

BELEIDSEVALUATIES: WE KUNNEN HET WÉL LEUKER MAKEN

P.W.C. Koning*

Samenvatting

Nut en noodzaak van beleidsevaluaties worden zelden betwist. De praktijk van de beleidsevaluatie blijft veelal achter bij de vooraf soms hooggespannen verwachtingen. Zowel onderzoekers als beleidsmakers zijn in de regel vervuld van goede bedoelingen, maar zij weten elkaar niet te vinden bij de evaluatie van beleidsevaluaties. In deze bijdrage passeren zeven oorzaken de revue, die verklaren waarom het mis kan gaan bij evaluaties. Tevens wordt aangegeven hoe het beter zou kunnen. Ter illustratie wordt verwezen naar evaluaties op het beleidsterrein van het Ministerie van Sociale Zaken en Werkgelegenheid, met name beleid dat is gericht op (re-)integratie van uitkeringsontvangers in de arbeidsmarkt.

Trefwoorden: beleidsevaluatie, re-integratiebeleid

1 Inleiding

Als één type onderzoek ambtelijke ontwerpers van beleid, beleidsmakers in het parlement en onderzoekers bij elkaar zou moeten brengen, dan zijn dit wel beleidsevaluaties. De potentiële maatschappelijke, politieke en wetenschappelijke verdiensten van beleidsevaluaties staan buiten discussie. De Tweede Kamer jaagt veel evaluatieonderzoek aan. Het gaat kennelijk ergens over. Voor de betrokken onderzoekers vormt het uitvoeren van beleidsevaluaties een belangrijke uitdaging. Welke variabelen zeggen iets over de effecten van op de rails gezet beleid, onder welke voorwaarden mag gevoerd beleid geslaagd heten, en wanneer en in hoeverre kunnen gemeten effecten eenduidig worden toegeschreven aan de inzet van beleidsmaatregelen? Ingewikkelde, boeiende vragen, waarop wetenschappers hun tanden kunnen stukbijten.

Nut en noodzaak van beleidsevaluaties worden zelden betwist. De praktijk van de beleidsevaluatie blijft veelal achter bij de vooraf soms hooggespannen verwachtingen. Een relatief gering aantal evaluaties scoort enerzijds hoog in termen van wetenschappelijke standaarden en wordt anderzijds door beleidsmakers als zinvol ervaren. Het evaluatielandschap kent een tweedeling. Aan de ene kant staan de evaluatie-experts, die vaak verongelijkt door het leven gaan. Zij gedragen zich als de beste stuurder aan wal, of hadden geen toegang tot gegevens die nodig waren voor een deugdelijke analyse. Aan de andere kant staan beleidsmakers, die stellen dat uitgevoerde evaluaties geen of onvoldoende oog hebben voor de context waarin besluitvorming plaatsvindt. Kortom: zowel onderzoekers als beleidsmakers zijn vervuld van goede bedoelingen, maar zij weten elkaar niet te vinden bij de evaluatie van beleidsevaluaties. De grondoorzaak is dat zij verschillende succescriteria hanteren.

Deze observatie is niet nieuw. Zo maakt de Wetenschappelijke Raad voor het Regeringsbeleid (2011) zich al enkele jaren sterk voor het project ‘Lessen van evaluaties’. De vraagstelling van dit project is constructief: ‘Hoe kan de overheid de kracht van beleidsevaluaties beter benutten en de zwaktes vermijden?’. Nog iets concreter:

‘De vraag is hoe kan worden voorkomen dat evaluatie ten onder gaat in het politieke gewoel. Hoe kan de overheid de evaluatiefunctie zo inrichten dat haar kernfuncties, zoals verantwoording afleggen en lessen leren, zo goed mogelijk worden vervuld, terwijl tegelijk onbedoelde gevolgen van evaluatie zoveel mogelijk worden voorkomen,

zoals een doorgeschoten afrekencultuur of overspannen verwachtingen van wat de overheid vermag?’

Het bedoelde WRR-project loopt nog. Het betreft ‘werk in uitvoering’. Over de oorzaken van deze lange doorlooptijd kan rijkelijk worden gespeculeerd. Een flauwe grap maken is verleidelijk. Tijdigheid is tenslotte een belangrijk criterium voor een goede beleidsevaluatie. De positieve kant van het trage verloop van dit project is dat buitenstaanders nog over de mogelijkheid beschikken om de WRR van enige input te voorzien – en dat is precies de bedoeling van deze bijdrage. Ik som zeven oorzaken op, waarom het mis kan gaan bij evaluaties. Tevens geef ik aan waar en hoe het beter zou kunnen. Wellicht kan de WRR hier nog zijn voordeel mee doen.

2 Evalueren is terugkijken

Kenmerkend voor evaluaties (ex post) is dat zij terugkijken op gevoerd beleid. Volgen kabinetten elkaar snel op, dan is de koppeling van beleidsresultaten naar bewindspersonen minder evident. Het komt daarom geregeld voor dat bewindslieden een evaluatie terzijde schuiven met het argument dat inmiddels ‘heel ander beleid’ wordt gevoerd. Niet gespeend van een vleugje opportunisme kan het ook slim zijn resultaten juist te omarmen, wanneer dat goed uitkomt, vooral als een beleidsinitiatief niet gewerkt heeft in het verleden en ‘om die reden het beleid dan ook reeds is bijgesteld’.

Voor onderzoekers is zo’n opstelling natuurlijk niet echt bevredigend. Zij voelen zich een speelbal van beleidsmakers, soms worden ze genegeerd en soms doodgeknuffeld. Bovendien wijzen onderzoekers er – veelal terecht – op dat beleid echt niet zo dramatisch verandert in de loop van de tijd en dat het verleden dus wel degelijk relevante lessen leert. Zo is het wat merkwaardig te denken dat Instroom-Doorstroombanen in tegenstelling tot de voormalige Melkertbanen wél tot een forse doorstroom van uitkeringsontvangers naar reguliere arbeid zullen leiden. Het gaat bij beide programma’s immers om mensen uit dezelfde doelgroep. Bovendien rapporteert de internationale literatuur bijna als een wetmatigheid dat dit soort banen de arbeidsmarkt niet efficiënter maakt.

Dat onderzoekers bij evaluaties achterom kijken, biedt een belangrijk voordeel. Hebben de uitkomsten van de evaluatie betrekking op de resultaten van ‘oud beleid’, dan ligt afrekenen op huidig beleid niet voor de hand. Gevolg is dat alle aandacht uit kan gaan naar de ‘werkzame en niet-werkzame bestanddelen’ van eerder ingezet beleid. Evaluaties krijgen dan een constructiever karakter. Bij de praktijk van aanpassing van arbeidsomstandigheden (arbo) en in de re-integratiepraktijk is dit van belang: wat werkt voor wie en wanneer?

Resultaten van onderzoek op dit terrein voeden een kennisbasis die tijdloos is en zich niet zoveel aantrekt van de nukken van de komende en gaande beleidsmaker. Uiteindelijk kunnen beleidsmakers van dit type beleidsonderzoek profiteren. Mijn pleidooi is dus om beleidsevaluaties vooral in te zetten op dossiers waarbij kennisopbouw essentieel is. Discussies over de zin en onzin van nieuw beleid gaan dan uit van een deugdelijke kennisbasis, en zijn niet gebaseerd op de eenzijdige vraag of ‘beleid X’ nu werkt of niet. Zo zit de wereld ook niet in elkaar.

3 Beleid evalueren is een vak

Ogenschijnlijk telt Nederland voldoende instanties die er alles aan gelegen is tot goede evaluaties te komen: de Algemene Rekenkamer (AR), inspecties, het Ministerie van Financiën, financiële directies bij beleidsministeries en de planbureaus. Daarnaast is, zoals eerder gesteld, de Tweede Kamer een belangrijke – zo niet de belangrijkste – aanjager en opdrachtgever van evaluatieonderzoek. Al deze partijen hebben een heilig geloof dat evaluaties het beleid kunnen versterken. Maar dat niet alleen: er leeft ook een sterke

overtuiging dat zonder het bestaan van deze instanties de evaluatiefunctie (sterk) zou verslappen. Veel tijd gaat daarom op aan het ‘achter de broek zitten’ van beleidsmakers.

Juist op dit laatste punt gaat het volgens mij ook vaak fout: te veel nadruk ligt op de vraag óf en wanneer er evaluaties zijn, niet hoe die evaluaties plaats moeten vinden en of evaluaties wel altijd tot zinvolle resultaten leiden. De kunst van het evalueren verschilt wezenlijk van het uitoefenen van controle of houden van toezicht. Bij beleidsevaluaties draait het om het opzetten van geloofwaardige onderzoekontwerpen, en niet primair om de vraag of financiële middelen rechtmatig zijn besteed. Van origine zijn instanties als de AR en inspecties juist bevolkt met mensen die opgeleid zijn als jurist, bestuurskundige of financieel controller – zij komen vooral van de laatste school. Voor hen is het lastig te doorgronden wat een deugdelijke evaluatie is – liefst natuurlijk door groepen die deelnemen aan een programma te vergelijken met controlegroepen die er niet aan hebben deelgenomen. Maar hoe regel je zoiets?

Het gevolg is dat moeizame discussies ontstaan met beleidsmakers. De toezichthouders willen wel iets van de grond krijgen, maar het hoe en wat staat hen niet precies voor ogen. Want waarom zou een ‘nulmeting’ altijd zinvol zijn, welke variabelen wil je meten, wanneer meet je echt toegevoegde waarde die ondubbelzinnig aan gevoerd beleid kan worden toegeschreven? Allemaal vragen die bij de direct betrokken ambtenaren een meer dan basale kennis van statistiek of econometrie vergen. En vragen waar beleidsmakers – de opdrachtgevers van veel evaluatieonderzoek – zelf eveneens het antwoord op schuldig plegen te blijven.

De oplossing voor dit probleem ligt voor de hand: zet meer en vaker evaluatie-experts in – in deeltijd of in voltijd – bij toezichthoudende instanties die ertoe doen. Voor mijn gevoel is in Den Haag geleidelijk aan een beweging zichtbaar in deze richting, maar het blijft een traag verlopend proces, zeker zolang onvoldoende massa van evaluatie-experts werkzaam is om echt een evaluatiecultuur tot stand te brengen. Het zou werkelijk helpen, wanneer experts op het terrein van beleidsevaluatie veel vaker over de vloer komen bij toezichthouders.

4 Evalueren is committeren

Steeds breder leeft het besef – niet alleen onder onderzoekers – dat beleidsexperimenten behulpzaam kunnen zijn bij het evalueren van beleid. Het idee is rechttoe-rechtaan: op willekeurige gronden, bijvoorbeeld het BSN-nummer¹ van mensen (even of oneven), vindt toewijzing van een beleidsinterventie plaats. Zodoende is selectie uitgesloten en kan de situatie van door het beleid geraakte groepen eenvoudig worden vergeleken met de situatie van controlegroepen, om langs deze weg de effectiviteit van beleid te meten. Beleidsexperimenten leveren daardoor uitkomsten op die vrijwel iedereen kan begrijpen. Dit contrasteert met allerlei geavanceerde, non-experimentele methoden die veel verbeelding vergen en tot allerlei hoofdbrekens leiden over hoe alsnog tot geschikte controlegroepen te komen.

Toch gaat het vaak om een andere reden mis. Zo helder als de spelregels voor de algemene opzet van experimenten zijn, zo lastig is het om die spelregels te implementeren. Dit heeft alles van doen met zogenaamde ‘*non-compliance*’: uitvoerders van het beleid die zich niet houden aan de gerandomiseerde toewijzing van interventies. In dat geval treedt mogelijk verwatering van effecten op: de door het beleid bestreken groep gaat meer op de controlegroep lijken, en andersom. Het is in deze gevallen op zijn minst noodzakelijk om te weten of, en zo ja bij wie *non-compliance* heeft plaatsgevonden – in dat geval kan met enkele veronderstellingen het echte effect ‘teruggerekend’ worden.

¹ BSN = Burger Service Nummer.

Experimenteel beleidsonderzoek staat in elk geval bij het Ministerie van Sociale Zaken en Werkgelegenheid (SZW) nog in de kinderschoenen. Recente ervaringen met gemeentelijke experimenten op het terrein van re-integratie stemmen daarbij nog niet tot volle tevredenheid. Bij sommige gemeenten was sprake van *non-compliance*, en dan nog wel zonder dat is geregistreerd in welke opzichten van de toegewezen groep uitkeringsontvangers is afgeweken.² Het gevolg laat zich raden: effectmetingen die moeilijk zijn te interpreteren en waarschijnlijk het werkelijke effect van ingezet beleid onderschatten.

Naar verwachting zal de komende jaren de inzet van experimenten op het domein van het Ministerie van SZW verder groeien. Daarmee komt dus een volgende uitdaging nog prangender in het vizier: commitment van uitvoerders bij op poten gezette evaluaties. Mijn inschatting is dat dat commitment cruciaal is om een blijvende cultuur van beleidsexperimenten te realiseren. Zonder commitment bestaat gerede kans dat resultaten onnodig teleurstellen, en daarmee neemt het draagvlak voor toekomstige experimenten verder af.

Het is zaak deze spiraal te voorkomen en zo nodig te doorbreken, maar hoe? Eén berijdbaar spoor bestaat eruit de prikkels tot commitment vanuit het ministerie te versterken, met name door cofinanciering van decentrale initiatieven uit de rijksbegroting uitsluitend toe te passen op voorwaarde dat beleidsevaluerend onderzoek deugdelijk is uitgevoerd.³ Wellicht helpt dit gemeenten om door te pakken. Een ander spoor is dat van ruchtbaarheid geven aan ‘*best practices*’, door goed onderzoek te benoemen en te promoten. Hopelijk leidt dit vervolgens tot een inktvlek-effect: steeds meer uitvoerders die ook mee willen doen.

5 Evalueren is abstraheren

Vanuit het standpunt van de rijksoverheid gezien, zijn beleidsevaluaties bij naar lokale en provinciale overheden gedecentraliseerd beleid vooral relevant om de rol van het betrokken ministerie als ‘systeemverantwoordelijke’ op effectiviteit te toetsen. Het gaat om het stellen van de juiste regels en randvoorwaarden. De centrale overheid heeft bij gedecentraliseerd beleid veel minder greep op de concrete keuze van instrumenten, die lokaal of provinciaal worden ingezet. Immers, het zijn uitvoerders – zoals klantmanagers, lokale organisaties en gemeenten – die hierover gaan. Bij het evalueren van arbeidsmarktbeleid leidt het beginsel van systeemverantwoordelijkheid bijvoorbeeld vaak tot onmin bij uitvoerders op decentraal niveau.

Zo dient het Ministerie van SZW zich eens in de zoveel tijd van de schone taak te kwijten om de algehele effectiviteit van re-integratiemiddelen vast te stellen. Om dat te doen is wel voldoende ‘massa’ nodig, dat wil zeggen de beschikbaarheid van voldoende gegevens om statistisch verantwoorde uitspraken te doen over het algehele beeld in Nederland. In de praktijk komt het er dan veelal op neer dat re-integratiestatistieken van het Centraal Bureau voor de Statistiek worden uitgelopen. De CBS-statistiek kent een grofmazige indeling van typen instrumenten. Wat vervolgens gebeurt, is dat klantmanagers bij het Uitvoeringsinstituut Werknemersverzekeringen (UWV) en bij de gemeenten zich niet of nauwelijks herkennen in het beleidsonderzoek van SZW. De uitkomsten zouden te abstract zijn, ze zeggen weinig of niets over de re-integratiepraktijk waarbij de uitvoerders ‘met de poten in de modder staan’, en te maken hebben met andere klanttypen en een andere lokale arbeidsmarkt.

Dit heeft natuurlijk iets onbevredigends: hoe goed evaluatieonderzoek ook is uitgevoerd, decentrale partijen vinden niet dat het hen verder brengt. Een zekere desinteresse en mate van cynisme bij de direct betrokkenen zijn hiervan het ongewenste gevolg.

Het merkwaardige is dat klantmanagers van het UWV en bij de gemeente doorgaans stellig geloven dat hun maatwerk uiteindelijk wel degelijk evalueerbaar is – of althans: dat de

² De Koning *et al.* (2014).

³ Ministerie van Sociale Zaken en Werkgelegenheid (2014).

wenselijkheid van dit beleid via onderzoek aannemelijk is te maken. Vanuit statistisch perspectief is dat echter onmogelijk: altijd zal enige abstractie nodig zijn om het *gemiddelde* effect voor een aantal cliënttypen en uiteenlopende omstandigheden verantwoord te kunnen schatten. Wat kan, is dat klantmanagers na verloop van tijd leren wat op hoofdlijnen de zogenaamde ‘werkzame bestanddelen’ van beleid zijn. Die kunnen zij als algemene richtlijn toepassen, maar ergens houden de inzichten op die van grootschalige beleidsevaluaties mogen worden verwacht. Mijn pleidooi is daarom uitvoerders beter te scholen in het trekken van lering uit evaluatieonderzoek. Dat betekent niet dat zij die lessen onverkort moeten overnemen, maar wel dat zij die als algemene richtlijnen zouden dienen te hanteren. Zolang dat niet gebeurt, blijft de afstand tussen ‘Den Haag’ en de lokale beleidspraktijk onnodig groot.

6 Evalueren is conserveren

Een betrekkelijk onomstreden uitgangspunt in de statistiek is dat een vergelijking van resultaten plaatsvindt met een ‘nulhypothese’, waarin een gegeven beleid geen effect heeft. Beleid zou dan moeten leiden tot significant andere uitkomsten. Dit betekent echter niet dat het effect van beleid ook substantieel moet zijn – een misverstand dat gauw opgeld doet (juist bij econometristen). Significantie wordt namelijk eerder gevonden bij onderzoek op basis van gegevens uit grote steekproeven. Of omgekeerd geredeneerd: soms zal het bijna onmogelijk zijn om een significant resultaat te vinden, louter en alleen omdat de bij de evaluatie gebruikte steekproef een te kleine omvang heeft.

Werken met nulhypothesen heeft als logische bijwerking dat evalueren neigt naar conservatisme: het vraagt nogal wat om te bewijzen dat iets werkt, en tot die tijd werkt het bij veronderstelling niet. Bij het formuleren van beleidsambities kan dit nogal frustrerend – en soms ook disciplinerend – werken. Zonder bewijs dat een beleidsmaatregel effectief is (kan zijn), zullen planbureaus en andere wetenschappelijke instanties dit nieuwe beleid afremmen. Wat daarom bij evaluaties al flink zou helpen, is naar hun toetsingskracht te kijken en verwachtingen te managen. Vragen zijn dan: wanneer volstaat het geëvalueerde beleid, wat hebben we nodig voor een gezaghebbende evaluatie? In dit verband kan het ook instructief zijn om de effectiviteit van alternatief beleid als referentiepunt te nemen – en dus niet zozeer een situatie zonder beleid.

Daarnaast kan het helpen om kosten en baten van beleid explicieter te maken, om zo – wanneer dat nodig is – de gevonden beperkte effectiviteit te kunnen relativiseren. Een treffend voorbeeld hierbij is de inzet van scholing en training voor werklozen, waarvoor slechts zelden volgens de standaardnorm significante effecten zijn gevonden. Zou echter het ‘*break-even*’ effect worden berekend dat nodig is om de kosten van een re-integratietraject goed te maken, dan zou snel duidelijk worden dat zeer omvangrijke databestanden nodig zijn om zo’n effect als significant te kunnen bestempelen. Dat zet schattingsresultaten toch al snel in een ander perspectief. Conclusie: onderdeel van evalueren is ook goed interpreteren. Dit brengt me als vanzelf op het volgende punt.

7 Evalueren is interpreteren

De standaardreflex bij vastgestelde geringe of negatieve effectiviteit van een maatregel is dat die maatregel beter kan worden afgeschaft. Die neiging bestaat zeker bij organisaties met een afrekencultuur, en dat in nog sterkere mate wanneer de ambtelijke en politieke leiding op zoek is naar bezuinigingen. Effectmeting kan evenwel slechts het begin van het verhaal zijn. Een interventie is niet te vergelijken met een pot pindakaas die overal hetzelfde smaakt, ongeacht in welke supermarkt hij is gekocht. Vooral uitvoerders zelf zijn medebepalend voor de effectiviteit van interventies – zij kunnen deze maken of breken. Dit

verhaal is dus complex. Het is zaak dat uitvoerders van harte aan beleid meewerken, willen optimale resultaten worden geboekt.

De notie dat deze vorm van commitment zo belangrijk is, heeft als logische implicatie dat deze factor expliciet onderdeel van evaluatieonderzoek moet zijn. Bij evaluatie-experts vraagt dit om blikverlegging. Als draagvlak onder de uitvoerders zo bepalend is voor het welslagen van beleid, dan betekent dit dat ‘harde’ kennis, eerder verkregen bij de evaluatie van soortgelijk beleid, soms misleidend kan zijn. Het is dan ook van belang om te doorgronden of alle relevante partijen – uitvoerders voorop – er een duidelijk belang bij hebben om beleid tot een succes te maken. Een treffend voorbeeld is vervat in het advies van de Sociaal Economische Raad (2002) over de Wet op de arbeidsongeschiktheidsverzekering. Het Centraal Planbureau (2002) kon destijds niet veel anders doen dan de verwachting uitspreken dat de *impact* van het SER-advies op het aantal uitkeringen beperkt zou zijn. Door de *sense of urgency* die partijen voelden, is nadien wel degelijk een fors effect bereikt.⁴

8 Evalueren is kiezen

Gezien vanuit het perspectief van wetenschappers hebben beleidsdoorlichtingen en evaluaties van de overheid veelal een megalomaan, ongericht karakter. Om een voorbeeld te noemen: de beleidsdoorlichting van de wet SUWI⁵ zou in beginsel de effectiviteit van het UWV, de Sociale Verzekeringsbank en de gemeenten – en veranderingen daarin – in beeld moeten brengen. Dat is een ondoenlijke opgave.

De gestelde beleidsvragen zijn doorgaans legitiem, maar zoals het nu gaat zitten er te weinig *checks* en *balances* in het systeem om evaluaties ook wetenschappelijk geloofwaardig te houden. Voor iedere vraag zou je moeten denken aan het onderzoeksdesign dat nodig is om het antwoord deugdelijk vast te stellen. De praktijk is echter dat beleidsmedewerkers allemaal kosteloos extra vragen kunnen invoegen in de startnotities van evaluaties en doorlichtingen. Het gevolg is een onontwarbare kluwen van niet-evalueerbare vragen. Wetenschappers houden zich daarom vaak afzijdig van een evaluatie: dat kan helemaal niet, dus daar ga ik mijn naam niet aan verbinden.

Ook hier ligt het recept voor verbetering voor de hand: er moet veel meer tegenwicht ontwikkeld worden bij het formuleren van evaluatievragen. Het helpt daarbij al als evaluatie-experts in een vroegtijdig stadium meedenken over wat zinvol en haalbaar is. Vervolgens zouden vragen te rangschikken zijn naar de mate van evalueerbaarheid, met experimenteel onderzoek als hoogste trede op de ‘evaluatieladder’. Dit alleen al zou sterk disciplinerend kunnen werken, zowel naar beleidsmakers als naar toezichthouders. Bovendien voorkomt deze aanpak dat ook goed onderzoek niet de erkenning krijgt die het verdient, aangezien het ondersneeuwt in de grote berg van resultaten die vooral stelen op ‘*expert guesses*’ of anekdotisch materiaal.

9 Slot

Evaluaties zijn onontbeerlijk voor goed beleid. Door aanpassingen van de huidige toetsingspraktijk wordt evalueren leerzamer en een inspirerende ervaring. Een zekere ontspanning in de onderlinge verhouding van beleidsmakers en onderzoekers zou helpen om grotere stappen in de goede richting te zetten. Beleidsmakers, toezichthouders en onderzoekers zullen af en toe naast hun schaduw moeten stappen. Als vanzelf worden beleidsevaluaties dan leuker.

⁴ Fenger en Koning (2014).

⁵ SUWI = Structuur Uitvoering Werk en Inkomen.

Pierre Koning

* *Chief Science Officer* bij het Ministerie van Sociale Zaken en Werkgelegenheid en bijzonder hoogleraar Arbeidsmarkt en Sociale Zekerheid aan de Vrije Universiteit Amsterdam.

LITERATUUR

- Centraal Planbureau (2002), *Analyse van het SER-advies 'Werken aan arbeidsgeschiktheid'*, CPB Notitie 18 maart, Den Haag: CPB
- Fenger, H.J.M., en P.W.C. Koning (2014), Policy analysis by negotiation: trade unions, employers' organisations and the Social and Economic Council, in F. van Nispen en P. Scholten (Eds.), *Policy Analysis in the Netherlands*, Bristol: Policy Press - International Library of Policy Analysis, 193-204
- Koning, J. de, P. de Hek, L. Mallee, F. Rosing, en M. Groenewoud (2014), *Uitkomsten en ervaringen experimenten netto-effectiviteit re-integratie*, Amsterdam: SEOR/Regioplan
- Ministerie van Sociale Zaken en Werkgelegenheid (2014), *Kennisontwikkeling voor gemeentelijke uitvoeringspraktijk*, 2014-0000152134
- Sociaal Economische Raad (2002), *Werken aan arbeidsgeschiktheid. Voorstellen WAO-beleid*, Den Haag: SER
- Wetenschappelijke Raad voor het Regeringsbeleid (2011), *Werkprogramma 2011-2012*, Den Haag: WRR