

Praktijkevaluatie in de geestelijke gezondheidszorg: onvolledige en afwijkende cyclusvormen

door A.A.J.M. Beenackers

Samenvatting

In de geestelijke gezondheidszorg hebben we te maken met vele vormen van praktijkevaluatie. In een onlangs in dit tijdschrift gepubliceerd artikel hebben we al die vormen teruggebracht tot dezelfde grondvorm: een cyclisch proces van exploratie, normering, operationalisering, toetsing en bijsturing. Thans presenteren we een typologie van evaluatieprocessen die niet aan dit model beantwoorden. Het model zelf gaat dan als norm fungeren, en de evaluatieprocessen die niet aan deze norm voldoen, worden als minder adequaat beschouwd. Aan de hand van voorbeelden laten we tevens zien hoe zo'n minder adequaat evaluatieproces stap voor stap getransformeerd kan worden in een evaluatieproces dat wél aan de gestelde norm beantwoordt.

Inleiding

De laatste decennia is er een grote verscheidenheid aan evaluatievormen ontstaan: goal evaluation, goal-free evaluation, goal-fixed evaluation, goal-attainment evaluation, input evaluation, output evaluation, effort evaluation, impact evaluation, accountability evaluation, etc. etc. Publikaties, zoals onlangs van Sijben (1987) en Broekman, Schaap en Schippers (1987), laten zien hoe moeilijk het is zich een duidelijk beeld te vormen van de essentie van praktijkevaluatie, en van de verschillen die er tussen de diverse varianten bestaan. In een tweetal artikelen proberen we hierin verandering te brengen: in het eerste artikel – dat eerder in dit tijdschrift gepubliceerd werd (Beenackers 1987) – lag daarbij de nadruk op de essentie van praktijkevaluatie; in dit tweede artikel – dat daar een vervolg op is – bekijken we de verschillen die er tussen de diverse varianten bestaan.

Om de essentie van praktijkevaluatie zichtbaar te maken, hebben we in ons eerste artikel de veelheid van evaluatievormen teruggebracht tot een eenheid. We hebben daartoe een evaluatiemodel gepresenteerd, dat beschouwd kan worden als een abstractie van allerlei vormen van praktijkevaluatie, althans voor zover deze in overeenstemming zijn met het terugkoppelingsprincipe uit de cybernetica.

We hebben dit de modelcyclus genoemd. In dit tweede artikel presenteren we een typologie van evaluatieprocessen die niet aan de modelcyclus beantwoorden. De modelcyclus gaat dan als norm fungeren, en alle cyclusvormen die niet aan deze norm voldoen, worden als onvolledig of afwijkend getypeerd. Aan de hand van voorbeelden laten we tevens zien hoe een onvolledige of afwijkende cyclus getransformeerd kan worden in een modelcyclus. Aldus kan deze typologie ook een hulpmiddel zijn bij het oplossen van problemen die zich bij lopende evaluatieprocessen voordoen.

Alvorens de onvolledige en afwijkende cyclusvormen te beschrijven, recapituleren we uit ons eerste artikel de systematiek van de modelcyclus en de betekenis van het begrip 'structurele grondslag'.

De modelcyclus

Evaluatie is een vorm van terugkoppeling. Terugkoppeling houdt in dat van een bepaalde activiteit, die achtereenvolgens is geprogrammeerd en uitgevoerd, het bereikte resultaat wordt vergeleken met het beoogde resultaat. Indien het bereikte resultaat niet gelijk is aan het beoogde resultaat, wordt de activiteit opnieuw geprogrammeerd en uitgevoerd, waarna vervolgens het bereikte resultaat weer vergeleken wordt met het beoogde resultaat. Enzovoort: net zolang tot het bereikte resultaat gelijk is aan het beoogde resultaat. We hebben dan te maken met een cyclisch proces van programmering, uitvoering en terugkoppeling. Bij dit proces is in principe een drietal instanties betrokken: een instantie die het programma vaststelt: *de beleidsinstantie* (B); een instantie die het programma uitvoert: *de uitvoerende instantie* (U); en een instantie die het resultaat terugkoppelt van de uitvoerende instantie naar de beleidsinstantie: *de terugkoppelende instantie* (T). De modelcyclus heeft uitsluitend betrekking op de evaluerende instanties (B en T); de rol van de geëvalueerde instantie (U) wordt besproken in het kader van de structurele grondslag.

Bij de modelcyclus is de beleidsvorming (B) opgesplitst in een drietal fasen: exploratie (fase 1), normering (fase 2) en bijsturing (fase 5). De terugkoppeling (T) is opgesplitst in een tweetal fasen: operationalisatie (fase 3) en toetsing (fase 4). We krijgen dan het in figuur 1 weergegeven evaluatiemodel. Ieder evaluatieproces dat volgens dit grondpatroon is opgebouwd, beschouwen we als een modelcyclus. Cycli die niet volgens dit grondpatroon zijn opgebouwd, noemen we onvolledig (indien bepaalde fasen ontbreken) of afwijkend (indien de verschillende fasen niet in de juiste volgorde staan).

De structurele grondslag

Onze typologie van evaluatieprocessen (de verschillende cyclusvormen) is gebaseerd op methodologische aspecten: aan organisatorische

Figuur 1: De modelcyclus

Fase 1	Instantie B	Exploratie: het beschrijven van de beoogde praktijk
Fase 2	Instantie B	Normering: het formuleren van een aantal normen, waaraan de beoogde praktijk moet voldoen
Fase 3	Instantie T	Operationalisatie: het vertalen van de normen in toetsbare criteria
Fase 4	Instantie T	Toetsing: het aan de hand van gegevens, die systematisch worden verzameld en verwerkt, nagaan of aan de criteria wordt voldaan
Fase 5	Instantie B	Bijsturing: het bijsturen van de praktijk, indien niet aan de criteria wordt voldaan

aspecten wordt voorbijgegaan, terwijl in de praktijk steeds weer blijkt dat het juist de organisatorische context is die in belangrijke mate bepaalt wat methodologisch haalbaar is. Het begrip 'structurele grondslag' voorziet in deze leemte.

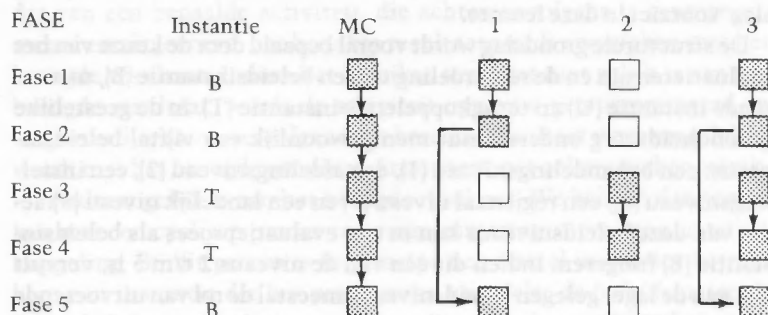
De structurele grondslag wordt vooral bepaald door de keuze van het evaluatieterrein en de rolverdeling tussen beleidsinstantie (B), uitvoerende instantie (U) en terugkoppelende instantie (T). In de geestelijke gezondheidszorg onderscheidt men gewoonlijk een vijftal beleidsniveaus: een behandelniveau (1), een afdelingsniveau (2), een instellingsniveau (3), een regionaal niveau (4) en een landelijk niveau (5). Ieder van deze beleidsniveaus kan in een evaluatieproces als beleidsinstantie (B) fungeren. Indien dit één van de niveaus 2 t/m 5 is, vervult één van de lager gelegen beleidsniveaus meestal de rol van uitvoerende instantie (U). Uitzonderingen daargelaten, zijn op niveau 1 (behandelniveau) B en U in één persoon verenigd: de behandelaar die de behandeling zowel vaststelt (B) als uitvoert (U). Van bijzonder belang is voorts de positie die de terugkoppelende instantie (T) inneemt ten opzichte van zowel B als U. Een goede structurele grondslag vereist dat al deze relaties expliciet en goed geregeld zijn. Indien dat niet het geval is, is het haast onmogelijk een modelcyclus tot stand te brengen, ook al doet de onderzoeker (de terugkoppelende instantie) nog zo zijn best. Ter adstructie van deze laatste stelling verwijzen we naar de ervaringen die Sijben (1987) onlangs in zijn proefschrift heeft beschreven.

De onvolledige cyclus

In het voorgaande hebben we gesteld dat als één of meerdere fasen ontbreken, we te maken hebben met een onvolledige cyclus. Men kan hierbij een tweetal hoofdtypen onderscheiden, welke rechtstreeks terug te voeren zijn op de rolverdeling tussen beleidsinstantie (B) en terugkoppelende instantie (T). Zoals we zojuist gezien hebben, komen de fasen 1 (exploratie), 2 (normering) en 5 (bijsturing) voor rekening van de beleidsinstantie, terwijl fasen 3 (operationalisatie) en 4 (toet-

sing) uitgevoerd worden door de terugkoppelende instantie. Bij een onvolledige cyclus is over het algemeen óf de terugkoppeling deficiënt (type 1), of de beleidsvoering (type 2). In het eerste geval hebben we te maken met beleidsvoering zonder terugkoppeling (BzT); in het tweede geval met terugkoppeling zonder beleidsvorming (TzB). Een combinatie van deze twee hoofdtypen geeft twee onvolledige cycli te zien, die samen geen volledige cyclus vormen, omdat ze inhoudelijk weinig met elkaar te maken hebben: beleidsvoering (B) en terugkoppeling (T) staan los van elkaar (type 3). In figuur 2 zijn deze onvolledige cyclusvormen schematisch weergegeven. Ter vergelijking is ook de modelcyclus (MC) afgebeeld. Ieder type zullen we afzonderlijk bespreken.

Figuur 2: Modelcyclus (MC) versus onvolledige cyclusvormen: (1) beleidsvoering zonder terugkoppeling (BzT); (2) terugkoppeling zonder beleidsvorming (TzB); (3) gebrek aan integratie (BzT & TzB).



Inhoudelijk en proportioneel consistent – De modelcyclus is een volledige cyclus, dat wil zeggen dat alle fasen aanwezig zijn. Een volledige cyclus is zowel proportioneel als inhoudelijk consistent. Proportioneel consistent betekent dat de verschillende fasen gelijkmatig tot ontwikkeling gebracht zijn: ze zijn als het ware van gelijke omvang. Van proportionele inconsistentie is sprake als er bijvoorbeeld vele criteria geformuleerd worden, terwijl de praktijk slechts aan enkele daarvan wordt getoetst; of ingeval er vele criteria zijn waaraan niet wordt voldaan, terwijl slechts ten aanzien van enkele daarvan wordt bijgestuurd.

Inhoudelijk consistent betekent dat de inhoud van de verschillende fasen met elkaar in overeenstemming is. Als bij voorbeeld de normering betrekking heeft op de produktiviteit, hebben ook alle andere fasen betrekking op de produktiviteit. Indien in zo'n geval de toetsing betrekking zou hebben op de kwaliteit, zou er sprake zijn van inhoudelijke inconsistentie. Iedere cyclus die zowel proportioneel als inhoudelijk consistent is, is een volledige cyclus; iedere cyclus die óf proportioneel óf inhoudelijk inconsistent is, is een onvolledige cyclus. Een

voorbeeld van een volledige cyclus is de wijze waarop de aanwending der financiële middelen wordt geëvalueerd. De jaarlijkse begroting is een beschrijving van de beoogde aanwending der financiële middelen (normering), terwijl de jaarrekening achteraf aangeeft hoe de financiële middelen daadwerkelijk zijn aangewend (toetsing). Ingeval er een verschil tussen begroting en jaarrekening ontstaat (of dreigt te ontstaan), wordt er door de desbetreffende beleidsinstantie bijgestuurd. Het systeem is zowel proportioneel als inhoudelijk consistent. We hebben daarom te maken met een volledige cyclus. Omdat alle fasen ook in de juiste volgorde staan, is deze volledige cyclus tevens een modelcyclus.

Beleidsvoering zonder terugkoppeling (BzT) – In alle gevallen waarin een praktijksituatie wordt vormgegeven, zonder dat er systematisch gegevens verzameld worden door middel waarvan de desbetreffende praktijksituatie wordt getoetst aan vooraf vastgestelde criteria, is er sprake van beleidsvoering zonder terugkoppeling (BzT). De fasen 3 (operationalisatie) en/of 4 (toetsing) zijn in dat geval onderontwikkeld. We hebben dan te maken met een vorm van proportionele inconsistentie.

Van BzT is sprake in al die situaties waarin beleidsmaatregelen genomen worden, zonder dat er adequaat wordt nagegaan of deze maatregelen het beoogde effect hebben. Alle beleidsinstanties die te maken hebben met een eindprodukt dat niet exact gedefinieerd is, gaan waarschijnlijk min of meer aldus te werk. Controle op het eindprodukt is dan immers niet eenvoudig, en wellicht is dat een belangrijke overweging om er maar van af te zien. Maar men kan dit ook omkeren: indien het eindprodukt niet gecontroleerd wordt, is het niet nodig om het exact te definiëren. Bij een dergelijke omkering wordt ervan uitgegaan dat het eindprodukt wel te definiëren is, maar dat men daartoe kennelijk onvoldoende gemotiveerd is. Leenen (1979) heeft dit als volgt geformuleerd: 'De stelling dat men kwalitatief goede zorg verleent, impliceert dat men standaarden en normen heeft waarop deze uitspraak is gebaseerd; bestaan die standaarden en normen niet, dan kan men ook geen oordeel over de kwaliteit van het eigen handelen geven' (933). De uitspraak van Leenen heeft betrekking op de in Nederland toegepaste vorm van intercollegiale toetsing, maar ze is uiteraard ook van toepassing op andere vormen van terugkoppeling. Aldus geredeneerd moeten er nog andere factoren zijn, die ertoe bijdragen dat beleidsinstanties vaak zo weinig belangstelling hebben voor terugkoppeling. Een aantal van deze factoren is van principiële aard, of wordt althans zo gepresenteerd. Een voorbeeld hiervan is de professionele autonomie van de hulpverlener, welke zich niet met bepaalde vormen van terugkoppeling zou verenigen. Volgens Raat en Goudriaan (1985) speelt ook deze factor een belangrijke rol bij de in Nederland gekozen vorm van intercollegiale toetsing.

Daarnaast zijn er factoren die maken dat terugkoppeling niet strikt noodzakelijk is. In de allereerste plaats speelt het ontbreken van het marktmechanisme daarbij een belangrijke rol. Omdat het marktmechanisme ontbreekt, wordt het beleid niet primair bepaald door de eigenschappen van het produkt, maar door allerlei andere factoren. Voor wat het overheidsbeleid betreft, wijzen Van Houten en Simonis (1984) in dit verband onder andere op de rol van politieke, bureaucratische en maatschappelijke factoren. En voor zover kennis daarbij een rol speelt, gaat het volgens De Jong (1984) niet alleen om objectieve kennis, maar ook om common sense-opvattingen en de publieke opinie. In de kwartaire sector is het beleid vaak primair het resultaat van onderhandelingen tussen belangengroeperingen: het gaat om het volgens bepaalde procedures bereiken van een compromis, en objectieve informatie over het geleverde produkt speelt daarbij een ondergeschikte rol.

Niet controleren kan daarnaast ook voordelen opleveren: wat niet gecontroleerd wordt, kan ook niet als onvoldoende aangemerkt worden. Tjeenk Willink, tot voor kort regeringscommissaris voor de reorganisatie van de rijkdienst, wijst er in zijn Jaarbericht 1985 op, dat een dergelijke controle op het eindprodukt bij de overheid in het algemeen ontbreekt. Zelfs voor de uitvoering van het beleid bestaat geen belangstelling: 'Ministers en parlementsleden hebben vooral belangstelling voor nieuw beleid. Nieuwe ontwikkelingen volgen elkaar in snel tempo op. Nieuwe beslissingen worden genomen, terwijl de voorafgaande nog nauwelijks zijn verwerkt' (p. 14). Hupe (1984) memoreert dat uitspraken van bewindslieden in het algemeen door het publiek als beleidsintenties opgevat worden, en dat – ook al is men het niet met deze uitspraken eens – men er in het algemeen van uitgaat dat de effecten van uitgevoerd beleid in overeenstemming zijn met de beleidsintenties (p. 287). Indien dit juist is, is het te begrijpen waarom er bij politici zo weinig belangstelling bestaat voor de beleidseffecten: bij een evaluatie daarvan wordt het voordeel van de twijfel ingewisseld tegen het risico dat men door de mand valt. En wat voor politici geldt, geldt waarschijnlijk voor alle beleidsinstanties die met een niet exact gedefinieerd eindprodukt te maken hebben. Wie de literatuur op het gebied van de ggz erop naleest – en daartoe verwijzen we naar Smits (1984) – komt precies dezelfde factoren tegen als die welke we zojuist genoemd hebben: beleidsinstanties hebben weinig belangstelling voor terugkoppeling omdat het moeilijk is, omdat het niet strikt noodzakelijk is, en omdat er weinig eer aan te behalen valt.

Terugkoppeling zonder beleidsvorming (TzB) – In alle gevallen waarin voor evaluatieve doeleinden gegevens worden verzameld, verwerkt en gerapporteerd, terwijl de desbetreffende praktijksituatie vooraf niet is genormeerd en/of achteraf niet wordt bijgestuurd, is er sprake van terugkoppeling zonder beleidsvorming (TzB). De fasen 1 (exploratie), 2 (normering) en/of 5 (bijsturing) zijn dan onderontwikkeld. We hebben

ook hier te maken met een vorm van proportionele inconsistentie.

Van TzB is sprake in al die situaties waarin gegevens worden verzameld c.q. onderzoek wordt verricht, terwijl deze gegevens c.q. resultaten geen of nagenoeg geen rol spelen bij het bepalen van het beleid. Waarschijnlijk is dit een algemeen probleem, dat zich voordoet zowel binnen de gezondheidszorg (Van Etten 1984; De Jong 1984), als in vele andere sectoren (Hupe 1984; Van Houten en Simonis 1984). En ook hier geldt dat de kern van het probleem gelegen is in het niet exact definiëren van het eindprodukt. Ook binnen de ggz zien we dat de teruggekoppelde gegevens slechts in geringe mate de beleidsvorming beïnvloeden. Van der Grinten (1983) komt na het bestuderen van 25 onderzoeken, die alle in Nederland uitgevoerd waren (18) of werden (7), tot de conclusie 'dat bescheidenheid ten aanzien van de betekenis van onderzoek voor het beleid gepast is' (p. 835).

Behalve met onderzoek hebben we in de ggz ook te maken met informatiesystemen. Dekker (1984) stelt zelfs 'dat de gezondheidszorg al verschillende terugkoppelingsmechanismen kent, waarvan registratie en het machtigingensysteem de belangrijkste zijn'. Maar, aldus Dekker, 'deze systemen worden niet of nauwelijks gebruikt om informatie terug te voeren met de bedoeling om de hulpverlening of het beleid bij te sturen' (p. 30). Elders hebben we uiteengezet dat dit ook volstrekt onmogelijk is (Beenackers 1983). Ook Bayer (1984) heeft hier onlangs op gewezen: het formuleren van doelstellingen, en het op basis van deze doelstellingen ontwikkelen van normen, zijn essentiële voorwaarden om verzamelde gegevens tot beleidsinformatie te kunnen verwerken (p. 275). En omdat aan deze voorwaarden niet wordt voldaan, is het volgens Bayer ook geenszins verwonderlijk dat de registratie voor evaluatiedoeleinden nauwelijks bruikbaar is (p. 269).

In de Verenigde Staten is de situatie overigens niet veel beter. De laatste jaren wordt in Nederland veelvuldig aandacht gevraagd voor de door de Community Mental Health Centers (CMHC's) toegepaste vorm van 'program evaluation' (o.a. Van der Grinten 1983; Schnabel 1984; Dekker 1984). Door DiNitto, McNeece en Johnson (1984) is bij 42 CMHC's nagegaan welke rol evaluatieresultaten spelen bij het nemen van beleidsbeslissingen. Daaruit bleek dat van de 140 gerapporteerde programmawijzigingen er 16 toegeschreven werden aan de evaluatieresultaten. Volgens de auteurs is dit erg weinig: 'data on program efficiency or effectiveness were *never* mentioned as a reason for eliminating a service and only rarely as a reason for modifying a service' (p. 182). Ook in de Verenigde Staten geldt dat terugkoppelende instanties niet veel verder komen dan een TzB-cyclus. Maar ook het onderzoek van DiNitto e.a. is daar in feite een voorbeeld van: het omvat immers alleen de fasen 3 (operationalisatie) en 4 (toetsing).

Zoals ook door Bayer (1984) is opgemerkt, is normering een essentiële voorwaarde om gegevens te kunnen vertalen in beleidsmaatregelen. De normeringsfase is de flessehals van ieder evaluatieproces: in

deze fase moet het produkt gedefinieerd worden, of het produktieproces, of de produktiemiddelen, of de randvoorwaarden – als er maar érgens een begin gemaakt wordt. En gezien de huidige evaluatiepraktijk is dat kennelijk erg moeilijk. De meeste evaluatieprocessen sneuvelen juist in deze fase: evaluatie zonder normering levert immers altijd een TzB-cyclus op.

Gebrek aan integratie – In het voorgaande hebben we twee onvollledige cyclusvormen besproken: beleidsvoering zonder terugkoppeling (BzT) en terugkoppeling zonder beleidsvorming (TzB). Een combinatie van deze twee cyclusvormen levert géén volledige cyclus op, omdat de twee samenstellende delen niet op elkaar zijn afgestemd: de beleidsvoering en de terugkoppeling zijn dan twee autonome processen; ieder met een eigen inhoud. De inconsistentie is in dit geval niet proportioneel van aard (alle fasen zijn immers aanwezig), maar inhoudelijk. Type 3 is geen op zichzelf staand type, maar het gelijktijdig aanwezig zijn van de typen 1 (BzT) en 2 (TzB). Gebrek aan integratie betekent dat beleidsinstantie (B) en terugkoppelende instantie (T) er niet in slagen gezamenlijk één geïntegreerd evaluatieproces tot stand te brengen. Afgezien van de vele methodologische problemen, heeft dit vooral te maken met structurele problemen; met name aangaande de rolverdeling tussen beleidsinstantie (B), terugkoppelende instantie (T) en uitvoerende instantie (U). Schnabel (1984): 'De verzelfstandiging van de onderzoeksfunctie uit de beleidsfunctie heeft, net zoals eerder de verzelfstandiging van de beleidsfunctie uit de hulpverleningsfunctie, ertoe bijgedragen dat de relatie tussen de functies moeizaam is geworden. Onderzoek van voorzieningen en instellingen is niet meer per definitie beleidsrelevant of beleidsgericht, evenmin als het beleid in de instellingen of het beleid ten aanzien van de voorzieningen per definitie hulpverleningsgericht is of bijdraagt tot de genezing van de patiënt' (p. 103). En dat dit vele problemen geeft, blijkt wel uit de literatuur. Onderzoek dat dicht bij het veld wordt uitgevoerd, levert vooral 'resistentieliteratuur' op: artikelen waarin de onderzoeker zich afvraagt waarom zijn werk in het veld zoveel weerstanden oproept. Onderzoek dat op afstand wordt uitgevoerd, levert vooral 'existentieliteratuur' op: artikelen waarin de onderzoeker zich afvraagt of zijn werk voor het veld überhaupt wel bestaat. De eerste categorie artikelen gaat vooral over het problematische karakter van de relatie tussen terugkoppelende instantie (T) en uitvoerende instantie (U). Deze onderzoekers zitten kennelijk zo dicht op het veld dat de beleidsinstantie (B) slechts op de achtergrond aanwezig is. Voorbeelden hiervan zijn Weiss (1972), Zusman en Bissonnette (1973), Krowinsky en Fitt (1978) en Beenackers (1984a). De tweede categorie artikelen gaat vooral over het problematische karakter van de relatie tussen terugkoppelende instantie (T) en beleidsinstantie (B). Deze onderzoekers zitten kennelijk zo ver van het veld af dat de uitvoerende instantie (U) slechts op de achtergrond aanwezig is.

Voorbeelden hiervan zijn Schuyt (1984), Van Houten en Simonis (1984), Hupe (1984) en De Jong (1984).

Gezien het veelal problematische karakter van de samenwerking tussen B, T en U, kan gesteld worden dat een goede structurele grondslag een absolute voorwaarde is om met succes te kunnen evalueren. Een modelcyclus moet beschouwd worden als een coproductie van beleidsinstantie (B), terugkoppelende instantie (T) en uitvoerende instantie (U). Als één van deze instanties uit de boot valt, is de kans op succes nihil.

De afwijkende cyclus

Indien wél alle fasen aanwezig zijn – en het systeem zowel proportioneel als inhoudelijk consistent is – hebben we te maken met een volledige cyclus. Indien tevens alle fasen in de juiste volgorde staan, spreken we van een modelcyclus; staan de fasen niet in de juiste volgorde, dan hebben we te maken met een afwijkende cyclus. Van alle afwijkende cycli die theoretisch mogelijk zijn, is er maar één die voldoende realiteitswaarde heeft om besproken te worden; namelijk die waarbij de normering niet aan de toetsing voorafgaat, maar erop volgt. De volgorde der fasen is dan: (1)-(3)-(4)-(2)-(5). We onderscheiden daarbij een tweetal typen: de experimentele cyclus en de traditionele cyclus. Bij de experimentele cyclus wordt de norm gespecificeerd door de teruggekoppelde gegevens; bij de traditionele cyclus is dat niet het geval.

Het verschil tussen de modelcyclus en de twee afwijkende cyclusvormen is dus gelegen in het moment waarop de normering een feit is: bij de modelcyclus is dat het geval vóórdat begonnen wordt met het verzamelen van de gegevens, bij de experimentele cyclus is de normering een feit op het moment dat de gegevens verzameld en verwerkt zijn, en bij de traditionele cyclus kan de discussie over de normering pas beginnen nadat de onderzoeksuitkomsten aan de beleidsinstantie gerapporteerd zijn.

Voor ieder evaluatieproces geldt dat de verschillende fasen beschouwd kunnen worden als de schakels van een keten die zo sterk is als de zwakste schakel. Onvolledige cyclusvormen, zoals we die in de voorafgaande paragraaf besproken hebben, kunnen daarom nooit een adequaat evaluatieproces opleveren. Afwijkende cyclusvormen kunnen daarentegen geschikt zijn voor exploratieve doeleinden: voor het ontwikkelen van normen (zie Beenackers 1987). In deze paragraaf bekijken we daartoe eerst de experimentele cyclus, en vervolgens de traditionele cyclus.

De experimentele cyclus – In alle gevallen waarin de onderzoeksopzet zodanig is, dat de verkregen gegevens (uiteindelijk) de norm specificeren, spreken we van een experimentele cyclus. We zullen daar enkele voorbeelden van bespreken.

Indien men er ondanks alle genoemde risico's toch in slaagt een traditionele cyclus te voltooien, is de exploratieve betekenis ervan gelijk aan die van een geslaagde experimentele cyclus: in beide gevallen kan men dan vervolgens overschakelen op de modelcyclus. Een traditionele cyclus heeft alleen maar kans van slagen indien de verzamelde gegevens normatief geïnterpreteerd kunnen worden; dat wil zeggen dat als twee willekeurige uitkomstmogelijkheden met elkaar vergeleken worden, het voor de beleidsinstantie duidelijk is welke van de twee het meest gunstig is. De variabelen die bij satisfactiemetingen gebruikt worden, kunnen hier als voorbeeld dienen: als een onderzoeker aan 100 cliënten vraagt of ze tevreden zijn met de hulp die ze ontvangen hebben, gaat hij ervan uit dat als 72 cliënten hierop bevestigend antwoorden, dit een gunstiger resultaat is dan ingeval dit er 69 zijn. De normatieve interpretatie levert bij dit soort variabelen kennelijk geen probleem op. Normeren betekent in dit voorbeeld dat de beleidsinstantie bepaalt hoeveel cliënten er minstens tevreden moeten zijn om niet tot bijsturing te besluiten. En pas als er ook genormeerd en zonodig bijgestuurd is, is de cyclus voltooid. Uit onderzoeksverslagen kan niet altijd opgemaakt worden of men hier wél of niet in geslaagd is. Veel verslagen gaan immers niet verder dan de onderzoeksresultaten. Bij de traditionele cyclus betekent dit dat niet alleen de bijsturingsfase buiten het onderzoeksverslag blijft, maar ook de normeringsfase. Een voorbeeld hiervan is het satisfactieonderzoek van Bayer (1979). Alle door hem gehanteerde variabelen zijn normatief te interpreteren. Maar het onderzoeksrapport geeft geen antwoord op de vraag of er ook daadwerkelijk normering heeft plaatsgevonden. We weten dus niet of we hier met een traditionele cyclus te maken hebben of met een onvolledige cyclus.

Het transformeren van een onvolledige of afwijkende cyclus in een modelcyclus

Tot ongeveer tien jaar geleden geschiedde de terugkoppeling voornamelijk mondeling; meestal in de vorm van periodiek overleg. We hebben dan bijna altijd te maken met een BzT-cyclus. De laatste jaren wordt er ook op grote schaal schriftelijk teruggekoppeld; meestal in de vorm van registratiegegevens. We hebben dan bijna altijd te maken met een TzB-cyclus. Per saldo krijgen we voor de ggz als geheel een onvolledige cyclus van het derde type: gebrek aan integratie. Uitzonderingen daargelaten, geldt dit voor het reguliere systeem als geheel. Daarnaast vindt er incidenteel terugkoppeling plaats door middel van onderzoek. Als regel heeft dit de vorm van een afwijkende cyclus (zie paragraaf 5). Volgens deze diagnose zijn alle cyclusvormen aanwezig, behalve de modelcyclus.

Aan de hand van een drietal voorbeelden zullen we laten zien hoe een onvolledige of afwijkende cyclusvorm getransformeerd kan wor-

den in een modelcyclus. Vanwege de representativiteit zullen we van elk der drie genoemde terugkoppelingsvarianten een voorbeeld geven: terugkoppeling door middel van onderzoek (voorbeeld 1), door middel van registratie (voorbeeld 2) en door middel van overleg (voorbeeld 3). In het eerste voorbeeld gaan we uit van een afwijkende cyclusvorm; in het tweede voorbeeld van een TzB-cyclus; en in het derde voorbeeld van een BzT-cyclus.

Voorbeeld 1 – Op behandelingsniveau (niveau 1) is de behandelaar zowel beleidsinstantie (B) als uitvoerende instantie (U). En indien het ook de behandelaar is die de evaluatie uitvoert, fungeert hij eveneens als terugkoppelende instantie (T). Van belangentegenstellingen is dan geen sprake. Ieder evaluatiesysteem, dat gebaseerd is op het verschil tussen de feitelijke behandelingsuitkomst en de uitkomst die men zich vooraf tot doel heeft gesteld, is in zo'n geval gemakkelijk te transformeren in een modelcyclus. Als voorbeeld nemen we Goal Attainment Scaling (GAS). Bij GAS wordt het begrip 'uitkomst' als 'verwachting' gedefinieerd. Een goal attainment score geeft dus aan in hoeverre de verwachtingen zijn uitgekomen (o.a. Melief e.a. 1979). De modelcyclus vereist evenwel dat het begrip 'uitkomst' expliciet als 'doel' wordt gedefinieerd. Dit kan gerealiseerd worden door de goal attainment schalen, waarop het verschil tussen beoogde en gerealiseerde uitkomst wordt aangetekend, niet te construeren rond het niveau dat men met de behandeling *verwacht* te bereiken, maar rond het niveau dat men *wil* bereiken (Beenackers 1984b). We krijgen dan een modelcyclus; althans op behandelingsniveau. Iedere behandelaar weet dan immers, dat als de beoogde eindscore niet wordt gerealiseerd, of niet gerealiseerd dreigt te worden, hij zijn behandeling zo mogelijk moet bijsturen.

Maar ingeval het afdelingsmanagement zich niet met iedere behandeling afzonderlijk bemoeit, levert precies dezelfde procedure op afdelingsniveau (niveau 2) slechts een traditionele cyclus op. Het afdelingsmanagement moet dan immers achteraf – nadat de afdelingsresultaten bekendgemaakt zijn – eerst nog normen ontwikkelen, alvorens de resultaten te kunnen vertalen in beleidsmaatregelen. Omdat op afdelingsniveau het afdelingsmanagement als beleidsinstantie (B) fungeert, en de behandelaar als uitvoerende instantie (U), zou men op dit niveau bij voorbeeld een gemiddelde score per behandelaar als norm kunnen stellen. Als een behandelaar onder dat gemiddelde blijft, is het aan het afdelingsmanagement om de praktijk bij te sturen. Aldus levert de GAS-procedure ook op afdelingsniveau een modelcyclus op.

Maar ingeval de directie zich niet met iedere behandelaar afzonderlijk bemoeit, levert precies dezelfde procedure op instellingsniveau (niveau 3) slechts een traditionele cyclus op. Enzovoort. Als een evaluatieproces meerdere niveaus bestrijkt, moet de norm dus per niveau gedifferentieerd worden. En dat moet uiteraard zo gebeuren, dat er op

De wijze waarop de Inspectie dit preventieve toezicht uitoefent, is te karakteriseren als een BzT-cyclus. Dit toezicht wordt immers niet uitgeoefend door het systematisch verzamelen van gegevens, op grond waarvan de praktijk aan vooraf vastgestelde criteria wordt getoetst (zie de betreffende paragraaf).

Vanaf 1983 zien we dat hierin verandering komt. De Inspectie is toen begonnen met het ontwikkelen van zogenaamde 'referentiekaders' (GIGV 1984). Een referentiekader is een verzameling normen en aandachtspunten, op grond waarvan de praktijk wordt beoordeeld. Ze vormen de normatieve grondslag van het preventieve toezicht. In termen van ons model: ingeval we met een aandachtspunt te maken hebben, is er sprake van fase 1 (exploratie); ingeval we met een norm te maken hebben, is er sprake van fase 2 (normering).

De referentiekaders worden geschreven per categorie instellingen: één voor psychiatrische afdelingen van algemene ziekenhuizen (1983); één voor psychogeriatrische verpleeghuizen (1984); enzovoort. De verschillende referentiekaders worden zowel qua inhoud als vorm zo veel mogelijk op elkaar afgestemd, zodat er uiteindelijk een stelsel onderling samenhangende normen ontstaat, dat de hele ggz bestrijkt.

Ieder referentiekader wordt na verloop van tijd herschreven, zodat iedere versie in zekere zin beschouwd kan worden als een concept voor de volgende versie. Zo is de eerste versie van het referentiekader voor psychiatrische afdelingen van algemene ziekenhuizen inmiddels vervangen door een tweede versie (1985). Aldus krijgen we een cyclisch proces van exploratie (fase 1) en normering (fase 2), waarbij iedere cyclus twee à drie jaar duurt. In 1984 is er een begin gemaakt met het operationaliseren van de referentiekaders (fase 3). Het referentiekader psychogeriatrische verpleeghuizen – waartoe we ons in dit voorbeeld beperken – is daarbij vertaald in 169 criteria.

In 1985 is op beperkte schaal de toetsing van start gegaan (fase 4). Daarbij werd gewerkt met zogenaamde veldteams, bestaande uit 5 inspecteurs. Aldus werd per team, per dag één psychogeriatrisch verpleeghuis volledig doorgelicht. Voor het verzamelen van de gegevens werden onderling complementaire methoden gebruikt: observatie, interviews en documentenanalyse. De vijf verpleeghuizen bleken gemiddeld aan 114 criteria te voldoen: aan 44 criteria werd niet voldaan, terwijl ten aanzien van 11 criteria de uitslag onduidelijk of twijfelachtig was.

In beginsel is het zo, dat ten aanzien van ieder criterium waaraan bij de toetsing niet blijkt te worden voldaan, de instelling geadviseerd kan worden de praktijk bij te stellen (fase 5). Per verpleeghuis heeft de inspectie de resultaten op gemiddeld 25 criteria daadwerkelijk ter discussie gesteld. Gemiddeld werd per verpleeghuis ten aanzien van 13 criteria een min of meer concrete afspraak gemaakt over bijstelling van de praktijk.

Bij de fasen 4 en 5 was het de Inspectie die de praktijk van de ver-

pleeghuizen aan de criteria toetste, en indien de praktijk niet aan een criterium voldeed, de praktijk ter discussie stelde. Maar tegelijkertijd waren het de verpleeghuizen die voortdurend de criteria toetste aan de praktijk, en indien een criterium niet aan de praktijk voldeed, het criterium ter discussie stelde. Dit was geen vervelende bijkomstigheid, maar een essentieel onderdeel van het systeem. Het stelde de Inspectie namelijk in staat de criteria te verbeteren. Nadat cyclus 1 voltooid was, heeft de Inspectie daartoe 56 van de 169 criteria ter discussie gesteld. In overleg met de vijf getoetste verpleeghuizen werden vervolgens 40 van deze criteria geherformuleerd. Dit is fase 1 van cyclus II. De volgende stap is herschrijving van het referentiekader: fase 2. Enzovoort.

We zien hier hoe stap voor stap, fase na fase tot ontwikkeling gebracht wordt, precies volgens het schema van de modelcyclus. Iedere fase wordt consequent afgeleid uit de voorgaande fase. Het is tevens een voorbeeld van een systeem dat zijn invloed uiteindelijk op alle beleidsniveaus kan doen gelden. En omdat daarbij alle sectoren van de ggz bestreken worden, kan er aldus een complete evaluatiepiramide ontstaan, waarbij de beleidslijnen van boven naar beneden lopen, en de terugkoppellingslijnen van beneden naar boven. De rolverdeling wordt bepaald door de positie van de Inspectie, zoals deze wettelijk is geregeld. Het evaluatieterrein wordt gespecificeerd door de diverse wetten en regelingen, zoals deze door de daartoe bevoegde instanties zijn uitgevaardigd, en waarvoor geldt dat de Inspectie op de naleving ervan moet toezien. De rolverdeling en het evaluatieterrein vormen gezamenlijk de structurele grondslag.

Uitleiding

In een tweetal artikelen hebben we gesteld dat iedere vorm van evaluatie beschouwd kan worden als een cyclisch proces van exploratie, normering, operationalisatie, toetsing en bijsturing. We hebben daarbij een zestal cyclusvormen onderscheiden: een drietal volledige cycli (één modelcyclus en twee afwijkende cycli: de experimentele cyclus en de traditionele cyclus) en een drietal onvolledige cycli (beleidsvoering zonder terugkoppeling, terugkoppeling zonder beleidsvorming en gebrek aan integratie). Het voordeel van deze typologie is, dat deze normatief van aard is, waardoor het relatief eenvoudig wordt evaluatieprocessen te evalueren en zonodig bij te sturen. En ook hiervoor kan men de modelcyclus gebruiken. Het woord 'praktijk' (zie figuur 1) moet daartoe vervangen worden door het woord 'evaluatiesysteem'. Fase 1 wordt dan: 'het beschrijven van het beoogde evaluatiesysteem'. Fase 2: 'het formuleren van een aantal normen waaraan het beoogde evaluatiesysteem moet voldoen'. Enzovoort. We krijgen dan een cyclisch proces van exploratie, normering, operationalisatie, toetsing en bijsturing, waarbij het actuele evaluatiesysteem stap voor stap in over-