresultaten verklaart, zowel in RCT's als in behandelingen in naturalistische settingen (...). Zulke onderzoeksgegevens moeten EBP ongetwijfeld de weg wijzen.' (Westen en anderen, 2005, p. 430) Recentelijk hebben Duncan en Miller de aloude schatting van Lambert (1992), die stelde dat de algemene factoren ongeveer tweemaal zoveel bijdragen als de technische interventies, flink naar boven bijgesteld: de algemene factoren zouden volgens hen *zeven tot tien keer zoveel* variantie verklaren als de specifieke factoren (Duncan & Miller, 2005, p. 145).

Onderzoekers hebben stelselmatig nagelaten door randomisering te controleren voor effecten van de individuele therapeut, waardoor ze het risico liepen therapeuteffecten aan te zien voor therapie-effecten. De therapeut blijkt iedere keer weer, door allerlei onderzoek naar allerlei stoornissen heen, aanzienlijk meer variantie te verklaren dan de specifieke therapie, vooral als de therapievoorkeur van de onderzoeker(s) constant wordt gehouden (Wampold, 2001). Westen voegt daaraan toe:

*Als ons doel EBP is, moeten we aanmerkelijk meer geld en middelen inzetten om individuele verschillen tussen klinici te onderzoeken en expertsysteemmodellen van therapeutische vakbekwaamheid te ontwikkelen.* (Westen en anderen, 2005, p. 430)

Dat inzicht is inmiddels ook in de hoogste regionen van de APA doorgedrongen. Levant (2005, p. 10) constateert: 'De individuele therapeut heeft grote invloed op het therapieresultaat, zowel in RCT's als in praktijksettings.' Hij meent dat de kennis over klinische expertise zo zoetjes aan een stevig fundament begint te krijgen – 'Hoewel er nog altijd minder resultaten van onderzoek naar klinische expertise beschikbaar zijn dan naar psychologische interventies' – en hij somt de kennis op die er al is en die nakende is (ik zal die hier niet reproduceren; zie p. 14 van het rapport). De toekomstplannen wijzen overduidelijk op een andere route dan die van de roemruchte *Division 12 task force on promotion and dissemination of psychological procedures* uit 1995. De nieuwe *task force* begint haar lange verlanglijst van 'dringend noodzakelijk onderzoek' met:

- *De werkwijze bestuderen van klinici die de beste resultaten behalen in de GGZ, zowel in het algemeen als met bepaalde categorieën patiënten of problemen;*
- *De technische vaardigheden van kundige klinici bepalen bij de uitvoering van psychologische interventies die bewezen effectief zijn gebleken; (...)*

Het gevaar dat psychotherapie verwordt tot de blinde uitvoering van een protocol door een lager opgeleide psychotechnicus – wat volgens Westen 'het grootste geschilpunt is met de EST-beweging' (Westen en anderen, 2005, p. 430) – is misschien nog wel niet helemaal afgewend, maar een

dergelijke ontwikkeling wordt door de APA in elk geval als onwenselijk beschouwd. Toegegeven: dat is niet erg verwonderlijk voor een organisatie die hoogopgeleide psychologen en psychotherapeuten vertegenwoordigt. Maar toch, als het aan deze invloedrijke *task force* ligt, zal het accent in onderzoek verschuiven van de robottherapeut en de van A tot Z uitgeschreven protocollen naar de in zijn vak uitblinkende behandelaar, behandelprincipes, de werkalliantie en het therapeutische proces.

*Als onderzoekers willen dat klinici hun onderzoek serieus nemen, zullen ze klinici serieus moeten nemen. (...) We weten inmiddels genoeg over interne validiteit. Laten we nu iets te weten zien te komen over psychotherapie.* (Westen, 2005, p. 171)

## Epiloog

Ik heb altijd het (onbewezen) vermoeden gehad dat de populariteit van effectonderzoek naar psychotherapieën iets te maken heeft met het failliet van het popperiaanse project. Ook met Poppers falsificationisme is het immers *niet* gelukt het demarcatiecriterium voor wat wetenschap onderscheidt van pseudo-wetenschap ondubbelzinnig vast te stellen.<sup>xv</sup> Psychologische theorieën toetsen is ondankbaar werk (hoeveel onomstreden theorieën bestaan er eigenlijk in de psychologie?), want de concurrentie van andere invalshoeken is groot. ‘Dat doen we dus niet meer, te veel sores; laten we maar eens heel pragmatisch beginnen met te onderzoeken welke geneesmiddelen en therapieën werken’ – dat moet ongeveer de heersende gedachte zijn geweest.

Popper begreep niet veel van mensen: hij hield het voor wenselijk en mogelijk dat onderzoekers stelselmatig proberen hun eigen theorieën naar de afvalbak te verwijzen. Dat is duidelijk een brug te ver gebleken. Iets van dien aard zien we hier ook terug: onderzoekers zouden eigenlijk moeten proberen om het de therapie van hun voorkeur in onderzoek *zo moeilijk mogelijk* te maken. Maar dat is beslist te veel gevraagd van (veel, alle?) onderzoekers en therapeuten (zie Luborsky & Barrett, 2005). Een beroep op hun integriteit is nodig maar waarschijnlijk vruchteloos. Misschien moeten tijdschriften van hun auteurs eisen dat zij hun therapievoorkeur vermelden, naar analogie van de vraag die psychiaters en psychofarmacologen krijgen voorgelegd over mogelijke belangenverstrengeling. ‘Werkt u voor Eli Lilly? Da’s mooi, maar dan zullen we uw conclusies met een flinke schep zout nemen!’

Het is altijd een troost als de zaken er bij de concurrent slechter voorstaan dan in het eigen bedrijf. Als we Healy (1997, 2002) en Van den Burg (1994, 1995, 1998a, 1998b) mogen geloven, is de situatie in de psychiatrie een stuk beroerder: beiden komen tot een vernietigend oordeel over de psychofarmacologie en de biologische psychiatrie. De eerste noemt de psychofarmacologische literatuur – vaak geschreven door werknemers van de farmaceutische industrie, maar ondertekend door psychiaters die zich daarvoor lenen – ‘het psychiatrische equivalent van een *Big Mac*’. (Healy, 2002, p. 311) Zo erg is het met de psychotherapie gelukkig niet gesteld, maar het zou toch wat al te naïef zijn te veronderstellen dat het eigenbelang in psychotherapeutisch effectonderzoek buiten de deur kan worden gehouden. Ook hier staan aanzienlijke financiële belangen op het spel. Healy waarschuwt dan ook voor een al te simpele conclusie die de psychotherapie zou ontzien:

*Oppervlakkige lezing van deze geschiedenis [van de psychofarmacologie] suggereert wellicht dat dit het probleem is: een wetenschap die is gecorrumpeerd door het bedrijfsleven [de farmaceutische industrie]. Maar het feit dat een soortgelijke dynamiek het domein van de psychotherapie heeft aangetast, waarschuwt tegen de al te gemakzuchtige aanvaarding van een dergelijke conclusie. (...) Nu zo veel mensen met chloorpromazine of psychotherapie in hun levensonderhoud voorzien, is het onduidelijk of we wel in de juiste positie verkeren om goede oplossingen te herkennen als we ze tegenkomen.* (Healy, 2002, p. 332)

Om terug te komen op mijn aan- en inleiding: arrogantie en verblinding door eigenbelang is stellig geen prerogatief van cognitieve gedragstherapeuten, maar als je het tij mee hebt, wordt het wel verleidelijker om daarvan te profiteren en naast je schoenen te gaan lopen. Hoogmoed komt voor de val (psychoanalytici kunnen erover meepraten); iets meer bescheidenheid zou cognitieve gedragstherapeuten sieren. En iets meer solidariteit – als je dat woord anno 2006 nog mag gebruiken – met de andere ‘scholen’. Of ben ik nu even naïef-idealistisch als Popper?

## Literatuur:

- Ablon, J.S., & Marci, C. (2004). Psychotherapy process: The missing link. Comment on Westen, Novotny, and Thompson-Brenner. *Psychological Bulletin*, *130*, 664-668.
- Albersnagel, F. (2004). De tijd zal het leren. Reactie op Takens' bespreking van de CBO-conceptrichtlijn depressie. *Tijdschrift voor Psychotherapie*, *30*, 432-435.
- Balkom, A.J.L.M. van, Vliet, I.M. van, & Spinhoven, P. (2004). De CBO-richtlijn angststoornissen. Reactie op 'Richtlijnen en denkwijzen' van Roelf Jan Takens. *Tijdschrift voor Psychotherapie*, *30*, 436-438.
- Broeke, E. ten, Korrelboom, K., & Meijer, S. (2000). De verborgen agenda van Wouter Gomperts en het Nederlands Psychoanalytisch Instituut. *Tijdschrift voor Psychotherapie*, *26*, 36-40.
- Burg, W. van den (1994). De werkzaamheid van antidepressiva. Blikvernuwing in de biologische psychiatrie. *MGv*, *49*, 1195-1210.
- Burg, W. van den (1995). Blikvernuwing in de biologische psychiatrie. Antwoord aan Slooff en Den Boer. *MGv*, *50*, 288-292.
- Burg, W. van den (1998a). De (in)effectiviteit van psychofarmaca. *Dth*, *18*, 5-68.
- Burg, W. van den (1998b). De (in)effectiviteit van psychofarmaca. Antwoord aan Kahn en Cath. *Dth*, *18*, 196-211.
- Duncan, B.L., & Miller, S.D. (2005). Treatment manuals do not improve outcomes. In J.C. Norcross, L.E. Beutler & R.F. Levant (Eds.), *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions* (pp. 140-149). Washington DC: American Psychological Association.
- Eddy, K.T., Dutra, L., & Westen, D. (2004). *A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder*. Atlanta, GA: Emory University. (Ongepubliceerd manuscript)
- Goldfried, M.R., & Eubanks-Carter, C. (2004). On the need for a new psychotherapy research paradigm: Comment on Westen, Novotny, and Thompson-Brenner (2004). *Psychological Bulletin*, *130*, 669-673.
- Gomperts, W.J. (1999). De kortzichtige eenvoud van het depressieprotocol. Een kritisch commentaar vanuit het Nederlands Psychoanalytisch Instituut. *Tijdschrift voor Psychotherapie*, *25*, 87-103.
- Haaga, D.A.F. (2004). A healthy dose of criticism for randomized trials: Comment on Westen, Novotny, and Thompson-Brenner (2004). *Psychological Bulletin*, *130*, 674-676.
- Healy, D. (1997). *The anti-depressant era*. Cambridge: Harvard University Press.
- Healy, D. (2002). *The creation of psychopharmacology*. Cambridge: Harvard University Press.
- Herbert, J.D., Lilienfeld, S.O., Lohr, J., Montgomery, R.W., O'Donohue, W.T., Rosen, G.M., & Tolin, D.F. (2000). Science and pseudo-science in the development of eye movement desensitization and reprocessing: implications for clinical psychology. *Clinical Psychology Review*, *20*, 945-971. (Vertaling in: *Toegang tot de Psychotherapie Internationaal*, 2001, nr. 3, pp. 304-343).
- Jaspers, J.P.C. (2000). Cognitieve gedragstherapie: panacee of placebo? *Tijdschrift voor Psychotherapie*, *26*, 185-192.
- Klooster, G. ten (2005). Het paard en de wagen. Enkele reacties op de controverse over de NVM-profielinterpretatie. *Tijdschrift voor Psychotherapie*, *31*, 238-240.
- Lambert, M.J. (1992). Psychotherapy outcome: Implications for integrative and eclectic psychotherapists. In J.C. Norcross & M.R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 94-129). New York: Basic Books.
- Levant, R.F. (2005). *Report of the 2005 presidential task force on evidence-based practice*. Washington DC: American Psychological Association. ([www.apa.org/practice/ebpreport](http://www.apa.org/practice/ebpreport))
- Lietaer, G. (1995). Vergelijkend proces- en effectonderzoek in de psychotherapie: een pleidooi voor integratie? *Tijdschrift voor Psychotherapie*, *21*, 194-206.
- Luborsky, L.B., & Barrett, M.S. (2005). Theoretical allegiance. In J.C. Norcross, L.E. Beutler & R.F. Levant (Eds.), *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions* (pp. 257-267). Washington DC: American Psychological Association.
- Morrison, K., Bradley, R., & Westen, D. (2003). The external validity of controlled clinical trials of psychotherapy for depression and anxiety: A naturalistic study. *Psychology and Psychotherapy: Theory, Research and Practice*, *76*, 109-132.
- Muris, P. (2002). Eye movement desensitization and reprocessing (EMDR). Een reactie. *Psychopraxis*, *4*, 52-53.
- Norcross, J.C., Beutler L.E., & Levant, R.F. (Eds.) (2005). *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions*. Washington DC: American Psychological Association.
- Philipse, H. (2002). Lezing 'Karl Popper (1902-1994)', in de reeks 'Filosofen van de twintigste eeuw' (Studium Generale). Utrecht, 19 maart.
- SEPI (2004). On 'On psychotherapy research'. SEPI Forum, maart-april 2004. ([www.cyberpsych.org/sepi/research.html](http://www.cyberpsych.org/sepi/research.html))
- Slife, B.D., Wiggins, B.J., & Graham, J.T. (2005). Avoiding an EST Monopoly: Towards a pluralism of philosophies and methods. *Journal of Contemporary Psychotherapy*, *35*, 83-97
- Takens, R.J. (2004a). Richtlijnen en denkwijzen. Een reactie op de Conceptrichtlijnen depressie en angststoornissen. *Tijdschrift voor Psychotherapie*, *30*, 39-51.
- Takens, R.J. (2004b). Te kort door de bocht. *Tijdschrift voor Psychotherapie*, *30*, 439-444.
- Thompson-Brenner, H., Glass, S., & Westen, D. (2003). A multidimensional meta-analysis of psychotherapy for bulimia nervosa. *Clinical Psychology: Science and Practice*, *10*, 269-287.
- Thompson-Brenner, H., & Westen, D. (2004a). *Accumulating evidence for personality subtypes in eating disorders: Differences in comorbidity, adaptive functioning, treatment response, and treatment interventions in a naturalistic sample*. Boston University, Boston. (Ongepubliceerd manuscript)
- Thompson-Brenner, H., & Westen, D. (2004b). *A naturalistic study of psychotherapy for bulimia nervosa: Comorbidity, outcome, and therapeutic interventions in the community*. Boston University, Boston. (Ongepubliceerd manuscript)
- Veeninga, A., & Hafkenscheid, A. (2005). De plaats van EMDR in de behandeling van posttraumatische stressstoornis. *Gedragstherapie*, *38*, 275-284.
- Wampold, B.E. (2001) *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Erlbaum.

- Wampold, B.E. (2005). Not a scintilla of evidence to support empirically supported treatments as more effective than other treatments. In J.C. Norcross, L.E. Beutler & R.F. Levant (Eds.), *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions* (pp. 299-308). Washington DC: American Psychological Association
- VGCT (2005a). Brief bij 'Folders Gedragstherapie en Cognitieve Therapie'. 26 oktober.
- VGCT (2005b). Persbericht. 23 november.
- Westen, D. (2005). Patients and treatments in clinical trials are not adequately representative of clinical practice. In J.C. Norcross, L.E. Beutler & R.F. Levant (Eds.), *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions* (pp. 161-171). Washington DC: American Psychological Association.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 69, 875-899.
- Westen, D., Novotny, C.M., & Thompson-Brenner, H. (2004). The empirical status of empirical supported psychotherapies: Assumptions, findings and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.
- Westen, D., Novotny, C.M., & Thompson-Brenner, H. (2005). EBP ≠ EST: Reply to Crits-Christoph et al. (2005) and Weisz et al. (2005). *Psychological Bulletin*, 131, 427-433.

<sup>i</sup> Dit artikel verscheen oorspronkelijk in het *Tijdschrift voor Psychotherapie*, 5/2006, pp. 347-366. Ik heb bij het schrijven veel steun, aanmoediging en commentaar gekregen van Myriam Van Gael en Gerjan ten Klooster, waarvoor mijn grote dank. Paul Verhaeghe, Myriam Van Gael en Hans Snijders hebben mij belangrijke literatuur toegestuurd. Het woord 'evidencebeest' ontleen ik aan Ten Klooster (2005), die het op zijn beurt – naar eigen zeggen – van iemand uit de medische hoek heeft gejat.

<sup>ii</sup> Dat wil zeggen: de streepjes stonden in de brief echt op de verkeerde plaats.

<sup>iii</sup> Op mijn kritische e-mail aan de VGCT kreeg ik van de (toenmalige) directeur het volgende antwoord:

'De VGCT beweert natuurlijk niet dat de cognitieve gedragstherapie een panacee is voor alles. Maar kennelijk kan die boodschap wel gelezen worden in onze teksten. Wat we wel staande kunnen en willen houden is dat CGT aantoonbaar effectief is voor allerlei psychische aandoeningen en als (toegevoegde) behandeling bij allerlei medische aandoeningen ofwel, in de woorden van voorzitter Cas Schaap, voor het hele handboek psychopathologie en het handboek medische psychologie. Daarmee zegt de VGCT dan niets over andere vormen van psychotherapie die effectief (zouden kunnen) zijn. Wij zullen onze teksten bijstellen.'

Het was natuurlijk sportief dat zij na één briefje van mij hun teksten al wilden corrigeren. Minder geruststellend was het citaat van de nieuwe voorzitter van de VGCT, dat CGT aantoonbaar effectief is 'voor het hele handboek psychopathologie en het handboek medische psychologie'. Daaruit spreekt dezelfde arrogantie waar ik eerder over viel.

<sup>iv</sup> De *EBP task force* heeft uiteindelijk de volgende definitie omarmd: '*Evidence-based practice* in de psychologie (EBPP) is de integratie van het beste beschikbare onderzoek met klinische expertise in de context van kenmerken van de patiënt, diens cultuur en voorkeuren.' (Levant, 2005, p. 5) En: 'EBPP is het ruimere begrip. EST vertrekt vanuit een behandeling en vraagt of zij in welomschreven omstandigheden effectief is voor een bepaalde stoornis of een bepaald probleem. EBPP begint bij de patiënt en vraagt welk bewijsmateriaal (inclusief relevante resultaten van RCT's) de psycholoog kan helpen om het beste behandelresultaat te bereiken. Daarenboven zijn EST's specifieke psychologische therapieën waarvan is aangetoond dat ze effectief zijn in gecontroleerde klinische proeven, terwijl EBPP een bredere scala van klinische activiteiten omvat (zoals psychodiagnostiek, gevalsbeschrijving [*case formulation*], therapeutische-relatievorming). Als zodanig is EBPP de uitdrukking van een besluitvormingsproces met als doel verschillende soorten onderzoeksresultaten – met inbegrip van, maar niet beperkt tot RCT's – in het behandelproces te integreren.' (Levant, 2005, p. 6)

<sup>v</sup> Alle cursiveringen in citaten zijn van mij (ml.)

<sup>vi</sup> In hun proloog schrijven Norcross, Beutler en Levant (2005, p. 7): '60 tot 90% van de EST's die tot op heden dat keurmerk hebben gekregen op grond van de beslissingsregels van Divisie 12 [van de APA] zijn cognitief-gedragstherapeutische behandelingen (CGT's).'

<sup>vii</sup> Een nuancering is hier op haar plaats: de RCT, de adellijke vertegenwoordiger van het effectonderzoek, is *de facto* een theoretievrijdige vorm van onderzoek, maar niet *per definitie*. Er zijn wel degelijk theoretisch-wetenschappelijk verantwoorde RCT's mogelijk. Ze worden alleen niet of nauwelijks gedaan. De lezer moet dus bij het navolgende telkens bedenken dat de kwalen die hier aan RCT's worden toegeschreven niet *inherent* zijn aan gerandomiseerd-gecontroleerd experimenteel onderzoek, maar alleen aan de *huidige toepassing* ervan (ook aangeduid als de 'EST-methodologie'): RCT's die leiden tot panklare EST's.

<sup>viii</sup> 'Het prototypische onderzoek over depressie sloot patiënten uit met zelfmoordfantasieën of comorbide middelenmisbruik, terwijl beide symptomen veel voorkomen bij depressieve patiënten. Voor paniek waren de prototypische uitsluitingscriteria: gematigde tot ernstige agorafobische vermijding, iedere bijkomende as-I- of as-II-stoornis die onmiddellijke behandeling behoefde, ernstige depressie die primair werd geacht en recente eerdere therapie. Prototypische uitsluitingscriteria voor gegeneraliseerde angststoornis waren ernstige depressie, middelenmisbruik en suicidaliteit. (...) Voor alle drie de stoornissen tenderden de criteria naar uitsluiting van patiënten die ernstiger problemen hadden, meer comorbiditeit vertoonden en moeilijker te behandelen waren, zoals patiënten met *borderlinetrekken* met een grotere kans op suïcide en middelenmisbruik' (646). *Tel uit je winst*.

<sup>ix</sup> Ik heb '*in the community*' al naar het uitkwam nu eens met 'in de GGZ', dan weer met 'in de dagelijkse praktijk' vertaald.

<sup>x</sup> Inmiddels wordt aangenomen dat de algemene factoren zeven tot tien maal zo veel variantie verklaren als specifieke factoren (zie Duncan & Miller, 2005, p. 145).

<sup>xi</sup> *Allegiance* is hier afwisselend vertaald met 'voorkeur' en met 'loyaliteit'.

<sup>xii</sup> Dat in RCT's voor (ultra)korte behandelingen werd (en nog steeds wordt) gekozen, stelde aanvankelijk niet op enig empirisch fundament of theorie, maar had louter pragmatische redenen: zo was het bij vergelijking van farmaco- en psychotherapie van belang de behandelingsduur voor de twee condities min of meer gelijk te houden. Bovendien eist het

---

statistische design dat de varia(n)tie *binnen* de experimentele condities niet te groot wordt, waardoor langere therapieën moeilijk te onderzoeken zijn: hoe langer de behandeling des te meer variabiliteit. Naturalistisch (niet-experimenteel) onderzoek naar psychotherapie laat consequent een ‘dosis-responsrelatie’ zien, dat wil zeggen: langere therapieën, in het bijzonder die langer dan 1 of 2 jaar, zijn effectiever dan korte behandelingen. Zulk onderzoek laat ook zien dat in 5 tot 16 sessies weliswaar aanzienlijke verlichting van symptomen optreedt, maar dat voor duurzame ‘rehabilitatie’ een langere behandeling vereist is.

Westen concludeert uit zijn eigen naturalistische onderzoek (Morrison, Bradley & Westen, 2003): ‘Toen Morrison et al. (2003) klinici indeelden naar theoretische oriëntatie (...) bleken de kortste behandelingen, wat niet verrassend is, de cognitief-gedragstherapeutische te zijn. Maar zelfs deze behandelingen waren gemiddeld bijna twee keer zo lang als geprotocolleerde CGT’s voor dezelfde stoornissen [depressie, paniek en GAD].’ (Westen en anderen, 2004, p. 647) Westen verwerpt de mogelijkheid dat het hier om inefficiënte behandeling of om zakkenvullende therapeuten gaat: ‘[H]et resultaat dat RCT’s voor deze stoornissen consequent te zien geven – dat de gemiddelde patiënt aan het eind van het experiment met kortdurende psychotherapie niet vrij van symptomen is –, suggereert dat klinici naar alle waarschijnlijkheid reageren op de situatie dat patiënten dan nog altijd klinisch relevante symptomen vertonen. In de tweede plaats was [in het onderzoek van Morrison en anderen] comorbiditeit veeleer regel dan uitzondering. (...) Over het geheel van stoornissen en theoretische oriëntaties verdubbelde de gemiddelde therapieduur als de patiënt enige vorm van as-I- of as-II-comorbiditeit vertoonde en ook de aanwezigheid van klinisch belangrijke persoonlijkheidsproblemen voorspelde de therapieduur.’ (647)

Het naturalistische onderzoek van Thompson-Brenner & Westen (2004a, 2004b) was een verbeterde versie van bovenstaand onderzoek onder klinici, maar nu over boulimia: ‘Opnieuw rapporteerden klinici van alle theoretische oriëntaties dat zij patiënten veel langduriger behandelden dan de 16 tot 20 sessies die worden voorgeschreven in de meest onderzochte en verspreide protocollen. Hoewel cognitief-gedragstherapeutische behandelingen van kortere duur waren dan eclectisch/integratieve en psychodynamische therapieën omvatte de gemiddelde CGT 69 sessies, aanzienlijk meer dan de 19 die worden voorgeschreven in het protocol. (...) Deze onderzoeksgegevens stemmen overeen (...) met de dosis-responsrelatie, die in vrijwel alle naturalistische onderzoeken wordt gevonden en die aangeeft dat een patiënt gemiddeld gesproken een grotere verbetering vertoont naarmate de behandeling uitgebreider is, en dat geldt zeker als de patiënt persoonlijkheidsproblematiek heeft.’ (648)

<sup>xiii</sup> EMDR is hier opnieuw een mooie illustratie, maar ruimtegebrek belet mij om dieper op deze kwestie in te gaan. Zie bijvoorbeeld Herbert en anderen (2000) en Muris (2002).

<sup>xiv</sup> Ook op filosofische gronden valt goed te verdedigen dat een EST- en RCT-monopolie moet worden vermeden: leve het pluralisme van filosofieën en empirische methoden (Slife, Wiggins & Graham, 2005).

<sup>xv</sup> Deze formulering en de conclusie dat het popperiaanse project ‘failliet’ is, ontleen ik aan Philipse (2002).