De betekenissen van "evidence based handelen" en de aard van "evidence" *Lessen rond rugscholen en radicalisering*

2

HOOFDSTUK

Amy-Jane Gielen en **John Grin**
Universiteit van Amsterdam

Inleiding 60

1. Verschillende evaluatiemethoden: op naar een realistische synthese 61

2. Lage rugklachten en de rugschool 63
   2.1. Inleiding 63
      2.1.1. Lage rugpijn 63
      2.1.2. Rugschool 64
   2.2. Evidence over de effeciviteit van rugscholen: aard en beperkingen 65
   2.3. De (on)mogelijkheden van een realistische synthese 66

3. Radicalisering en preventieve interventies 69
   3.1. Inleiding 69
      3.1.1. Radicalisering 69
      3.1.2. Interventies 71
   3.2. Evidence over de effeciviteit van antiradicaliseringsbeleid: aard, beperkingen en beleidsimplicaties 72

Conclusies 75

Bibliografie 77

Inleiding

Evidence-based handelen is ooit geïntroduceerd in de Britse gezondheidszorg. Tevens is deze werkwijze ook – mede onder invloed van Tony Blair – verspreid naar andere terreinen, zoals onderwijs en hulpverlening, ook naar andere landen. In 1997 verkondigde Tony Blair namelijk '*what counts is what works*' met de intentie om een post-ideologische benadering op overheidsbeleid te ontwikkelen. Niet langer zouden ideologische motieven centraal staan bij de invoering van beleid, maar wetenschappelijke kennis (Davies e.a., 1999, 3).
Ook Nederlandse overheden zien in dit concept goede mogelijkheden om de effectiviteit en de efficiëntie van hun handelen te bevorderen op een wijze die inhoudelijk terdege onderbouwd is en (daarmee) tevens een zekere legitimiteit heeft. Daarmee is direct duidelijk dat "evidence" onder evidence-based beleid niet moet worden gezien als wetenschap, maar als beleidsanalyse, dat wil zeggen als kennis die een handelende actor moet ondersteunen en informeren (Lindblom, 1990; Dunn, 1994). Dat betekent dan direct ook, dat "evidence" niet eenduidig, langs technische weg, de strijd om problemen en oplossingen daarvoor kan beslechten, maar, integendeel, in die strijd een rol speelt.
Dat laatste blijkt ook uit het Nederlandse meta-debat in het *Tijdschrift voor Sociale Vraagstukken* over de zin en onzin van evidence-based handelen. Twee kwesties staan centraal. Het eerste betreft de vraag naar de betekenis van evidence voor het handelen. Van der Kooij & Rensen (2009, 23) zien het als een vertaling van de intentie van Sackett om eerstelijnswerkers ten dienste te staan in hun dagelijkse praktijk. En Kolmer (2009, 27) ziet het als een "ménage à trois" van wetenschappers, beleidsmakers en beroepsoefenaars die gezamenlijk werken aan het ontdekken van een patroon van een interventie, waarvan ze gefundeerd zeggen dat het werkt. Anderen zijn aanmerkelijk sceptischer. Van Reekum (2008, 24) classificeert het als hype en ziet het als een instrumenteel middel om beleid te legitimeren. Gradener (2009, 24) ziet het risico van evidence based werken in de vorm van databases en glanzende brochures. Weer anderen waarschuwen ervoor, dat evidence based standaarden en protocollen gemakkelijk de professionele rationaliteit, alsook daarmee het vermogen van professionals om kwalitatief hoogwaardig en cliëntgericht te werken, ondermijnen (Timmermans & Berg 2003; Tonkens, 2008). Ook genoemd wordt het risico van depolitisering en technocatisering van besluitvorming: politieke oordeelsvorming wordt verdrongen door "objectieve analyse" (Burssens, 2007).
In dit artikel doen we een poging om meer inzicht te geven in en een kritische analyse te presenteren van de voorwaarden waaronder "evidence" een basis kan zijn voor een meer effectief overheidsbeleid. We behandelen daartoe twee casus.

De eerste casus is er een uit de gezondheidszorg, het domein waar evidence-based handelen zijn oorsprong vindt. De tweede casus komt uit de sociale sector, waarin het voeren van een evidence-based beleid steeds meer tot norm wordt verheven. Het zal blijken dat de casussen meer op elkaar lijken dan men zou kunnen denken, in ieder geval bezien vanuit het vraagstuk van evidence-based beleid: in beide gevallen dient de aard van het probleem – en daarmee de interventie – principieel vanuit multidisciplinair perspectief bekeken te worden. De multidisciplinaire aard van de interventie heeft in beide casus gevolgen voor de methodiek die wordt gebruikt om de effectiviteit van de interventie te evalueren.

In dit artikel betogen wij dat de gangbare evaluatiemethoden onvoldoende geschikt blijken om de eventuele effectiviteit van de interventie te meten. Wij zullen in lijn van Pawson & Tilly (1997) laten zien dat het zinvoller is om te kijken naar mechanismen achter de interventie,

de invloed van de context en verschillende patronen van uitkomsten. Daarbij zullen we ook ingaan op zaken die Pawson (2002a, 2002b) onuitgepakt laat in de "realistische synthese" bij het evalueren van interventies met een multidisciplinair karakter.

We zullen eerst de gangbare evaluatiemethoden uiteenzetten en vervolgens ingaan op de "realistische synthese" van Pawson. Vervolgens zullen we overgaan tot besprekingen van de casuïstiek, waarbij we ingaan op de aard van de problematiek en interventie, naast de aard van evidence dat verzameld dient te worden.

1. Verschillende evaluatiemethoden: op naar een realistische synthese

Een belangrijk aspect van de discussie (en kritiek) op het voeren van een evidence-based beleid, is aard van de "evidence", meer precies de mate waarin sprake is van kennis die inderdaad informerend is voor, dat wil zeggen inhoud toevoegt aan, het handelen. Sommigen stellen dat de "evidence" bij evidence-based handelen vaak sterk gedomineerd wordt door een bepaald type kennis. Dat is dan het type kennis dat ten grondslag ligt aan het type analyses dat in de geneeskunde de basis vormt voor evidence-based handelen: randomized controled trials (hierna: rct's): formele, doorgaans kwantitatieve kennis van het verband tussen een interventie en het effect daarvan. Deze klassieke methode van het kwantitatief georiënteerde (quasi-)experiment verdeelt de doelgroep van het beleid (in principe aselect) over een experimentele en controlegroep, waarbij de eerste wel en de tweede niet aan het beleid wordt blootgesteld. Het beleidseffect wordt uitgedrukt in het verschil tussen de na verloop van tijd gemeten effectvariabelen in beiden groepen. Met deze methode kan in beginsel getoetst worden of er een causale relatie bestaat tussen beleid en verandering. Omdat de experimentele groep en de controle groep voorafgaand aan de interventie vrijwel hetzelfde zijn, kan iedere geconstateerde verandering worden toegeschreven aan de interventie.

Om een overzicht te krijgen van de effecten van een bepaalde interventie, worden systematische reviews en meta-analyses uitgevoerd. In een systematische review wordt via een systematisch literatuuroverzicht een stand van zaken gemaakt van het wetenschappelijk onderzoek over een specifiek onderwerp. De verzameling en de beoordeling van de literatuur moet transparant zijn en gebeurt bij voorkeur aan de hand van richtlijnen, zoals bijvoorbeeld gehanteerd door de Campbell Collaboration. Dit is een vooraanstaande, onafhankelijke, internationale, non-profit organisatie die systematische reviews over het effect van interventies uitvoert en ter beschikking stelt. Een meta-analyse bouwt voort op een systematische review. Hierbij wordt aan de hand van statistische formules een algemeen "effect" berekend. Systematische reviews betreffende de thema's opvoeding en onderwijs, "social welfare" en criminaliteit zijn te vinden op de website van de Campbell Collaboration.[1] Een Nederlands overzicht van effectieve jeugdinterventies volgens wetenschappelijk onderzoek is ook te vinden op een website van het Nederlands Jeugdinstituut.[2]

1 http://www.campbellcollaboration.org

2 www.jeugdinterventies.nl

Op deze meta-analyses die voortbouwen op systematische review is ook kritiek. Kern van die kritiek is dat sociaal beleid zich niet afspeelt in een geïsoleerde klinische omgeving, maar binnen een sociale context die zeer sterk van invloed is op het wel of niet slagen van de interventie. In allerlei gevallen, waarin sprake is van meervoudige veroorzaking en/of causale relaties die geen eenduidige richting hebben, is kwantitatieve kennis veel minder van toepassing (Yanow, 2006; Burssens, 2007). Bovendien gaat ze vaak voorbij aan de rol die interpretaties in handelen spelen, zodat dezelfde interventie in verschillende gevallen tot verschillende resultaten kan leiden (Yanow, 2006). Anderen wijzen erop, dat zowel bij deze benadering als bij andere, meer narratieve benaderingen, er veelal sprake is van een onderontwikkeld begrip van *waarom* bepaalde interventies werken.

In de literatuur worden deze verschillende evaluatiemethoden naast elkaar gepresenteerd als onverenigbare polen (zie bijvoorbeeld Guba & Lincoln, 1989; Pawson & Tilley, 1997). Zo wordt de klassieke rationele evaluatie tegenover de post-moderne constructivistische evaluatie geplaatst. Kritisch realisten (Pawson & Tilly, 1997; Pawson 2002a, 2002b) hebben getracht deze "paradigma-oorlog" te slechten. Zo stellen ze dat waar formele, kwantitatieve kennis ten onrechte vaak contextuele invloeden weglaat, meer narratieve benaderingen vaak weer onvoldoende in staat zijn om te komen tot voor meerdere contexten relevante lessen. Om dit laatste dilemma te overstijgen, is het noodzakelijk dat analyses een beter zicht bieden op de (complexe) verbanden tussen mechanismen achter de interventie, de invloed van de context en verschillende patronen van uitkomsten (Pawson & Tilley, 1997). De identificatie en de verklaring van combinaties of het samenspel van contexten (C), van onderliggende mechanismen (M) en van de resultaten (O) is in het kritisch realistische onderzoek een kerntaak.

Pawson (2002b, 341-346) vertaalt deze benadering van evaluaties in het algemeen naar een methode voor het genereren van kennis voor evidence based handelen. Eerst behandelt hij de beperkingen van twee veel gebruikte methoden: meta-analyse en narratieve analyses. Meta-analyse, meestal van een reeks rct's omtrent interventies in een bepaalde problematiek, spitsen toe op een direct causaal verband tussen een bepaald type interventie en een uitkomst. Om dat te kunnen doen moeten de eerder onderzochte interventies worden geclassificeerd. Het resultaat is dan kennis over die klasse(n) die op de grootste schaal, het beste effect hebben gehad. Narratieve analyses richten zich niet op een enkelvoudig, causaal verband, maar op configuraties van eigenschappen van interventies, doelgroepen, omgeving enzovoorts. Daarmee geven ze nauwkeurig inzicht in elk van de bestudeerde interventies, maar reiken weer weinig hulp aan bij het overzetten van de lessen naar een andere context.

Pawson stelt daarom een realistische synthese voor, die een generatieve benadering van veroorzaking hanteert. Dat wil zeggen dat de analyse zich richt op de mechanismen en context elementen die tezamen de uitkomsten voortbrengen. Het resultaat van zo'n synthese is een theorie over de werking van interventies rond een bepaald probleem. Van groot belang is dat daarmee aandacht wordt gevraagd voor de actieve bijdrage die actoren leveren aan het tot stand komen van de uitkomst ("agency"), in wisselwerking met de context.

In de rest van deze bijdrage willen we verder verkennen hoe zo'n theorie, waarin agency en context tezamen behandeld worden, tot stand kan worden gebracht.

2. Lage rugklachten en de rugschool

2.1. Inleiding

De geavanceerde verzorgingsstaat staat onder druk. De overheid oriënteert zich steeds meer op bepaalde kerntaken, te weten wet- en regelgeving, normstelling en handhaving. In Nederland heeft dat de afgelopen decennia tot een reeks hervormingen geleid. De uitvoering van sociale zekerheidsregelingen en gezondheidszorgregelen laat zij zoveel mogelijk over aan de private markt en spoort via financiële prikkels maatschappelijk partijen aan om zelf hun verantwoordelijkheid te nemen voor ziekte en arbeidsongeschiktheid. Overheidsregulering wordt vervangen door zelfregulering, publieke initiatieven door private initiatieven. De nieuwe Zorgverzekeringswet is de consolidatie van een langdurig proces van de terugtredende overheid op het gebied van gezondheidszorg en sociale zekerheid. In dat proces wordt tegelijk nadruk gelegd op de eigen verantwoordelijkheid van de maatschappij. De nieuwe Zorgverzekeringswet beoogt meer ruimte te geven om in te spelen op de behoefte van patiënten en cliënten. Zorgverzekeraars onderhandelen met zorgaanbieders over prijs, inhoud en organisatie van de zorg.
In deze context zou evidence-based handelen als basis voor kosten-effectieve interventies de verschillende spelers moeten helpen om de door hen toegedachte verantwoordelijkheid te dragen. In de rest van deze paragraaf gaan we nader in op een voorbeeld van een klasse problemen, waarin dit bijzonder pregnant naar voren komt omdat gezondheidsproblemen hier verbonden zijn met arbeidsomstandigheden, persoonlijke omstandigheden en leefstijl. Onder invloed van een toename van het aantal chronische zieken en levenstijl gerelateerde aandoeningen, is in deze casus het accent van curatief naar preventief verschoven. Centraal daarbij is dat ook preventieve interventies – willen zij vergoed worden door de zorgverzekeraar – "evidence based" moeten zijn en bovenstaande maatregelen moeten bezien worden vanuit de achtergrond om onze verzorgingsstaat effectiever en efficiënter te maken.

2.1.1. Lage rugpijn

Maar hoe werkt dat precies in de praktijk? Om dit te onderzoeken hebben wij gekeken naar een veel voorkomend gezondheidsprobleem uit deze klasse: lage rugpijn. Volgens het Nederlands Centrum voor Beroepziekten wordt in Nederland ongeveer 15% van het verzuim op het werk veroorzaakt door rugklachten. Meestal is dit verzuim kortdurend van aard, maar in ruim 10% van de gevallen ligt de verzuimduur tussen één en drie maanden en in een kleine 5% tussen drie en zes maanden. Ongeveer 1% van de verzuimepisodes duurt langer dan één jaar. Aandoeningen aan de rug verklaren ongeveer 12 tot 14% van het aantal langdurig arbeidsongeschikten in Nederland. In 90-98% van de gevallen zijn deze klachten "a-specifiek", dat wil zegen dat er niet een eenduidig lichamelijk defect (zoals een hernia) aan ten grondslag ligt (Vendrig, 1999, 1). Tegen de achtergrond van de gevoelde wenselijkheid om het beroep op zowel de sociale zekerheid als de gezondheidszorg is dit een interessante casus: ze is een voorbeeld van het recente streven naar meer preventieve aanpakken.
Het diagnostisch beleid van lage rugpijn was in de afgelopen decennia primair gericht op het uitsluiten van een specifieke pathologie. Traditioneel legden de artsen de aandacht op de oorzaak van de pijn in termen van een lichamelijke pathologie en poogden de oorzaak aan te pakken. Wanneer er geen lichamelijke pathologie geïdentificeerd kon worden, werd een psychologische oorzaak aangenomen, waarvoor de term psychogene pijn werd gebruikt. Vanwege het

ontbreken van een beter alternatief, werden mensen simpelweg geadviseerd te rusten, voorzichtig aan te doen en een pijnstiller in te nemen wanneer de pijn te erg werd. In het verleden was de enige medische aanbeveling die de medische beroepsgroep kon doen om "te leren leven met de pijn" en een beroep te doen op de ziektewet. Deze traditionele scheiding tussen lichaam en geest is inadequaat gebleken. Traditionele medische of paramedische behandelingen die geen psychologische factoren bevatten, zijn ineffectief gebleken voor de behandeling van chronische lage rugpijn (Vendrig, 1999, 1).
Over het algemeen wordt in de wetenschappelijke literatuur erkend dat naast eventuele lichamelijke factoren ook psychosociale factoren een rol spelen bij de etiologie als de prognose van lage rugpijn (Keijsers e.a., 1989, 112). Dat staat ook wel bekend als het biopsychosociale model (Bekkering e.a., 2001, 2):
'*Biopsychosociale factoren kunnen het afwijkende verloop veroorzaken en/of in stand houden. Voorbeelden zijn een verminderde mobiliteit, verminderde spierkracht of een verminderde coördinatie (biologische factoren), bewegingsangst of irreële gedachten over de lage rugpijn (psychische factoren) en de werksituatie of steun/acceptatie in de omgeving (sociale factoren).*'
Kortom, rugpijn heeft een meervoudige oorzaak en vereist dus een multidisciplinaire aanpak.

2.1.2. Rugschool

Een multidisciplinaire behandelmethode voor lage rugklachten is de rugschool, in de jaren '70 van de vorige eeuw in Zweden ontwikkeld door Zachrisson. Het geheel is gebaseerd op de etiologische theorie dat mechanische belasting een belangrijke factor is bij het ontstaan van lage rugklachten. Het programma leert dan ook om de mechanische belasting van de rug zo laag mogelijk te houden. De doelstellingen van de lessen worden door Zachrisson als volgt omschreven (Terpstra & Bouter, 1988, 60):
- De patiënten in staat stellen een actieve rol te spelen in het verbeteren van hun werksituatie, met als doel hun rugproblemen te verminderen;
- Kennistoename omtrent de rugproblematiek bij de patiënt. Dit zou leiden tot een beter begrip en verlaagt het risico van een ongeschikte behandelingskeuze door de patiënt;
- Secundaire preventie van rugklachten en het hiermee verminderen van de vraag naar medische en sociale voorzieningen.

In Nederland is het aanbod van rugscholing de laatste jaren enorm toegenomen. In 2001 zijn er naar schatting rond de 400-450 locaties, zowel in de eerste lijn als in de tweede lijn, van waaruit rugscholing wordt aangeboden. Rugscholing kent dan ook verschillende vormen (Baan e.a., 2003, 338):
- Van groepsgewijze vormen tot individueel aangepaste vormen;
- Laag-intensieve en hoog-intensieve programma's die vanuit verschillende organisaties aangeboden kunnen worden;
- Rugscholing voor acute, subacute en chronische rugpijn;
- Rugscholing op monodiciplinare wijze (bijvoorbeeld alleen fysiotherapie) en multidisciplinair (fysiotherapie en psychologie);
- En tot slot rugscholing aangeboden door eerste en tweedelijnswerkers of grote commerciële conglomeraten.

2.2. Evidence over de effeciviteit van rugscholen: aard en beperkingen

Er is veel onderzoek gedaan naar de effectiviteit van de rugschool. De uitkomst van deze onderzoeken schept echter geen duidelijk. De ene rugschool is wel effectief, de andere niet. In de literatuur wordt gesuggereerd dat methodologische kwaliteit van deze effectiviteitstudies niet deugt. Zo stellen Keijsers e.a. (1989) dat de onderlinge vergelijkbaarheid erg lastig is, omdat iedere rugschool namelijk een ander programma heeft. Ze hebben daarnaast kritiek op de opzet van de studies, die naar hun oordeel vaak niet voldoen aan de normale eisen van rct's:
'*Wat betreft de interne validiteit van de studies zijn onder andere de volgende lacunes geconstateerd. Er worden relatief weinig patiënten in de studies opgenomen, er wordt in de opzet van de studies geen gebruik gemaakt van een placebo-aandachtsgroep en een directe evaluatie van de doelstelling van rugscholen wordt niet uitgevoerd. Deze geconstateerde lacunes maken het vooralsnog onmogelijk tot een valide uitspraak te komen over de effectiviteit van de binnen de afzonderlijke studies onderzochte rugscholen.*' (Keijsers e.a., 1989, 112).

In een meta-studie naar rugscholing van Miedema en Van Loon (2003), komt naar voren dat de methodologische kwaliteit van het onderzoek de laatste tien jaar is toegenomen. Van de tot dan 25 gepubliceerde effectstudies, waarin men de effectiviteit van vijftien monodisciplinaire rugschoolinterventies en zestien multidisciplinaire Nederlandse rugscholingsprogramma's heeft bestudeerd, bleek ruim 60% op methodologie en rapportage een voldoende te scoren. Dat betekent dat de resultaten van het onderzoek naar een rugscholingsprogramma te zijn vergelijken met die van tenminste één controlegroep en dat alle patiënten bij aanvang van de behandeling rugklachten hadden. De meest voorkomende tekortkoming van de 40% van de studies die niet voldoende scoorde, was dat geen gebruik was gemaakt van een controlegroep. Aan de resultaten van deze studies is in de meta-analyse meer waarde gehecht dan aan de resultaten van studies met methodologische tekortkomingen.

De meta-analyse probeert, op grond van de beschouwde studies, te komen tot conclusies met betrekking tot de effecten van diverse rugscholen op pijn en functioneren. Daarbij wordt gekeken naar ziekteverzuim en werkstatus, zowel op de korte als de middellange termijn. Het valt niet mee om zo te komen tot eenduidige evidence-based richtlijnen. Het blijkt dat bij acute en subacute rugklachten de effecten van *monodiscipliniare rugscholingsinterventies* op pijn en functioneren beperkt zijn, maar dat er wel aanwijzingen zijn voor een effect op verzuim. Bij chronische klachten zijn er effecten op pijn en functioneren op de korte en middellange termijn. Een jaar na de interventie zijn deze effecten niet meer aantoonbaar. Ook bij patiënten met chronische klachten is een effect op verzuim aangetoond; dit effect houdt ook op de lange termijn aan. Succesvol zijn vooral de interventies waarin een actief oefenprogramma is opgenomen. Het samenvoegen van de studies met dit type interventies bij subacute en chronische patiënten leidt tot de conclusie dat deze combinatie duidelijk positieve effecten heeft op functie, ziekteverzuim en werkstatus bij patiënten met meer dan zes weken ziekteverzuim wegens rugklachten of chronische rugklachten ongeacht het verzuim.

De conclusie over de *laagintensieve, multidisciplinaire* rugscholingsprogramma's luidt dat deze programma's volgens het onderzoek tot nog toe alleen meerwaarde lijken te hebben voor de werkstatus van mannen met subacute klachten. Bij vrouwen met subacute klachten bleken de resultaten tegenstrijdig. Bij chronische patiënten zijn laagintensieve varianten van multidisciplinaire programma's niet zinvol. De achtergrond hiervan blijft onduidelijk.

Bij chronische patiënten laten *hoogintensieve* varianten van *multidisciplinaire* rugscholingsprogramma's op de korte termijn een positief effect zien op pijn en functioneren. Hiervoor bestaat sterk bewijs. Op de middellange termijn zijn deze effecten minder duidelijk en op de lange termijn ontbreekt bewijs voor effectiviteit. De resultaten ten aanzien van werkgerelateerde uitkomstmaten zijn wisselend. De rugscholingsprogramma's met de meest positieve resultaten hebben vooral gemeen dat een groot deel van de interventietijd besteed wordt aan actieve training. Hier ligt een duidelijke parallel met de conclusies over de monodisciplinaire rugscholingsprogramma's. Nader onderzoek moet uitwijzen voor welke chronische patiënten monodisciplinaire behandeling een voldoende adequate behandeling is en voor welke niet.

Er is veel onderzoek gedaan, toch blijven nog veel vragen, die juist voor handelende actoren relevant zijn, onbeantwoord:
- Het blijkt, vanwege de diversiteit in programma's en patiënten, alsook door verschillen in uitkomstmaten, lastig om de meest effectieve varianten of componenten te identificeren.
- Het blijft vaak onhelder wat de mechanismen zijn die tot succes leiden, zodat de relevantie voor handelende actoren beperkt is: moet men nu wel of niet vrouwen met subacute klachten laagintensief, multidisciplinair behandelen? Wat is het precies in trainingsprogramma's dat bijdraagt aan hoogintensieve, multidisciplinaire programma's?
- De niet-specifieke effecten van dergelijke programma's, zoals extra aandacht, zijn onbekend.

2.3. De (on)mogelijkheden van een realistische synthese

De besproken meta-casus is een goed voorbeeld van de valkuil waarover Pawson schrijft. Ze spitst toe op verschillende klassen interventies en stelt de vraag welke effecten interventies in die klasse hebben. Pawson (2002a) zet uiteen dat die vraag *a priori* niet een goede is, omdat de effecten van een interventie doorgaans afhangen van (i) de context en (ii) de wijze waarop actoren reageren op de interventie. Hij pleit daarom voor een "realistische synthese", een meta-analyse die niet op klassen interventies is toegespitst, maar op klassen van mechanismen waarmee actoren worden beïnvloed.

De vraagstelling in zo'n "realistische synthese" moet dan zijn hoe die mechanismen, in wisselwerking met de context, verschillende actoren beïnvloeden en zo tot effecten leiden. De centrale idee achter Pawson's methode is "realistische synthese". Dat wil zeggen dat men nu juist de inconsistenties tussen bestaande analyses gebruikt om theorie te bouwen over de wijze waarop interventies leiden tot mechanismen die actoren beïnvloeden. Het is plausibel dat zo'n "meta-analyse nieuwe stijl" onduidelijkheden als de in paragraaf 2.2 genoemde onbeantwoorde vragen zou kunnen helpen oplossen. In zo'n synthese zou worden gezocht naar de mechanismen die kunnen verklaren waarom mannen meestal en vrouwen soms goed op laagintensieve programma's reageren; naar de mechanismen die de samenhang tussen middellange termijn effecten en het opnemen van actieve training in het programma beschrijven, et cetera. Het is ook duidelijk dat realistische synthese, met zijn nadruk op de relatie tussen context, mechanismen en uitkomsten goed past op het type problemen waarvan lage rugklachten een voorbeeld zijn: problemen waarbij de determinanten van gezondheid en ziekte en hun effecten niet alleen in het lichaam zijn gelokaliseerd, maar in de wisselwerking tussen lichaam, geest, leefwijze en

leefomgeving. Maar daarmee vraagt ze ook om evidence die is gebaseerd op theorie die gericht is op het begrijpen van die wisselwerking. Pawson gaat er vanuit dat zo'n theorie in een realistische synthese uit de onderliggende analyses kan worden geconstrueerd, via in wezen een "grounded theory" (Glaser & Strauss, 1967) benadering. Enige verdere reflectie op onze casus suggereert dat dit echter gecompliceerder ligt dan Pawson onderkent.

Het punt is, dat het gebruik van rct's nauw aansluit bij de dominante theoretische benadering in de medische wetenschap. Daarin worden als voornaamste determinanten van gezondheid en ziekte gezien lichamelijke defecten en een interventie – door een medische professional – die dat defect verhelpt of zijn effecten vermindert (Van der Wilt, 1995; Grin, 2004, 2007). Rct's als dominante methodologie beoogt het verband tussen een interventie en zijn effecten vast te stellen en is daarmee nauw verwant aan dit type kennis. Deze dominante benadering past naadloos op de klassieke idee van de verzorgingstaat: zorg via een aanbod van interventies aan een in beginsel passieve patiënt/werknemer.

De behandeling van lage rugklachten door rugscholen (en overigens ook door fysiotherapeuten) kan nu juist aan een andere, meer preventieve aanpak bijdragen, omdat ze veel meer nadruk legt op de agency van de patiënt/werknemer. Deze behandeling lokaliseert ziekte en gezondheid dus niet slechts in een lichaam, maar in de wisselwerking tussen lichaam, context en agency. Daaronder ligt een ander type kennis: het zogenoemde biopsychosociale model. Desgevraagd zegt de rugschool Winnock daarover dat:

'*Langdurige pijn- en vermoeidheidsklachten zijn zelden alleen lichamelijk en zitten net zo min tussen de oren. Lichamelijke, mentale en sociale factoren staan voortdurend met elkaar in wisselwerking. Daarom wordt net als bij de paramedici gebruik gemaakt van het biopsychosociaal model, dat naar het totaalplaatje kijkt en niet alleen naar een van deze factoren, waarbij wordt verondersteld dat het ontstaan van de klachten niet hetzelfde is als het voorbestaan van de klachten. Bij het ontstaan van de klachten spelen namelijk andere factoren een rol dan bij het aanhouden van de klachten. Soms verdwijnen de klachten, maar soms houden ze langer aan. Indien de klachten aanhouden, komt door allerlei processen in een negatieve spiraal terecht. Het zoeken naar dé oorzaak heeft dan geen zin meer, beter is het om de vicieuze cirkel/negatieve spiraal op zo veel mogelijk punten te verbeteren.*'[3]

Merk op dat het biopsychosociale model daarmee een heel geschikt kader biedt voor realistische synthese zoals voorgesteld door Pawson: het kijkt niet naar een enkelvoudig causaal verband tussen interventie en ziekte, maar neemt de wisselwerking met agency en context mee. Juist daardoor geeft het een veel beter zicht op de effectiviteit van de interventie. Ook suggereert dit citaat dat juist zo ook de effecten op middellange termijn kunnen worden begrepen. Maar juist tegen die achtergrond is tevens duidelijk dat rct's een aantal ernstige tekortkomingen hebben als nuttige "evidence." In de eerste plaats leggen ze een verband tussen interventie en effect, voorbijgaand aan de wijze waarop dat verband wordt beïnvloed door mechanismen en context. Dit probleem zou kunnen worden ondervangen met rct's waarin (i) relevante onderscheiden worden gemaakt tussen verschillende contexten en (ii) de vraag naar de relaties tussen interventie, relevante mechanismen en effecten wordt gesteld. De voorwaarde is dan wel, dat in theoretisch opzicht een model beschikbaar is dat inzicht biedt in wat "relevante"

3 Interview met bedrijfsarts en tevens medewerker business development en productmanagement sociale zekerheid van een grote zorgverzekeraar, 7 juli 2006.

onderscheiden en mechanismen zijn. En waarschijnlijk zijn (ook) andere methoden dan rct's gewenst om de bijpassende "evidence" te verzamelen.

Pawson ziet zo'n model als product van een realistische synthese. Dat is een zinnige invalshoek, maar in ons voorbeeld leidt het wel tot twee belangrijke kanttekeningen. De eerste is, dat de mate waarin een dergelijke synthese werkelijk zicht geeft op effectiviteit mede afhangt van wie haar maakt, alsook op grond van welke kennis. We hebben gezien dat het biopsychosociale paradigma, vooral ontwikkeld en gebruikt door paramedici, daarvoor veel beter geschikt is dan het paradigma dat nog steeds dominant is in de medische wetenschap. De besproken meta-analyse laat zien wat er mis kan gaan als (impliciet) dat laatste paradigma als uitgangspunt wordt genomen. (zie ook Vos e.a., 2002). Om te vermijden dat routineus wordt gekozen voor deze benadering is het noodzakelijk dat de "realistische synthesemaker" inzicht heeft in de verhouding tussen deze interventie en verschillende theoretische benaderingen van dit type aandoeningen. Let wel: dat blijkt al van belang bij een casus uit de gezondheidszorg, het gebied bij uitstek waar de oorspronkelijke idee van evidence based handelen op grond van meta-analyses vandaan komt. Dat zou ons op andere terreinen des te waakzamer moeten maken.

De tweede kanttekening is dat zo'n theorie, eenmaal ontwikkeld in een realistische synthese, waarschijnlijk zal leiden tot een beter begrip van de mechanismen en contexten die er toe doen. Zo'n nieuw ontwikkelde theorie zou, idealiter, dus ook moeten leiden tot een andere opzet van een "nieuwe generatie" rct's, of zelfs andersoortige evaluaties.

Beide kanttekeningen tezamen roepen indringende vragen op omtrent de institutionele kant van evidence-based handelen. Wie mag op grond waarvan bepalen hoe een realistische synthese wordt gemaakt en op grond van welke kennis? Kan dit alleen op grond van formele kennis? Of is ook juist waar zo'n theorie nog maar rudimentair bestaat, de interpretatie van onderliggende evidence gebaat bij kennis van en vanuit de praktijk? Ons voorbeeld suggereert dat men het biopsychosociale model ook zou kunnen opvatten als een heuristisch kader voor een realistische synthese. Immers, dat model voegt iets toe aan ervaringen uit de praktijk. Zo kan zo'n synthese tevens bijdragen aan de verdere ontwikkeling van het model. Dat is immers nog relatief onderontwikkeld, omdat in de wederzijdse beïnvloeding tussen verzorgingsstaat en medische wetenschap het klassieke model veel meer aandacht (en middelen) kreeg. Met andere woorden, men zou het biopychosociale model zoals we het nu kennen, in het bijzonder rond rugklachten, als uitgangspunt kunnen gebruiken voor verdere verrijking door realistische synthese. Aldus kan men een stap verder komen dan via de minder a priori geïnformeerde "grounded theory"-benadering die Pawson suggereert. Zo kunnen theorievorming en realistische synthese elkaar in de loop der tijd informeren. Het aardige van de biopsychosociale model is dat, hoewel nog onderontwikkeld, een redelijk model is dat past op de interventie – omdat het de verbindingen tussen C, M en O thematiseert – wel degelijk voor handen is. Maar hoe doen we dat als dat nog nauwelijks het geval is? Daarover gaat de volgende paragraaf.

3. Radicalisering en preventieve interventies

3.1. Inleiding

In Nederland worden er miljoenen euro's uitgetrokken om via meerjarige onderzoeksprogramma's de effectiviteit van interventies in de sociale sector te onderzoeken, om zo een evidence based aanpak te kunnen ontwikkelen. In toenemende mate financieren ministeries en gemeenten in Nederland alleen nog maar sociale interventies die evidence based zijn. Een toenemende trend daarbinnen, is de ontwikkeling van databases waarin evidence based interventies zijn opgenomen. Zo heeft bijvoorbeeld MOVISIE (een landelijk kennisinstituut voor maatschappelijke ontwikkeling) in opdracht van de overheid een meerjarig onderzoeks- en implementatieprogramma opgezet, dat gericht is op een substantiële verhoging van de kwaliteit en effectiviteit van het uitvoerende werk in de sociale professies. Een onderdeel daarvan is een onderzoek naar stimulering van maatschappelijke binding van jongeren, waarin een verkennend onderzoek wordt gedaan naar sociale interventies op het terrein van radicalisering en culturele spanningen. Lub (2009) concludeert in dat rapport dat slechts een aantal van deze interventies is geëvalueerd (Gielen 2008a, 2008b, 2009a, 2009b; Zannoni, 2009), maar over het gros van interventies is met betrekking tot effectiviteit weinig bekend. Daarnaast concludeert hij dat meer toetsende onderzoeksmethoden (volgsystemen, controlegroepen, longitudinaal onderzoek et cetera) niet worden toegepast. Het rapport doet daarom een aantal suggesties voor vervolgonderzoek. Immers, de hamvraag van het rapport – "what works?" – is vanwege het beperkte aantal effectevaluaties onbeantwoord gebleven. Het doel was om meer bruikbare informatie te verzamelen die tot een sterke bewijsvoering van de effecten van sociale interventies op het terrein van radicalisering en culturele spanning in Nederland kunnen leiden. De twee voorstellen die in het rapport worden gedaan om deze doelstelling in de toekomst alsnog te realiseren, sluiten nauw aan op de realistische synthese van Pawson. Enerzijds wordt gepleit voor meer "theory based evaluation", dat de theoretische vooronderstellingen en impliciete assumpties van sociale interventies expliciet maakt en toetst. Anderzijds wordt het voorstel gedaan om te kijken naar de context van de interventies en de verschillende perspectieven van partijen die betrokken zijn bij de interventies. Kortom, de drie elementen van Pawson: mechanismen (theorie), context en agency (perspectieven) zijn in dit voorstel verwerkt.

Op basis van de analyse en conclusie van de voorgaande casus, achten wij dit voorstel voor effectevaluatie zeer zinnig. Het rapport laat echter onuitgepakt hoe dat precies in de praktijk z'n uitwerking krijgt. Immers, de casus van de rugschool heeft ons geleerd dat het voor de realistische synthesemaker van belang is om inzicht te hebben in de verhouding tussen de interventie en de verschillende theorieën die de aard van de problematiek, in deze casus radicalisering, kunnen verklaren. Kortom, kennis over hoe radicalisering wordt veroorzaakt, is een noodzakelijke randvoorwaarde om een goede effectmeting te kunnen ontwikkelen volgens de realistische synthese.

3.1.1. Radicalisering

Welke theoretische kennis is bekend over het fenomeen radicalisering? Radicalisering is een relatief nieuw fenomeen. Sinds de Nederlandse samenleving met de moord op Theo van Gogh, op eigen bodem werd geconfronteerd met religieus geïnspireerd terrorisme, zijn er diverse onderzoeken verschenen over de factoren die van invloed zijn op het radicaliseringsproces bij

(Marokkaans-Nederlandse) moslimjongeren (AIVD 2006; Buijs e.a., 2006; Slootman & Tillie 2006; Groen & Kranenberg 2006). Uit deze onderzoeken is naar voren gekomen dat radicalisering een meervoudige en gelaagde problematiek is. Allerlei factoren op het persoonlijke, interpersoonlijk en collectieve niveau en binnen een nationale en internationale context, blijken een rol te spelen. Kortom, de aard van radicalisering bleek multidisciplinair te zijn. Hoe de verschillende mechanismen en triggers van radicalisering op verschillende aggregratieniveaus op elkaar inwerkten, bleef echter onduidelijk.

In bovengenoemde literatuur werd een centrale rol weggelegd voor "identiteit" in het radicalisingsproces. Gielen (2008a) heeft daarom geprobeerd om de wisselwerking van die veelheid aan mechanismen, agency en context inzichtelijk te maken door het begrip identiteit en identiteitsvorming radicaliseringsproces verder uit te pakken. De verschillende aspecten van identiteit en hoe dat van invloed is op het radicaliseringsproces[4] zullen nu nader worden besproken.

In de eerste fase ("pre-radicalisering"), is er altijd sprake van een zekere voedingsbodem. In deze fase zijn de persoonlijke en interpersoonlijke aspecten van identiteit bepalend of iemand vatbaar is voor een radicale boodschap. Met persoonlijke aspecten doelen we hier op een aantal demografische kenmerken. Te denken valt aan leeftijd (pubers/adolescenten zijn vatbaarder, omdat ze middenin het proces van identiteitsvorming zitten) en lage sociale-economische positie. Verder zijn er ook een aantal psychologische kenmerken. Zo merkten Buijs e.a. (2006a) op dat jongeren vatbaarder zijn voor radicalisering naarmate zij een sterk rechtvaardigheidsgevoel hebben, meer behoefte hebben aan binding en gemakkelijk beïnvloedbaar zijn. In deze fase speelt de kwaliteit van sociale relaties eveneens een belangrijke rol, ook wel het interpersoonlijke aspect van identiteitsvorming genoemd. Hoe minder hechte sociale relaties (binnen gezin, school, vrije tijdsbesteding) een jongere heeft, des te vatbaarder hij/zij is voor radicalisering. Deze demografische, sociale en psychologische factoren kunnen vervolgens tot gevoelens leiden van wat "relatieve deprivatie" wordt genoemd: een psychologische interpretatie van de persoonlijke omstandigheden, waarbij het gevoel overheerst dat desbetreffende jongere het slechter heeft en meer wordt benadeeld dan anderen. Het gevolg is dat gevoelens van onrechtvaardigheid gaan overheersen.

Deze gevoelens van onrechtvaardigheid zijn vooral voor de jongere zelf vervelend, maar (nog) niet gevaarlijk. Anders wordt het wanneer een bepaalde ingrijpende gebeurtennis, "een trigger", zich in het leven van de jongere voordoet. Triggers kunnen een reflexieve fase (2) op gang brengen, waarbij een jongere zijn "ik" gaat evalueren: wie ben ik, wat wil ik, wat geloof ik? Triggers kunnen allerlei vormen aannemen: persoonlijk (dood van je moeder, zoals bij Mohammed B.) politiek (Gaza-oorlog of aanhoudende negatieve beeldvorming) en sociaal-economisch (gevolgen economische crisis). Gevoelens van oneerlijkheid treden in deze fase naar de voorgrond. Indien de desbetreffende persoon het idee heeft dat hij voldoende middelen heeft om binnen de structuren zijn/haar leven ten goede te veranderen, dan zal er geen (verdere) radicalisering optreden.

4 De verschillende fases van radicalisering zijn mede gebaseerd op een door Gielen & Petri (2009) ontwikkelde vragenlijst om zorgpunten van radicalisering te meten. Aan de ontwikkeling van deze vragenlijst is onder andere internationaal literatuuronderzoek vooraf gegaan (Borum, 2003; Stahelski, 2004; Siber & Bhatt, 2007; Dalgaard-Nielson, 2008).

Indien de persoon in kwestie echter niet het gevoel heeft dat iets aan zijn/haar situatie kan veranderen, is hij/zij vatbaar voor allesomvattende religieuze of ideologische wereldvisies, zoals het nationaal-socialisme of het islamitisch-radicale gedachtegoed (fase 3). De soort wereldvisie die de jongere ontwikkelt is afhankelijk van de context waarin de jongere zich bevindt. Hierbij valt te denken aan persoonlijke contacten ("rekruteurs"), de eigen vriendengroep of bepaalde internetpagina's die worden bezocht. Het belang om ergens bij te horen – ofwel het collectieve aspect van identiteit – gaat in deze fase overheersen. Gevoelens van onrechtvaardigheid veranderen hier in gevoelens van oneerlijkheid.

Dat kan aanleiding geven tot een verschuiving in de identiteit (fase 4). Waar men normaal gesproken een meervoudige identiteit heeft, wordt deze rigide: het denken en handelen wordt volledig bepaald op basis van de allesomvattende wereldvisie. Het leidt tot "deindividualisering": andere personen worden niet langer als individu beschouwd maar als groep. Er ontstaat wij/zij denken. De media kan dit gevoel nog verder versterken. Veel negatieve aandacht over de wij-groep kan leiden tot het versterken van het wij-gevoel. Hierdoor kan een duidelijke afzetting tegen de zij-groep ontstaan. In dit geval kan angst en intimidatie er toe leiden dat men zich sterker betrokken gaat voelen bij een bepaalde groep en zich gaat afzetten tegen de ideeën van de andere groep (Donselaar & Tanja, 2000). Kortom, er is sprake van polarisatie. Het rigide identiteitsaspect komt verder tot uiting, doordat een jongere wegtrekt uit netwerken en kerninstituties waar niet wordt gedacht en gehandeld volgens de allesomvattende wereldvisie. Dit neemt soms vrij extreme vormen aan: het contact met de ouders wordt verbroken, omdat zij niet ideologisch of religieus genoeg zijn. Banen en stages worden opgezegd en studies worden afgebroken, omdat het niet verenigbaar is met de ideologie/religie (denk aan een studie rechten die wordt afgebroken omdat de Nederlandse rechtsstaat niet wordt erkend). Deze fase is bijzonder zorgelijk te noemen, omdat jongeren zich afzonderen van de maatschappij.

Ten slotte kunnen dehumanisering en demonisering optreden (fase 5). Positieve eigenschappen voor de "wij-groep" (zuiver) en negatieve eigenschappen van de "zij-groep" (ongelovigen) worden verheerlijkt. In extreme gevallen wordt de zij-groep gelijkgesteld aan de vijand. De allesomvattende wereldvisie rechtvaardigt vervolgens het gebruik van geweld tegen de vijand. Het radicaliseringsproces is hier als het ware "voltooid", waarbij extremisme en terrorisme niet zijn uitgesloten.
Kortom, het radicaliseringsproces verloopt in stappen en op verschillende aggregatieniveaus. Ons model, dat je zou kunnen opvatten als het analogon voor radicalisering van het biopsychosociale model, impliceert dat ook radicalisering dus – net als lage rugklachten – interventies op meerdere niveaus vraagt.

3.1.2. Interventies

Laten we vanuit dit rudimentaire model eens de al bekende interventies onder de loep nemen. Er zijn veel rapporten verschenen op het gebied van radicalisering, maar over welke interventies eerstelijnswerkers kunnen gebruiken om radicalisering te voorkomen of tegen te gaan, is weinig duidelijkheid. Lub (2009) publiceerde een overzicht van een aantal specifieke interventiemethoden en Lousberg e.a. (2009) heeft zeer recentelijk een lijst gepubliceerd waarin alle interventies op het gebied van het voorkomen van moslimradicalisering én rechtse radicalisering, zijn gestructureerd. Ze kunnen (nog) niet als best practice fungeren, maar wel als voorbeeld dienen van interventies die mogelijk ingezet zouden kunnen worden.

Lousberg e.a. (2009, 11) maakt een onderscheid tussen drie verschillende interventies. Een preventieve interventie is bedoeld om radicalisering te voorkomen. Een curatieve interventie is bedoeld om radicalisering tegen te gaan zodra tekenen van radicalisering gesignaleerd zijn en gericht op het gebruik van zachte maatregelen. Een repressieve interventie is tot slot bedoeld om radicalisering tegen te gaan zodra tekenen van radicalisering gesignaleerd zijn, gericht op het gebruik van harde maatregelen (zoals het handhaven van regels of straffen). In totaal zijn er 213 interventies in kaart gebracht. Daarvan kunnen er 106 toegepast worden bij rechts-radicalen en 174 bij moslimradicalen. Er zijn 67 interventies waarbij geen onderscheid gemaakt wordt tussen rechts-radicalen of moslimradicalen. Conform de methode van systematische review is geprobeerd een onderscheid te maken van verschillende klasse(n) interventies. De verschillende interventies zijn geclusterd tot 15 verschillende soorten, die zijn weergeven in het onderstaande figuur.[5]

Figuur 2.1 – *Overzicht van interventies*

Preventief	Curatief	Repressief
Voorlichting aan jongeren over discriminatie, et cetera	Individuele netwerkaanpak	Sancties en regels
Training aan eerstelijnswerkers	Stage of werkplek	
Gesprek met (potentieel) radicale jongeren	Individuele mentoring	
Activiteiten van jongeren voor jongeren	Deradicalisering	
Contacthypothese (interreligieuze of -culturele dialoog)		
De inzet van rolmodellen/high potentials		
Interventies speciaal voor vrouwen		
Opvoedingsondersteuning		
Vergroten religieuze weerbaarheid		
Inzet op weerbaarheid-empowerment		

3.2. Evidence over de effeciviteit van antiradicaliseringsbeleid: aard, beperkingen en beleidsimplicaties

Zoals gezegd, is nog een zeer beperkt aantal effectevaluaties uitgevoerd. De vraag is daarom interessant hóe je deze interventies zou kunnen evalueren.

Toen het evidence based werken ook zijn intrede deed in de sociale sector is de aard van de evidence ook veel diffuser geworden, omdat rct's niet op ieder beleidsprobleem toepasbaar zijn. Veel interventies lenen zich technisch en ethisch gezien niet voor de onderzoeksopzet van een rct. In tegenstelling tot de exacte wetenschappen, speelt sociaal-wetenschappelijk onderzoek zich zelden af in een geïsoleerde ruimte of laboratorium. De interventie en het effectonderzoek vinden plaats middenin de complexe samenleving. Er zijn dus continu storende invloeden aanwezig. Daarnaast is het in de sociale sector ook ontzettend moeilijk voor een interventie bepaalde doelgroepen te bereiken, laat staan voor effectonderzoek. Bepaalde doelgroepen hebben vaak een enorme huiverigheid voor onderzoekers, zeker als het gaat om een doelgroep als "potentieel radicale jongeren". Soms moet je als onderzoeker content zijn dat je überhaupt toegang

5 Voor een nadere toelichting op de interventies en essentiële randvoorwaarden voor een effectieve inzet, zie Lousberg e.a. (2009).

krijgt tot de doelgroep en op de koop toe aannemen dat werken met een geradicaliseerde controlegroep niet tot de mogelijkheden behoort (Gielen & Junne, 2008). Desalniettemin worden óók in de sociale sector rct's nog wel als de gouden standaard gezien, onder andere omdat men behoefte heeft aan gekwantificeerde resultaten.

Nu we een model gepresenteerd hebben die theoretisch de oorzaken van radicalisering in kaart brengt en expliciet maakt hoe verschillende mechanismen in het proces van radicalisering, context en agency op elkaar inwerken, is het interessant te kijken hoe we de realistische synthese en de eerder geschetste aanvullende voorwaarden kunnen gebruiken om tot een goede effectevaluatie te komen. Kortom, de centrale vraag is hoe het identiteitsmodel van radicalisering behulpzaam kan zijn als heuristisch kader om tot een realistische synthese te komen.

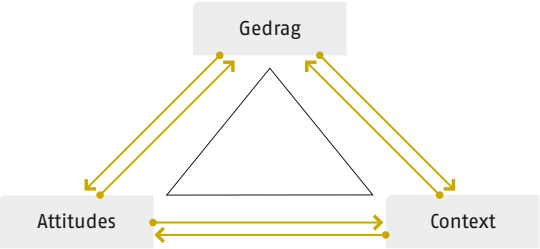
Het eerste wat het heuristische kader ons laat zien, is dat het alleen zinvol is de gepresenteerde interventies in te zetten op de juiste doelgroep. Een doelgroep die vanwege de context waarin zij zich bevinden, vatbaar zijn voor radicalisering: jongeren die zich in het proces van identiteitsvorming bevinden (pubers/adolescenten), zich begeven in een context van naasten met een lage sociaal-economische positie, specifieke psychologische kenmerken, weinig hechte (sociale) relaties binnen en buiten het gezin en tot slot toegang tot danwel nabijheid van een context waarin het radicale gedachtegoed aangeboden zou kunnen worden. Evaluaties dienen zich op deze doelgroep toe te spitsen; andere doelgroepen leveren geen geldig vergelijkingsmateriaal op.

Als de vatbaarheid voor radicalisering afhankelijk is van een wisselwerking van verschillende contexten, dan heeft dat als gevolg dat het weinig zinvol is om alleen een enkele interventie op een enkele specifieke context in te zetten. Alleen het inzetten van opvoedingsondersteuning voor ouders, zal niet voldoende zijn om jongeren minder vatbaar te maken voor radicalisering. Het is van belang het brede palet van interventies in te zetten, die zich vervolgens ook nog moeten richten op de relevante doelgroep.

Volgens de realistische synthese zouden het effect van de combinatie van verschillende interventies (M1, M2, M3 et cetera) in voorgenoemde context (C1, C2, C3 et cetera), moeten blijken uit de uitkomsten (O) dat de vatbaarheid voor radicalisering is afgenomen en er geen terroristische aanslagen zijn gepleegd. Het identiteitsmodel als heuristisch kader laat echter zien dat radicalisering een (langdurig en gelaagd) proces is in een complexe samenleving waarin verschillende mechanisme op elkaar inwerken die van invloed zijn op het gedachtegoed en gedrag van personen. Hierbij kan het plegen en het gebruik van geweld het eindstadium zijn, maar dat is niet per definitie het geval. Derhalve is niet zinvol om alleen "afname van terroristische aanslagen" als uitkomstmaat van deze interventies te hanteren. In plaats van de invloed van genoemde interventies in zeer algemene en brede zin te proberen te meten, is het interessanter te kijken naar attitude- en gedragsveranderingen, acties, activiteiten van personen en organisaties, die direct met de interventie in aanraking zijn gekomen. Hoe heeft de doelgroep gereageerd? Gedraagt zij zich nu anders? Is het gedrag van omstanders veranderd (vriendengroepen, medescholieren, buren, et cetera)? Zien de contactpersonen (leraren, politieagenten, jeugdwerkers) een verandering? Op deze manier wordt in kaart gebracht, welke veranderingen in de directe omgeving nog het meest met het project zelf in verband kan worden gebracht. Speculaties over verdergaande maatschappelijke repercussies worden vermeden, omdat hierover geen goed onderbouwde uitspraken kunnen worden gedaan. Kortom, de effectevaluatie

van de interventie(s) moet focussen op veranderingen (O1, O2, O3 et cetera) in attitude, gedrag en context en de wisselwerking tussen deze componenten, zoals is weergeven figuur 2.2.

Figuur 2.2 – *Aggregatieniveaus voor effectmeting*



Dat impliceert de ontwikkeling van indicatoren om verschillende uitkomsten van interventies te meten en bewijst nog maar eens de noodzaak van een gedegen kennis van de aard van het probleem (radicalisering). Gielen & Junne (2008, 150) hebben op basis van het identiteitsmodel een aanzet gegeven tot het formuleren van deze indicatoren.

Figuur 2.3 – *Indicatoren voor effectmeting*

Indicatoren op het individuele niveau:
- toename van contacten met andersdenkenden,
- toenemende belangstelling voor andere aspecten van het leven (familie, werk, school, sport, kunst, buurt, etc.)
- hoger tolerantieniveau tegenover andersdenkenden
- vaker op bezoek bij ouders/familieleden
- kennisname van verschillende media
- meer meldingen van discriminatie
- meer meldingen van mogelijke radicalisering

Indicatoren op het niveau van de groep:
- grotere openheid van de groep (voor bezoekers, vrienden, leden met andere opvattingen)
- uitnodiging van sprekers met andere meningen
- minder hiërarchie en meer diversiteit in de groep
- minder strikt verwachtingspatroon t.a.v. het handelen van groepsleden
- breder programma van de groep, nieuwe discussiethema's
- minder neiging zich af te zetten tegen andere groepen
- nieuwe rolmodellen
- uitnodigingen door andere groepen

Indicatoren op het niveau van de omgeving:
- positieve waarneming door jeugdwerkers, leraren, politieagenten, imams
- meer contact met de buren, leerkrachten, et cetera.
- actievere participatie in de samenleving
- minder discriminatie
- positievere berichtgeving in de media

Het interessante is dat uit dit model direct al een aantal belangrijke lessen voor beleidsmakers volgt:

- Het is allereerst van belang dat alle betrokkenen – wetenschappers, beleidsmakers en professionals – gedegen kennis hebben over de aard van de problematiek.
- Kennis over de aard van de problematiek is niet alleen essentieel om vervolgens goede interventies te kunnen ontwikkelen, maar ook gedegen en vooral ook bruikbare effectevaluaties.
- Het is van belang de interventies in te zetten op relevante doelgroepen in relevante contexten.
- Dat vereist dat verschillende interventies tegelijkertijd en parallel aan elkaar dienen te worden ingezet.
- Het streven naar algemene (abstracte) uitkomsten zoals "het voorkomen van radicalisering of een terroristische aanslag", zou plaats moeten maken voor een evaluatie van verandering in attitudes, gedrag en de (directe) context die toe te schrijven zijn aan de interventie.
- Dat impliceert, tot slot, dat onderzoekers al worden betrokken vóórdat de interventie wordt ingezet, om zo in de gelegenheid te worden gesteld om iets te kunnen zeggen over attitudes, gedrag en context voordat de interventie wordt geïmplementeerd.

Merk op, dat deze manier van evidence-based analyse inderdaad leidt tot, analytisch en empirisch gefundeerde, wezenlijke correcties en aanvullingen op eerder handelen en dus naar we mogen verwachten tot meer effectief beleid.

Conclusies

In de inleiding bespraken we het debat over evidence-based handelen en welke vragen dat opwerpt voor de aard van de evidence. Voor wat betreft de aard van de evidence zagen we dat rct-achtige kennis maar beperkt past op de problematieken waarom het ging. Dat geldt net zo goed voor de casus uit de gezondheidszorg (waar de rct zijn basis heeft) als voor de casus uit de sociale sector. Met name de casus over rugscholen laat zien dat deze beperking niet alleen te maken heeft met de methode (meta-analyse van rct-achtige studies) maar ook met de aard van onderliggende medische kennis. In beide gevallen geldt dat die beperking te maken heeft met de fundamentele kritiek van Pawson (2002a) op een benadering die toespitst op interventies en voorbijgaat aan context en aan de agency van betrokken actoren. Over het debat over de dominantie van een bepaalde benadering leren we dus i) dat inderdaad het "rct discours" domineert en ii) ofwel leidt tot een gebrekkige relevantie voor de praktijk (rugschool), danwel (bijna) niet toepasbaar is in sociale praktijken (radicalisering).

Hoewel de twee casussen uit verschillende disciplines kwamen, laten ze, in ieder geval vanuit het vraagstuk van evidence-based beleid, zien dat in beide gevallen de aard van het probleem – en daarmee de interventie – principieel vanuit multidisciplinair perspectief bekeken dienen te worden. De multidisciplinaire aard van de interventie heeft in beide casus gevolgen voor de methodiek die wordt gebruikt om de interventie te evalueren.

De realistische synthese van Pawson biedt daartoe belangrijke aanknopingspunten: een analyse die zich richt op de mechanismen en context, elementen die samen de uitkomsten voortbrengen. Het resultaat van zo'n synthese is een theorie over de werking van interventies rond een

bepaald probleem. Van groot belang is dat daarmee aandacht wordt gevraagd voor de actieve bijdrage die actoren leveren aan het tot stand komen van de uitkomst ("agency"), in wisselwerking met de context.

De casussen hebben ons laten zien dat het echter wel van belang is dat de "realistische synthesemaker" inzicht heeft in de verhouding tussen deze interventie en verschillende theoretische benaderingen van de aard van het probleem waarvoor de interventies zijn ontwikkeld. Zeker bij meer complexe vraagstukken is het verstandig om als uitgangspunt van de realistische synthese een model te hebben dat verschillende invalshoeken op een doordachte wijze synthetiseert. Deze kennis zal leiden tot een beter begrip van de mechanismen en contexten die er toe doen. Zo'n nieuw ontwikkelde theorie zou, idealiter, dus ook moeten leiden tot een andere opzet van andersoortige evaluaties en uiteindelijke leiden tot een doel(groep)gerichte en effectievere inzet van evaluaties en daarmee – uiteindelijk – tot meer effectief beleid.

Misschien wel het belangrijkst is dat dergelijke nieuwe evaluatiemethoden ook recht doen aan de betekenis die de betrokken praktijkmensen toekennen aan evidence-based handelen. Zowel voor- en tegenstanders van het evidence based werken zijn het erover eens dat interventies alleen effectief kunnen zijn in de relatie tussen cliënten en professionals, met concepten en theorieën die zijn ontleend aan hun handelen. Die relatie of anders gezegd, mechanismen kan alleen in kaart worden gebracht als wetenschappers, beleidsmakers en professionals gezamenlijk werken aan het ontdekken van mechanismen van een beleid in een bepaalde context, waarvan gezamenlijk kunnen constateren dat het werkt. Dat vereist dan natuurlijk een goed geformaliseerde aanpak. Daartoe moet echter wel het nodige aan Pawson's aanpak worden toegevoegd, ook in principiële zin. In de eerste plaats, zo zagen we bij de rugschool casus, moet rekening worden gehouden met het feit dat kennis in meerdere paradigma's komt. Bij het maken van een "realistische synthese" moet kritisch worden bezien welke paradigma's worden weerspiegeld in de gebruikte evidence en hoe die zich verhoudt tot de betreffende problematiek. In de tweede plaats is "realistische synthese" gediend bij inbreng vanuit de realiteit, *inclusief* praktijkkennis.

Het is dus niet eenvoudig zo dat "de" wetenschap de praktijk kan informeren over wat de meest effectieve en efficiënte omgang met maatschappelijke problemen is, zoals de oorspronkelijke, enigszins naïeve opvatting over evidence-based handelen wil (zie Wildavsky, 1979; Lindblom, 1990; Fischer 1995). Deze vaststelling vergt nadere reflectie – en niet alleen over het missie statement van evidence-based handelen. Immers, een meer dialectische verhouding tussen evidence en praktijk leidt ook tot nieuwe vragen. Hoe voorkomen we dat de praktijken meer de evidence gaan bepalen dan omgekeerd? Wat is, voor "realistische synthese", de ideale verhouding tussen de ontwikkeling van evidence-based richtlijnen en de praktijken die ze beogen te beïnvloeden?

BIBLIOGRAFIE

AIVD (2006), *De gewelddadige jihad in Nederland. Actuele trends in de islamitisch-terroristische dreiging*. Den Haag: Algemene Inlichtingen- en Veiligheidsdienst.

Baan, C.A., J.B.F. Hutten & P.M. Rijken (2003), *Afstemming in de zorg: Een achtergrondstudie naar de zorg voor mensen met een chronische aandoening*. Bilthoven: RIVM.

Buijs, F.J., F., Demant & A. Hamdy (2006a), *Strijders van eigen bodem. Radicale en Democratische moslims in Nederland*. Amsterdam: Amsterdam University Press.

Burssens, D. (2007), Hoe evident is evidencebased beleid? In: *Alert*, 33, (3), 51-63.

Davies, H.T.O, Nutley, S.M. & Smith, P.C. (1999). What Works? The Role of Evidence in Public Sector Policy and Practice. In: *Public Money & Management*, 19 (1), 3-5.

Fischer, F. (1995). *Evaluating public policy*. Chicago: Nelson-Hall.

Gielen, A.J. (2008a). *Radicalisering en identiteit*. Amsterdam: Aksant.

Gielen, A.J. (2008b). *Religie en cultuur in de (Marokkaanse) opvoeding. Evaluatie cursus opvoedingsondersteuning in Stadsdeel Slotervaart*. Almere: A.G. Advies.

Gielen, A.J. & Junne, G. (2008), Evaluatie van antiradicaliseringsprojecten. Hoe meet je of radicalisering wordt tegengegaan? In: Verlet, D. & Devos, C. (red). *Over beleidsevaluatie: van theorie naar praktijk en terug*. Brussel: Studiedienst van de Vlaamse Regering, 131-162.

Gielen, A.J. (2009a). *Een kwestie van identiteit. Evaluatie training Identiteit & Weerbaarheid voor moslima's*. Almere: A.G. Advies.

Gielen, A.J. (2009b). *Uitgesproken meningen. Evaluatie van de religieus-seculierekring in Stadsdeel Slotervaart*. Almere: A.G. Advies.

Glaser, B.G. & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Piscataway/NJ: Aldine Transaction.

Gradener, J. (2009). Professionals leren niet uit een database. In: *Tijdschrift voor Sociale Vraagstukken*, 63(3), 22-24.

Grin, J. (2004). Health technology assessment between our health care system and our health. Exploring the potential of reflexive TA. In: *Poiesis & Praxis* 2 (2-3),157-174.

Grin, J. (2007), Van klassieke kennismaatschappij naar maatschappelijke wijsheid? In: Funtowicz, S., Goorden, L., Grin, J. & Leroy, P. *Wetenschap, maatschappij, politiek: wie stuurt wie?* Antwerpen/Delft: viWTA/Eburon, 49-68.

Groen, J. & Kranenberg, A. (2006). *Strijdsters van Allah. Radicale moslima's en het Hofstadnetwerk*. Amsterdam: Meulenhoff.

Keijsers, J., Bouter, L.M., Steenbakkers, W. & Meertens, R.M. (1989). Methodologische kwaliteit en onderlinge vergelijkbaarheid: van onderzoek naar de effectiviteit van rugscholen. In: *Ned. T. Fysiotherapie*, 99 (5), 112-116.

Kolmer, D. (2009). Wat werkt wordt bepaald door de context. In: *Tijdschrift voor Sociale Vraagstukken*, 63 (3), 25-27.

Lindblom, C.E. (1990). *Inquiry and Change. The troubled attempt to understand and shape society*. New Haven: Yale University Press.

Lousberg, M., van Hemert, D. & Langelaan, S. (2009). *Ingrijpen bij radicalisering. De mogelijkheden van de eerstelijnswerker*. Soesterberg: TNO Veiligheid en Defensie.

Lub, V. (2009). *Stimulering van maatschappelijke binding van jongeren. Een verkenning naar sociale interventies op het terrein van radicalisering en culturele spanningen*. Utrecht: Movisie.

Mackenbach, J. (2003). Meer aandacht voor preventie. In: *Tijdschrift voor sociale geneeskunde*, 81 (3), 121-122.

Oers, J. van (2002). *Gezondheid op koers? Volksgezondheid Toekomstverkenning*. Bilthoven: RIVM.

Pawson, R. (2002a). Evidence-based Policy: In Search of a Method. In: *Evaluation, 8 (1), 157 – 181*.

Pawson, R. (2002b). Evidence-based Policy: The Promise of `Realist Synthesis'. In: *Evaluation*, 8 (3), 340-358.

Reekum, R. van (2008). Wetenschappelijk gefundeerd beleid is een hype. In: *Tijdschrift voor Sociale Vraagstukken*, 62 (9), 24-27.

Rensen, P. & van der Kooij, A. (2009). Evidence based spook bestaat niet. In: *Tijdschrift voor Sociale Vraagstukken*, 63 (1-2), 22-25.

Slootman, M. & Tillie, J. (2006). *Processen van radicalisering. Waarom sommige Amsterdamse moslims radicaal worden*. Amsterdam: Instituut voor Migratie en Etnische studies.

Tonkens, E. (2008). *Mondige burgers, getemde professionals. Marktwerking en professionaliteit in de publieke sector*. Amsterdam: Van Gennep.

Vendrig, A. (1999), The prediction of outcome in the multimodel treatment of chronic back pain: the role of emotional distress and personality (proefschrift). Utrecht : Elinkwijk.

Vries, H. de (2003). Investeren in preventie, een uitdaging voor het volgende kabinet. In: *Tijdschrift voor Sociale Geneeskunde*, 81 (2), 125-126.

Vos, R., Houtepen, R. & Horstman, K. (2002). Evidence-Based Medicine and Power Shifts in Health Care Systems. In: *Health Care Analysis*, 10 (3), 319-328.

Wildavsky, A. (1979). *The Art and Craft of Policy Analysis*. Londom: McMillan.

Wit, P. de & Hoedeman, R. (2004). Minder medische consumptie – bedrijfsartsen benaderen gezondheidsproblemen anders. In: *Medisch Contact*, nr. 16.

Yanow, D. (2006), Evidence-based policy. In: Bevir, M.W. (red.), *Encyclopedia of Governance*. Newbury Park, CA: Sage.