

Verslag van Bevindingen 5

29 april 2025

Inleiding

In dit vijfde verslag wordt de volgende fase van de evaluatie door Berenschot van een wetenschappelijke reflectie voorzien. Het betreft de conceptrapportage, zoals deze op 22 april 2025 door Berenschot is opgeleverd.

Dit vijfde verslag is op 29 april 2025 voorgelegd aan Berenschot.

De aanbevelingen tot dusver

In de eerste vier verslagen van bevindingen worden enkele aanbevelingen gedaan om het wetenschappelijke gehalte van de evaluatie te borgen. Deze aanbevelingen kunnen als volgt worden weergegeven.

Aanbevelingen bij het eerste verslag

Sommige concepten kunnen voorzien worden van een nadere operationalisatie. Het hanteren van te abstracte concepten levert immers het risico op dat hetzelfde fenomeen op verschillende manieren geclassificeerd en beoordeeld wordt, hetgeen de betrouwbaarheid in een latere onderzoeksfase kan reduceren.

Aanbevelingen bij het tweede verslag

In de concept definitieve beleidstheorie wordt niet adequaat verwezen, zodat niet altijd duidelijk wordt, waar de reconstructie van de beleidstheorie op is gebaseerd. Voor de betrouwbaarheid van de evaluatie is het echter van belang, dat bevindingen getraceerd en geverifieerd kunnen worden. Daarnaast is het voor de uiteindelijke analyse en conclusies van belang, dat bevindingen solide onderbouwd kunnen worden. De uitspraak van één respondent geeft bijvoorbeeld een minder solide onderbouwing dan formeel vastgesteld beleid, terwijl de indruk van de onderzoeker zelf in het geheel niet aangemerkt kan worden als bevinding.

Niet altijd wordt duidelijk in hoeverre de bevindingen betrekking hebben op formeel vastgesteld beleid of op de daadwerkelijke praktijk.

Een theorie wint aan wetenschappelijk gehalte als deze gefalsificeerd kan worden. Als de doelenboom wordt gezien, valt op, dat dit eigenlijk niet mogelijk is. Het gebruik van vrij abstracte werkwoorden maakt namelijk, dat men altijd kan betogen dat de activiteit is uitgevoerd, dat de resultaten zijn behaald en dat de doelstellingen zijn bereikt. Als niet kan worden vastgesteld in hoeverre activiteiten, resultaten of subdoelen niet zijn gerealiseerd, kan het onderlinge verband (of het ontbreken daarvan) tussen deze concepten niet worden aangetoond of gefalsificeerd. Het gevolg is dan, dat de validiteit van de gehanteerde beleidstheorie bijzonder lastig te evalueren is en hier zal zich ongetwijfeld een uitdaging voordoen tijdens de latere analysefase.

De interviewleidraad bevat een zeer duidelijke inleiding. Niet duidelijk wordt echter hoe de respons wordt geanalyseerd en wordt verwerkt. Zo weten respondenten niet in hoeverre zij later traceerbaar zijn. Het is van zeer groot belang, dat respondenten vooraf op de hoogte worden gesteld van de manier waarop de interviewrespons wordt verwerkt en in hoeverre deze respons geanonimiseerd wordt.

Daarnaast verdient het aanbeveling om in de inleiding aan te kondigen, dat van ieder interview een verslag wordt gemaakt, dat ter verificatie wordt aangeboden aan de respondent. De respondent heeft dan de mogelijkheid het verslag te beoordelen op onjuistheden en/of onvolledigheden. Dit werkt de betrouwbaarheid van de evaluatie in de hand.

Als de vragen inhoudelijk worden gezien, valt op dat niet altijd expliciet naar de relaties tussen de concepten wordt gevraagd. Aanbevolen wordt om systematisch te vragen naar de relatie tussen activiteiten en resultaten, tussen resultaten en subdoelen en tussen subdoelen en doelstelling.

Bij de uitvoering wordt terecht gevraagd naar neveneffecten, maar deze vraag zou specifiekere geformuleerd kunnen worden. Zo kan expliciet gevraagd worden naar ongewenste/voorziene en ongewenste/onvoorziene effecten. Daarnaast is het van belang dat ook gevraagd wordt naar gewenste/onvoorziene effecten (de zogeheten bijvangst).

Aanbevelingen bij het derde verslag

Aanbevolen wordt, om de interviewbevindingen zoals deze verwerkt worden in de hoofdstukken 3 van de deelevaluaties (de uitvoering), individueel voor te leggen aan de respondenten. Daarbij dient de mogelijkheid geboden te worden om hier individueel op te reageren. Vervolgens zal Berenschot de tijd moeten krijgen om deze reacties adequaat te verwerken. Deze aanbeveling betreft feitelijk dezelfde werkwijze die is toegepast bij de totstandkoming van de definitieve beleidstheorie.

Individuele reacties kunnen worden opgehaald voordat een terugkoppeldag plaatsvindt, maar ook daarna (zoals het voorbeeld bij Luchtkwaliteit laat zien). Een terugkoppeldag kan echter niet in de plaats treden van de mogelijkheid om individueel te reageren. Men kan er immers niet vanuit gaan, dat iedere respondent de vrijheid voelt tijdens een plenaire sessie te reflecteren.

Aanbevelingen bij het vierde verslag

Aanbevolen kan worden om bij de subconclusies die betrekking hebben op de doeltreffendheid en doelmatigheid deze twee concepten eerst van een eenduidige conceptualisatie en operationalisatie te voorzien. Uiteraard dienen de subconclusies vervolgens in lijn te worden gebracht met de geopteerde operationalisatie. Doordat dit nu nog niet is gedaan, kan een lezer die een andere conceptualisering en operationalisering hanteert, tot andere subconclusies komen.

Aanbevolen kan verder worden om los van het oordeel van de respondenten, een zelfstandige oordeel te vormen over de doeltreffendheid en doelmatigheid en dit duidelijk te maken in de tekst. Zo wordt duidelijk in hoeverre de subconclusies gebaseerd zijn op dit zelfstandige oordeel.

Aanbevelingen samengevat

Sommige aanbevelingen komen in verschillende verslagen terug. De aanbevelingen kunnen als volgt worden samengevat.

Aanbeveling	Toelichting
Eenduidige operationalisatie van begrippen.	Het hanteren van te abstracte concepten levert het risico op dat hetzelfde fenomeen op verschillende manieren geclassificeerd en beoordeeld wordt, hetgeen de betrouwbaarheid in een latere onderzoeksfase kan reduceren.
Verwijzingen.	Voor de betrouwbaarheid van de evaluatie is het van belang, dat bevindingen getraceerd en geverifieerd kunnen worden. Daarnaast is het voor de uiteindelijke analyse en conclusies van belang, dat bevindingen solide onderbouwd worden.
Onderscheid formeel vastgesteld beleid en daadwerkelijke praktijk.	Aangezien men er niet vanuit mag gaan, dat het formeel vastgesteld beleid conform wordt geïmplementeerd, dient duidelijk een onderscheid gemaakt te worden tussen beleid en praktijk.
Mogelijkheid tot falsificatie.	Als niet kan worden vastgesteld in hoeverre activiteiten, resultaten of subdoelen niet zijn gerealiseerd, kan het onderlinge verband (of het ontbreken daarvan) tussen deze concepten niet worden aangetoond of gefalsificeerd. Het gevolg is dan, dat de validiteit van de gehanteerde beleidstheorie bijzonder lastig te evalueren is.
Analyse en verwerking interviewrespons.	Het is van zeer groot belang, dat respondenten vooraf op de hoogte worden gesteld van de manier waarop de interviewrespons wordt verwerkt en in hoeverre deze respons geanonimiseerd wordt.
Vastlegging interviews.	Het verdient aanbeveling om in de inleiding aan te kondigen, dat van ieder interview een verslag wordt gemaakt, dat ter verificatie wordt aangeboden aan de respondent. De respondent heeft dan de mogelijkheid het verslag te beoordelen op onjuistheden en/of onvolledigheden. Dit werkt de betrouwbaarheid van de evaluatie in de hand.
Verifiëren causale relaties.	Aanbevolen wordt om systematisch te vragen naar de relatie tussen activiteiten en resultaten, tussen resultaten en subdoelen en tussen subdoelen en doelstelling.
Oog voor ongewenste en onvoorziene effecten.	Tijdens de interviews kan expliciet gevraagd worden naar ongewenste/voorziene en ongewenste/onvoorziene effecten. Daarnaast is het van belang dat ook gevraagd wordt naar gewenste/onvoorziene effecten (de zogeheten bijvangst).
Voorleggen bevindingen aan respondenten.	Aanbevolen wordt, om de interviewbevindingen zoals deze verwerkt worden in de hoofdstukken 3 van de deelevaluaties (de uitvoering), individueel voor te leggen aan de respondenten. Daarbij dient de mogelijkheid geboden te worden om hier individueel op te reageren.
Conclusies doelmatigheid en doeltreffendheid.	Aanbevolen kan worden om bij de subconclusies die betrekking hebben op de doeltreffendheid en doelmatigheid deze twee concepten eerst van een eenduidige conceptualisatie en operationalisatie te voorzien. Uiteraard dienen de subconclusies vervolgens in lijn te worden gebracht met de geopteerde operationalisatie.
	Aanbevolen kan verder worden om los van het oordeel van de respondenten, een zelfstandige oordeel te vormen over de doeltreffendheid en doelmatigheid en dit duidelijk te maken in de tekst. Zo wordt duidelijk in hoeverre de subconclusies gebaseerd zijn op dit zelfstandige oordeel.

Reflectie op de methoden en technieken van onderzoek

De gehanteerde methoden van dataverzameling en -analyse, i.e. een documentenanalyse, aangevuld met interviews en groepsgesprekken, zijn toereikend om een beleidsevaluatie adequaat vorm te geven. Het aantal geanalyseerde documenten en het aantal respondenten dat deel heeft genomen aan interviews en groepsgesprekken geven een goed beeld van zowel het gevoerde beleid (inclusief de gehanteerde beleidstheorie) als de daadwerkelijke implementatie van dat beleid.

In paragraaf 2.2 worden de definities van doeltreffendheid en doelmatigheid uitgewerkt. Daarbij valt op, dat doeltreffendheid wordt gedefinieerd als de mate waarin de beleidsdoelstelling dankzij de inzet van de onderzochte beleidsinstrumenten wordt gerealiseerd. Doorgaans verwijst doeltreffendheid naar de realisatie van gewenste en voorziene effecten als gevolg van het gevoerde beleid en het is opmerkelijk dat in de gehanteerde definitie niet naar deze effecten wordt gerefereerd. Beleidsdoelstellingen kunnen immers ook betrekking hebben op de prestaties die geleverd moeten worden, het proces dat gevoerd moet worden of zelfs op de input die beschikbaar is (bij een ombuigingsoperatie bijvoorbeeld). Doelmatigheid wordt gedefinieerd als de relatie tussen de effecten van het beleid en de kosten van het beleid. Hoewel deze definitie ook elders binnen het openbaar bestuur wordt gehanteerd, wordt door sommige auteurs aangegeven, dat deze definitie te breed is en te weinig basis biedt voor een adequate analyse. Feitelijk beslaat deze definitie namelijk het gehele causale verband tussen input, proces, output en effect. Bij een waargenomen wijziging van doelmatigheid aldus gedefinieerd, kan dan niet snel worden geconstateerd waar in de keten de oorzaak voor deze wijziging ligt. Bovendien kunnen deelverhoudingen contrair aan elkaar wijzigen met als gevolg dat een constante mate van doelmatigheid niet direct hoeft te betekenen, dat de onderliggende deelverhoudingen eveneens constant zijn. Daarnaast gaat de gehanteerde definitie van doelmatigheid voorbij aan het gegeven dat effecten ook zonder overheidsinterventie gerealiseerd kunnen worden (doelbereiking).

Het grootste probleem van de gehanteerde definities is echter dat de realisatie van beleidsdoelstellingen en effecten niet worden afgezet tegen een norm. Het realiseren van beleidsdoelstellingen leidt volgens de gehanteerde definitie tot doeltreffendheid, maar het is zeer wel mogelijk dat die beleidsdoelstellingen minder ambitieus zijn geformuleerd. Doeltreffendheid zou dan eerder een indicator zijn voor minder ambitieus geformuleerde doelstellingen. Evenzo zegt de relatie tussen kosten en effecten nog niet zo veel over de daadwerkelijke doelmatigheid. Als de effecten gerealiseerd worden binnen de begrote kosten, kan het immers zeer wel zo zijn, dat de beoogde effecten minder ambitieus zijn geformuleerd, of dat de budgetten vooraf zijn gemajoreerd.

Bovenstaande bemerkingen leiden dan doorgaans tot een andere definitie van de begrippen doelmatigheid en doeltreffendheid. Doelmatigheid wordt dan aangemerkt als de verhouding tussen gerealiseerde productiviteit en beoogde productiviteit, waarbij productiviteit wordt opgevat als output-input verhouding. Doeltreffendheid wordt aangemerkt als de verhouding tussen gerealiseerde effect-productiviteit en beoogde effect-productiviteit, waarbij effect-productiviteit wordt opgevat als effect-output verhouding. De beoogde productiviteit en effect-productiviteit bevatten een normstelling die idealiter zo objectief mogelijk vorm is gegeven (Klaassen e.a., 2010).

Bij de onderzoeksmethoden die voor de evaluatie zijn ingezet, wordt aangegeven dat gesprekken zijn gevoerd. Positief daarbij is, dat de lijst met gesprekspartners geanonimiseerd is. Aan de hand van bijlage 3 zijn de respondenten niet te traceren (met uitzondering wellicht van de ambtelijk opdrachtgever), hetgeen correspondeert met de communicatie richting de respondenten ten tijde van de uitnodiging om deel te nemen aan de interviews. Respondenten kregen vooraf te horen dat hun bijdrage niet getraceerd kon worden en dat namen geanonimiseerd zouden worden. Hoewel deze

mededeling mondeling is gedaan en niet is opgenomen in de schriftelijke uitnodiging, is steekproefsgewijs geverifieerd dat Berenschot deze mededeling (waarbij anonimiteit verzekerd werd) daadwerkelijk voorafgaand aan de gesprekken heeft gedaan. De deelname aan interviews en groeps gesprekken kan deels gebaseerd zijn op deze mededeling en Berenschot heeft correct en in lijn met deze mededeling deze evaluatie vormgegeven.

De interviewleidraad is traceerbaar vooraf gecommuniceerd richting de respondenten en dit geldt voor alle interviews in de verschillende fasen van de evaluatie. Zo wisten respondenten vooraf wat de doelstelling van de evaluatie was en welke vragen gesteld konden worden.

Bij de onderzoeksmethoden wordt aangegeven dat interne verslagen van de interviews zijn gemaakt. Deze verslagen zijn niet gedeeld en zijn dan ook niet te verifiëren. Ook zijn deze verslagen niet ter reflectie voorgelegd aan de respondenten zelf. Desgevraagd is aangegeven, dat dit een bewuste keuze betreft die in overleg met de provincie en Berenschot is gemaakt.

De concept deelproducties zijn echter wel voorgelegd aan de respondenten. Daarbij werd alle respondenten de mogelijkheid geboden om te reflecteren op deze deelproducties. Hierdoor ontstaat ook de mogelijkheid voor respondenten om na te gaan in hoeverre hun input tijdens de interviews correct is vertaald naar deze deelproducties. Van belang is daarbij dat respondenten dit schriftelijk en individueel kunnen doen, zodat het gevaar van 'group thinking' wordt weggenomen. Berenschot heeft expliciet respondenten de mogelijkheid geboden om schriftelijk en individueel te reflecteren op de deelproducties, waarmee de betrouwbaarheid en validiteit aanmerkelijk in de hand wordt gewerkt.

De respondenten hebben tijdens het verloop traceerbaar van deze mogelijkheid gebruik gemaakt. Vervolgens is nagegaan in hoeverre de schriftelijke input van de respondenten daadwerkelijk heeft geleid tot aanpassing van de concepten. Geconstateerd kan worden dat Berenschot de input juist en volledig heeft verwerkt. Tevens kan geconstateerd worden dat Berenschot daarbij een goed onderscheid heeft gemaakt tussen persoonlijke opvattingen van de respondenten en objectieve feitelijke input van respondenten, hetgeen overigens niet altijd eenvoudig is.

De terugkoppeldagen hebben traceerbaar plaatsgevonden en dit geldt evenzeer voor de afsluitende bijeenkomsten met het kernteam. Ook tijdens deze terugkoppeldagen en bijeenkomsten werd respondenten de mogelijkheid geboden om te reflecteren op conceptbevindingen. Ad random zijn enkele van deze dagen en bijeenkomsten bijgewoond en geconstateerd kan worden dat de sfeer tijdens deze sessies open, inhoudelijk en constructief was.

In de conclusies en aanbevelingen wordt niet verwezen naar de onderliggende documenten of naar de respons vanuit de interviews. Zo wordt gesteld dat bij het tegengaan van geluidhinder de provincie voornamelijk inzet op het opstellen en uitvoeren van het Actieplan Geluid Provinciale Wegen. Hier zou een verwijzing op zijn plaats zijn. Het maakt immers voor de robuustheid van conclusies uit, of deze uitspraak gebaseerd is op een beleidsdocument of op de uitspraak van een respondent. Deze bemerking geldt tevens voor de deelevaluaties die als onderlegger dienen voor de overkoepelende evaluatie.

In de toelichting op de conclusies wordt een duidelijk onderscheid gemaakt tussen beleid enerzijds en praktijk anderzijds. Duidelijk is voor de lezer in hoeverre een passage betrekking heeft op het geformuleerde beleid of op de daadwerkelijke uitvoering van dat beleid.

Kans is gezien om de gehanteerde beleidstheorie te traceren. Hiertoe zijn veel gesprekken gevoerd en documenten geanalyseerd. Tijdens de interviews is expliciet gevraagd naar de relatie tussen activiteiten, resultaten, subdoelen en doelstellingen. De theorie die op basis hiervan kon worden

gereconstrueerd, is voorgelegd aan de respondenten ter verificatie en de input van deze respondenten is traceerbaar verwerkt in een uiteindelijke beleidstheorie. Vervolgens wordt in de evaluatie correct en logisch gereflecteerd op deze beleidstheorie. De reconstructie van de beleidstheorie is aantoonbaar geverifieerd en geconstateerd kan worden dat dit onderdeel van de evaluatie adequaat is uitgevoerd.

Reflectie op de conclusies met betrekking tot doeltreffendheid

Bij de beoordeling van doeltreffendheid wordt correct aangegeven, dat niet altijd is vast te stellen in hoeverre het behalen van de doelstellingen daadwerkelijk te relateren is aan het handelen van de provincie. Het onderscheid tussen doeltreffendheid en doelbereik wordt hier adequaat gemaakt (p. 10). Waar de evaluatie over beleidsdoelstelling spreekt, worden overigens vermoedelijk gewenste en voorziene effecten bedoeld. In navolging op de eerdere opmerkingen met betrekking tot doeltreffendheid, kunnen beleidsdoelstellingen immers ook betrekking hebben op de output, het gevoerde proces en zelfs op de input. De conclusies met betrekking tot de doeltreffendheid verwijzen echter naar gewenste en voorziene maatschappelijke effecten en daarmee wordt toch aansluiting gevonden bij de gangbare definitie van doeltreffendheid: de verhouding tussen overheidsinterventie en voorziene en gewenste maatschappelijke effecten. In de toelichting worden enkele voorbeelden gegeven waarbij de doeltreffendheid minder goed is vast te stellen en voorbeelden waarbij dat juist wel kan. Op basis van deze voorbeelden kan niet zonder meer de conclusie worden getrokken dat over het geheel genomen de provincie doeltreffend heeft geopereerd (Conclusie 1, p. 11). Dit zou immers een analyse vereisen van alle gerealiseerde activiteiten en de realisatie van alle gestelde doelstellingen (voorzien en gewenst). Als deze analyse is uitgevoerd, zou hier zeker naar verwezen moeten worden in de toelichting op deze conclusie. Daarnaast wrekt zich hier het feit, dat de realisatie niet wordt afgezet tegen een objectieve externe norm, zodat een objectief oordeel over doeltreffendheid niet goed gegeven kan worden. De toevoeging op de conclusie dat de provincie de dingen goed doet, wekt verwarring, aangezien dit eerder een indicator zou zijn voor doelmatigheid, ook volgens de definitie die voor deze evaluatie is gehanteerd.

Na conclusie 1 wordt aangegeven dat uiteraard verbeteringen en optimalisaties mogelijk zijn. Deze worden uitgewerkt in de volgende conclusies. Zo wordt in conclusie 2 gesteld dat op het niveau van het beleidsdoel als geheel de verbinding aan kracht kan winnen door een overkoepelende visie op te stellen en beleidskeuzes en de uitvoering daarvan helder met elkaar te verbinden. Niet duidelijk is, in welke mate deze verbinding op dit moment ontoereikend is. Gesteld wordt immers dat de teams vaak de juiste afdelingen binnen de organisatie weten te vinden en vice versa.

Deze conclusie is voorts nog redelijk abstract geformuleerd en denkbaar is, dat deze conclusie ruimte biedt voor uiteenlopende interpretaties. Bovendien is niet geheel duidelijk hoe hiermee het geconstateerde knelpunt uit de toelichting (het borgen van samenwerking tussen de teams en afdelingen) verholpen kan worden. Dit geldt evenzeer voor de twee aanbevelingen die gekoppeld worden aan deze conclusie. De aanbeveling aan PS om te werken aan een overkoepelende milieuvisie en de aanbeveling aan GS en de ambtelijke organisatie om verbinding tussen de teams van de beleidskeuzes te organiseren, zijn vrij abstract geformuleerd. Daarnaast wordt niet geheel duidelijk hoe de teams en afdelingen nu verbonden zouden moeten worden.

Aan de andere kant zijn de derde conclusie en de aanbevelingen die hieraan gekoppeld worden, zeer adequaat toegelicht en concreet geformuleerd. De conclusie wordt gedragen door de toelichting en de aanbevelingen bieden voldoende richting voor een handelingsperspectief, zowel voor PS als voor GS en de ambtelijke organisatie.

Bij de opmaat naar de vierde en vijfde conclusie wordt terecht geconstateerd, dat niet altijd concrete resultaten zijn gekoppeld aan het omgevingsbeleid van de provincie. De vijfde conclusie sluit hier logisch op aan en dit geldt evenzeer voor de aanbeveling. De vraag rijst in dit verband echter waarom concrete resultaten niet altijd worden opgenomen. Het achterwege laten van output-gerelateerde informatie kan immers tal van oorzaken hebben en het zou de evaluatie ten goede komen als hierop in was gegaan.

Bij het onderdeel doeltreffendheid zou de evaluatie aangevuld kunnen worden met een beschouwing over onvoorziene, maar gewenste effecten, voorziene en ongewenste effecten en, tot slot, onvoorziene en ongewenste effecten (van Twist en Verheul, 2010). Deze hebben zich namelijk ook ongetwijfeld voorgedaan en kunnen de perceptie van doeltreffendheid in zeker mate beïnvloeden.

Reflectie op de conclusies met betrekking tot doelmatigheid

Terecht wordt geconstateerd dat de informatie op product- en deelproductniveau niet altijd een beeld geeft van de hoeveelheid geld, die binnen de beleidskeuze, naar welk doel is gegaan. Ook binnen een beleidskeuze is er geen goed beeld te schetsen van de verdeling van de kosten, zo wordt terecht opgemerkt. De evaluatie constateert voorts, dat intern binnen de afdeling dit inzicht al wel (gedeeltelijk) aanwezig is. De conclusie (6) is dan logischerwijs, dat sturing op doelmatigheid moeilijk is. Dit verhoudt zich echter minder goed tot de constatering dat de provincie de 'dingen goed doet'. Uit de evaluatie komt naar voren dat er geen concrete indicaties zijn die wijzen op een ondoelmatig verloop, maar hiermee zij niet gesteld dat daarmee doelmatig wordt gehandeld. De constatering dat de provincie de dingen goed doet, zou dan ook beter kunnen worden toegelicht, waarbij zorgvuldiger kan worden aangegeven hoe men precies tot deze constatering is gekomen, zeker gelet op het feit, dat binnen een beleidskeuze geen goed beeld is te schetsen van de verdeling van de kosten. Wellicht heeft Berenschot hier een extracomptabele analyse uitgevoerd, maar dit blijkt verder niet uit de toelichting.

Het gegeven dat de operationalisatie van doelmatigheid geen norm bevat (zie boven), maakt een eenduidig en objectief oordeel met betrekking tot doelmatigheid lastig. Als men al kans ziet om de realisatie van beleidsdoelstellingen te koppelen aan ingezette (financiële) middelen, dan nog rijst de vraag of daarmee doelmatig is gehandeld. Het enkel afzetten van de realisatie ten opzichte van de begroting volstaat doorgaans niet, aangezien altijd de kans bestaat de budgetten gemajoreerd worden en in die zin minder bruikbaar zijn als objectieve norm.

Samenvatting bevindingen

De bovenstaande reflectie kan als volgt worden afgezet tegenover de eerdere aanbevelingen.

Aanbeveling	Toelichting	Reflectie
Eenduidige operationalisatie van begrippen.	Het hanteren van te abstracte concepten levert het risico op dat hetzelfde fenomeen op verschillende manieren geclassificeerd en beoordeeld wordt, hetgeen de betrouwbaarheid in een latere onderzoeksfase kan reduceren.	De definities van doeltreffend en doelmatigheid bevatten geen normstelling. De constatering dat beleidsdoelstellingen worden gerealiseerd kan een indicator zijn voor doeltreffendheid, maar

		ook voor minder ambitieus geformuleerde doelstellingen. Dit geldt tevens voor het concept doelmatigheid.
Verwijzingen.	Voor de betrouwbaarheid van de evaluatie is het van belang, dat bevindingen getraceerd en geverifieerd kunnen worden. Daarnaast is het voor de uiteindelijke analyse en conclusies van belang, dat bevindingen solide onderbouwd worden.	Zowel in de toelichting op de conclusies als in de onderliggende deelevaluaties kan veelvuldiger verwezen worden. Verwijzingen kunnen daarbij zowel in de hoofdtekst als in voetnoten worden opgenomen.
Onderscheid formeel vastgesteld beleid en daadwerkelijke praktijk.	Aangezien men er niet vanuit mag gaan, dat het formeel vastgesteld beleid conform wordt geïmplementeerd, dient duidelijk een onderscheid gemaakt te worden tussen beleid en praktijk.	In de toelichting op de conclusies wordt duidelijk in hoeverre een passage betrekking heeft op beleid of op de uitvoering van dat beleid.
Mogelijkheid tot falsificatie.	Als niet kan worden vastgesteld in hoeverre activiteiten, resultaten of subdoelen niet zijn gerealiseerd, kan het onderlinge verband (of het ontbreken daarvan) tussen deze concepten niet worden aangetoond of gefalsificeerd. Het gevolg is dan, dat de validiteit van de gehanteerde beleidstheorie bijzonder lastig te evalueren is.	De reconstructie van de beleidstheorie is geverifieerd en geconstateerd kan worden dat dit onderdeel van de evaluatie adequaat is uitgevoerd.
Analyse en verwerking interviewrespons.	Het is van zeer groot belang, dat respondenten vooraf op de hoogte worden gesteld van de manier waarop de interviewrespons wordt verwerkt en in hoeverre deze respons geanonimiseerd wordt.	Vooraf is medegedeeld dat input van respondenten geanonimiseerd zou worden en deze evaluatie is correct en in lijn met deze mededeling vormgegeven.
Vastlegging interviews.	Het verdient aanbeveling om in de inleiding aan te kondigen, dat van ieder interview een verslag wordt gemaakt, dat ter verificatie wordt aangeboden aan de respondent. De respondent heeft dan de mogelijkheid het verslag te beoordelen op onjuistheden en/of onvolledigheden. Dit werkt de betrouwbaarheid van de evaluatie in de hand.	De interne verslagen zijn niet voorgelegd aan de respondenten. Respondenten hebben echter wel de mogelijkheid gekregen om schriftelijk en individueel te reageren op deelproducties. Van deze mogelijkheid is inderdaad gebruik gemaakt en geconstateerd kan worden dat de input juist en volledig is verwerkt. Tevens kan geconstateerd worden dat daarbij een goed onderscheid is gemaakt tussen persoonlijke opvattingen van de

		respondenten en objectieve feitelijke input van respondenten. Daarnaast is tijdens de terugkoppeldagen de mogelijkheid geboden om te reflecteren op de conceptbevindingen en ook hier is gebruik van gemaakt.
Verifiëren causale relaties.	Aanbevolen wordt om systematisch te vragen naar de relatie tussen activiteiten en resultaten, tussen resultaten en subdoelen en tussen subdoelen en doelstelling.	Tijdens de interviews is expliciet gevraagd naar de relatie tussen activiteiten, resultaten, subdoelen en doelstellingen.
Oog voor ongewenste en onvoorziene effecten.	Tijdens de interviews kan expliciet gevraagd worden naar ongewenste/voorziene en ongewenste/onvoorziene effecten. Daarnaast is het van belang dat ook gevraagd wordt naar gewenste/onvoorziene effecten (de zogeheten bijvangst).	De evaluatie kan aangevuld worden met een beeld van ongewenste/voorziene en ongewenste/onvoorziene effecten. Ook gewenste, maar onvoorziene effecten komen niet expliciet aan bod in de evaluatie.
Voorleggen bevindingen aan respondenten.	Aanbevolen wordt, om de interviewbevindingen zoals deze verwerkt worden in de hoofdstukken 3 van de deelevaluaties (de uitvoering), individueel voor te leggen aan de respondenten. Daarbij dient de mogelijkheid geboden te worden om hier individueel op te reageren.	Deze mogelijkheid is traceerbaar geboden en hier is door respondenten traceerbaar gebruik van gemaakt.
Conclusies doelmatigheid en doeltreffendheid.	Aanbevolen kan worden om bij de subconclusies die betrekking hebben op de doeltreffendheid en doelmatigheid deze twee concepten eerst van een eenduidige conceptualisatie en operationalisatie te voorzien. Uiteraard dienen de subconclusies vervolgens in lijn te worden gebracht met de geopteerde operationalisatie.	De conclusies met betrekking tot doeltreffendheid en doelmatigheid worden vrij snel en stellig betrokken. De toelichting die op deze conclusies gegeven wordt, draagt de conclusies in mindere mate.
	Aanbevolen kan verder worden om los van het oordeel van de respondenten, een zelfstandige oordeel te vormen over de doeltreffendheid en doelmatigheid en dit duidelijk te maken in de tekst. Zo wordt duidelijk in hoeverre de subconclusies gebaseerd zijn op dit zelfstandige oordeel.	Omdat in de operationalisatie van de concepten doeltreffendheid en doelmatigheid geen expliciete en objectieve norm wordt gehanteerd, moet volstaan worden met een vergelijking tussen wat de provincie zelf aan middelen heeft begroot en als

		doelstelling voor ogen had enerzijds en de realisatie van kosten en doelstellingen anderzijds. Dit verkleint evenwel de mogelijkheid tot een objectieve beoordeling van de mate van doeltreffendheid en doelmatigheid.

Eindoordeel

De evaluatie is wetenschappelijk adequaat uitgevoerd. Daarbij zij opgemerkt, dat Berenschot zich zeer behulpzaam en transparant opstelde tijdens de totstandkoming van deze reflectie.

De noties met betrekking tot doelmatigheid en doeltreffendheid zijn ingegeven door publicaties die in de afgelopen jaren hierover zijn verschenen, ook in internationaal verband. Daarbij zij gesteld, dat deze noties conceptueel weliswaar zuiver en logisch zijn, maar dat de vertaling van deze noties naar de praktijk van het openbaar bestuur zeer gecompliceerd is. Zo is het construeren van een externe norm voor een output-effect verhouding of voor een input-effect verhouding in de praktijk welhaast onmogelijk, hoe theoretisch wenselijk een dergelijke norm ook is.

Aangehaalde publicaties

- Klaassen, H.L., Anderson, R., Maks, H., Doelmatigheid, Puzzels voor het openbaar bestuur, in *Efficiency en effectiviteit van de publieke sector in de weegschaal*, red. D. Verlet en C. Devos, Studiedienst van de Vlaamse Regering, Room, Sint-Niklaas, april 2010.
- Twist, M. van, Verheul, W., Onvoorziene opbrengsten, Bestuurswetenschappen, no. 37, 4, 2010.