

# Evaluatiemethodiek benutten

Deelrapport: Richtlijnen ex-post beleidsevaluaties

Datum 22 december 2010  
Kenmerk VW82

*MuConsult B.V.*  
Postbus 2054  
3800 CB Amersfoort  
Telefoon 033 – 465 50 54  
Fax 033 – 461 40 21  
E-mail [INFO@MUCONSULT.NL](mailto:INFO@MUCONSULT.NL)  
Internet [WWW.MUCONSULT.NL](http://WWW.MUCONSULT.NL)

# Inhoudsopgave

<b>Samenvatting</b>	<b>1</b>
<b>1. Inleiding</b>	<b>2</b>
1.1 Algemeen	2
1.2 Ex-post evaluatie	2
1.3 Eisen aan evaluatieonderzoek	4
1.4 Deze rapportage	5
<b>2. Kwaliteitseisen en -criteria ex-post onderzoek</b>	<b>6</b>
2.1 Inleiding	6
2.2 Probleemstelling van het onderzoek	6
2.3 Beleidstheorie	9
2.4 Keuze van de onderzoeksopzet (hoofdlijnen)	10
2.5 Dataverzameling	12
2.6 Analyse van de data	14
2.7 Rapportage	17
<b>3. Organisatie en uitvoering van evaluatie</b>	<b>19</b>

# Samenvatting

Het directoraat-generaal Mobiliteit (DGMO) van het Ministerie van Verkeer en Waterstaat heeft aan MuConsult en 4Cast opdracht verleend om te komen tot een goede methode van evalueren van benutting en benuttingsmaatregelen (ex-post en ex-ante). Daartoe wordt een overzicht gegeven van bestaande kennis van benuttingsmaatregelen en wordt een methodiek uitgewerkt waarmee ex-ante evaluaties kunnen worden uitgevoerd om een kosten-batenanalyse te kunnen maken over bredere uitrol, zowel binnen benutting als tussen benutting en andere pijlers als bouwen. Op basis daarvan worden aanbevelingen afgeleid voor verbeteringen ex-ante en ex-post evaluaties.

Het doel van deze deelrapportage is het opstellen van **een set van richtlijnen waaraan toekomstige ex-post evaluaties van benuttingsmaatregelen/-beleid moeten voldoen** om te komen tot een systematische programmering van deze ex-post evaluaties en vergelijkbaarheid in de uitkomsten van verschillende evaluatieonderzoeken en beantwoording van vragen over de beleidseffecten.

Daarbij gaat het om eisen ten aanzien van:

- De uitvoering van het onderzoek. Er worden stadia in het onderzoeksproces onderscheiden met specifieke kwaliteitsnormen;
- De organisatie van het onderzoek. Hierbij gaat het om de wijze waarop het onderzoek wordt uitgevoerd en ingebed in de organisatie.

In het rapport zijn checklists opgenomen met vragen die bij kwaliteitsbeoordeling worden gebruikt.

# 1. Inleiding

## 1.1 Algemeen

Het beter benutten van de huidige infrastructuur is naast bouwen en beprijzen één van de pijlers om de doelen uit de Nota Mobiliteit te halen. En deze pijler is nog belangrijker geworden, aangezien 'beprijzen' momentaal geen politiek haalbare richting is. Met het Beleidskader Benutten heeft de minister de beleidslijn ten aanzien van deze pijler aan de Tweede Kamer gepresenteerd. Een aantal concrete acties uit dit beleidskader worden momenteel uitgevoerd, zoals de Mobiliteitsaanpak, Actieprogramma Wegen en proeven Dynamax, dan wel voorbereid. Eén van de acties uit het Beleidskader Benutten is het ontwikkelen van een landelijk evaluatie- en monitoringsprogramma. Achtergrond van deze actie is dat een eenduidig beeld van de kosteneffectiviteit van benuttingsmaatregelen ontbreekt. Tevens is het evalueren van benuttingsmaatregelen complex gebleken, omdat het ongelijksoortige maatregelen betreft, waarvan de effectiviteit sterk afhangt van de specifieke situatie. Uit evaluaties blijkt dat dezelfde maatregel op verschillende locaties en onder verschillende omstandigheden verschillende effecten laat zien. Verder vertonen de diverse afzonderlijke maatregelen vaak een complexe, veelal vooralsnog onvoldoende doorgronde interactie die de effectiviteit positief of negatief kan beïnvloeden. De behoefte aan evaluatiegegevens is groot teneinde een goede beleidsafweging te kunnen maken over de (bredere) inzet/uitrol van benuttingsmaatregelen zowel binnen het geheel van benutting, als tussen benutting en andere pijlers zoals bouwen en beprijzen.

Het directoraat-generaal Mobiliteit (DGMO) van het Ministerie van Verkeer en Waterstaat heeft aan MuConsult en 4Cast de opdracht verleend om te komen tot een goede methode van evalueren van benutting en benuttingsmaatregelen (ex-post en ex-ante), welke moet leiden tot inzicht in:

- a) de kosteneffectiviteit van benuttingsmaatregelen en de condities waaronder de benuttingsmaatregelen en de gecoördineerde inzet hiervan goed werken;
- b) de generieke kosteneffectiviteit van groepen van benuttingsinstrumenten. Deze kosteneffectiviteit moet bepaald worden op basis van bijvoorbeeld de investerings- en exploitatiekosten en de maatschappelijke baten in termen van de doelstellingen, thans uit de Nota Mobiliteit (NoMo); en
- c) de effecten van benuttingsmaatregelen op het bereiken van de NoMo-beleidsdoelstellingen, zoals doorstroming, milieu, verkeersveiligheid, verlies of winst aan mobiliteit.

## 1.2 Ex-post evaluatie

Onderdeel van de opdracht is de vraag aan welke eisen ex-post beleidsevaluaties van benuttingsmaatregelen moeten voldoen. Daarbij gaan we uit van de eisen voortvloeiend uit functies van ex-post evaluaties in het beleidsproces. Deze zijn beschreven in de eerder uitgebrachte notitie (MuConsult, september 2010, Kader: beleidsdoelen, maatregelen en indicatoren). In deze notitie is aangegeven dat beleidsevaluatie twee belangrijke functies heeft bij de beleidsvorming:

1. Het afleggen van verantwoording ten aanzien van te maken en gemaakte keuzen. Door informatie te verschaffen aan het parlement en de samenleving, kan men laten zien dat beleidsmaatregelen doeltreffend en doelmatig zijn;

2. Het leren, waarbij resultaten van de onderscheiden typen evaluatieonderzoeken worden gebruikt om te komen tot vergroting van de doelmatigheid en doeltreffendheid van de beleidsmaatregelen.

#### **Uitwerking doelen ex-post evaluaties**

Het afleggen van verantwoording wordt uitgewerkt in de ministeriële regeling "Regeling prestatiegegevens en evaluatieonderzoek Rijksoverheid". Deze regeling komt voort uit het streven van de Rijksoverheid om resultaatgericht te sturen, waaraan een belangrijke impuls is gegeven door het project "Van beleidsbegroting tot beleidsverantwoording" (VBTB). Centraal daarin staan de vragen: wat willen we bereiken, wat gaan we daarvoor doen en wat mag dat gaan kosten. Achteraf worden de uitgaven verantwoord door een koppeling te leggen tussen de doelen van beleid (te bereiken effecten), de daartoe te leveren prestaties en de daarvoor in te zetten middelen. In de genoemde regeling zijn de volgende bepalingen opgenomen:

1. alle doelstellingen van beleid worden in de regel ten minste eens in de vijf jaar (nader) geëvalueerd;
2. evaluatieonderzoek voldoet aan kwaliteitseisen op het vlak van validiteit, betrouwbaarheid en bruikbaarheid;
3. de kwaliteit, alsmede de onafhankelijkheid waarmee een onderzoek is opgezet en/of uitgevoerd, moet achteraf kunnen worden beoordeeld, en;
4. de ambtelijke en politieke top wordt volgens een standaardprocedure op de hoogte gebracht van de uitkomsten van evaluatieonderzoek.

Achterliggend doel van deze uitwerking is om de omvang en kwaliteit van de evaluatieonderzoeken, en daarmee het beleid, te verbeteren. Uit een aantal evaluatieonderzoeken was namelijk gebleken dat deze kwaliteit nogal te wensen overliet (Algemene Rekenkamer, 20xx). Verantwoordelijk voor deze regeling is FEZ binnen ieder ministerie.

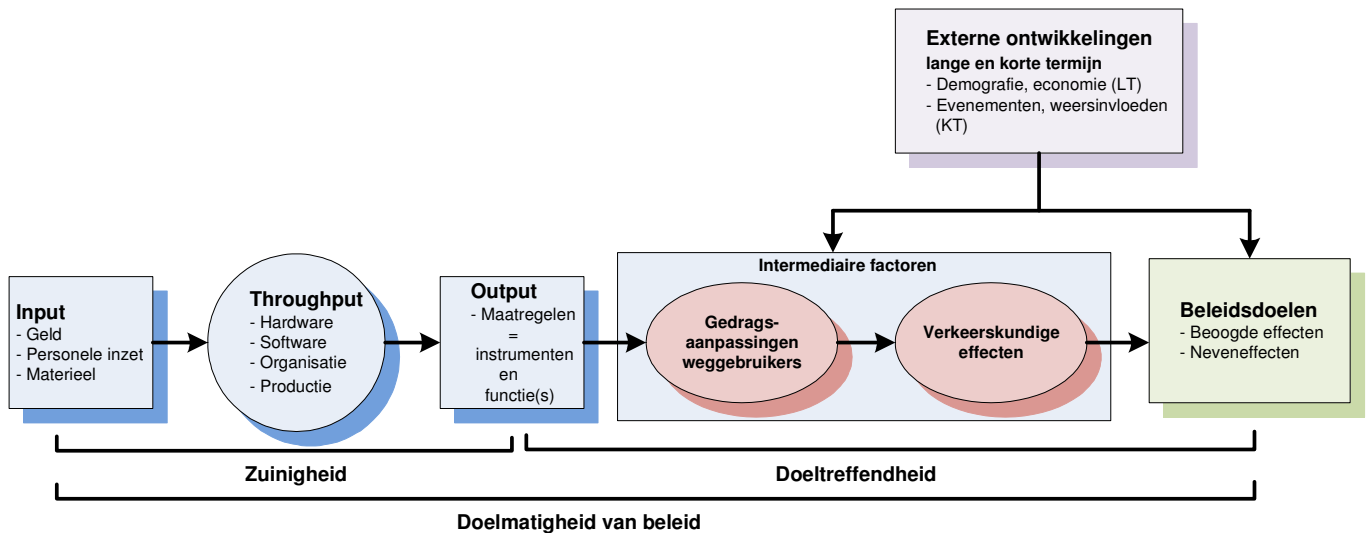
Het leerdoel is vooral gericht op het verbeteren van het beleid. Beleidsevaluaties beschrijven het probleem in alle relevante facetten. Op basis van resultaten kunnen maatregelen worden bijgesteld, kunnen doelstellingen worden bijgesteld en kunnen beslissingen worden genomen over eventuele verdere uitrol van de maatregelen. Mede daarom is niet alleen een terugkoppeling naar de beleidsdoelen van belang, maar dient ook informatie beschikbaar te komen bedoeld om besluitvorming over bijstelling van de maatregel en verdere uitrol mogelijk te maken. Zonder deze terugkoppeling is geen sprake van een *beleids*cyclus.

Vanuit deze algemene doelen richt ex-post evaluatieonderzoek zich op de doelbereiking, de doeltreffendheid en doelmatigheid van beleidsmaatregelen:

1. *Doelbereiking*: halen we de doelen?
2. *Doeltreffendheid*: in welke mate is het realiseren van de doelen toe te schrijven aan de beleidsmaatregelen? Dit is geen eenvoudige vraag, want daarbij dient te worden vastgesteld wat de bijdrage van de maatregel is geweest in de waargenomen ontwikkelingen;
3. *Doelmatigheid (kosteneffectiviteit)*. Hoe is de verhouding tussen de gerealiseerde effecten en de gemaakte kosten/inzet van menskracht. Merk op dat daarbij onderscheid gemaakt wordt tussen de doelmatigheid van beleid en de doelmatigheid van de bedrijfsvoering. Bij het laatste gaat het om de verhouding tussen de output (prestaties ofwel geleverde producten en diensten) en de middelen, bij doelmatigheid van beleid gaat het om de verhouding tussen effecten (in termen van beleidsdoelen en eventuele neveneffecten) en ingezette (overheids)middelen.

De relatie tussen de verschillende onderdelen wordt middels het volgende schema in kaart gebracht.

*Figuur 1.1: Onderdelen van de ex-post beleidsevaluatie*



Vraagstukken betreffende doelbereiking, doeltreffendheid en doelmatigheid kunnen op elk van de onderscheiden niveaus worden beantwoord, dus zowel op het niveau van beleidssporen, beleidsprogramma's alsmede afzonderlijke maatregelen/projecten. Bij elke evaluatie zullen steeds beleidsdoelen moeten zijn geformuleerd waartegen de maatregelen worden afgezet.

### 1.3 Eisen aan evaluatieonderzoek

De gedachte is dat de uitkomsten van evaluatieonderzoek ex-post bijdragen aan de kwaliteit van de besluitvorming over beleid en aan de bedrijfsvoering. In de bovengenoemde regeling wordt gesteld dat het evaluatieonderzoek moet voldoen aan de kwaliteitseisen die gelden voor wetenschappelijk onderzoek:

1. Validiteit, de algemene geldigheid van de onderzoeksopzet en de wijze waarop conclusies worden getrokken;
2. De betrouwbaarheid en nauwkeurigheid (van de gebruikte meetmethoden);
3. De bruikbaarheid van de onderzoeksresultaten in het beleidsproces;
4. Onafhankelijkheid van het onderzoek, waarbij de opstellers van het evaluatieonderzoek de onderzoeksvragen, resultaten en bevindingen kunnen formuleren zonder inmenging van ontwikkelaars en uitvoerders van de maatregel.

Het gaat hier – uiteraard – om zeer algemene eisen die in het navolgende verder worden ingevuld.

Op basis van de meta-evaluatie van bestaande ex-post evaluaties kunnen de volgende eisen worden geformuleerd ten aanzien van de ex-post evaluatie van benuttingsbeleid:

- **Doelen en gebruik van evaluatieresultaten helder benoemen**  
Naarmate de doelen van de evaluatie helderder worden geformuleerd zal de kwaliteit van het evaluatieonderzoek toenemen. Daarbij dient onder meer worden gedacht aan

de uitwerking van de wijze waarop evaluatieresultaten bij beleid en uitvoering worden gebruikt.

- **Versterken gerichtheid op beleidsdoelen**

Beleidsevaluaties vinden altijd plaats op basis van beleidsdoelen. Evaluatie van de werking van de maatregelen is daarbij een onderdeel. Dit betekent dat ook de inzet van de maatregel gericht zal zijn op realisatie van de beleidsdoelen. Dit betekent ook dat de beleidsdoelen zelf expliciet en helder geformuleerd moeten worden.

- **Duidelijke kwaliteitseisen vaststellen.**

Daarbij is van belang om eisen te stellen aan de kwaliteit van de proefopzet, de dataverzameling en aan de analysemethoden in het licht van het te bestuderen fenomeen en de heersende evaluatiedoelen. Dat daarbij geëigende competenties (kennis, vaardigheden en houding) noodzakelijk zijn spreekt voor zich.

Deze eisen zullen in het navolgende worden uitgewerkt.

## 1.4 Deze rapportage

Het doel van deze deelrapportage is het opstellen van **een set van richtlijnen waaraan toekomstige ex-post evaluaties van benuttingsmaatregelen/-beleid moeten voldoen** om te komen tot een systematische programmering van deze ex-post evaluaties en vergelijkbaarheid in de uitkomsten van verschillende evaluatieonderzoeken en beantwoording van vragen over de beleidseffecten.

De rapportage is geen kookboek van onderzoeksmethoden; daarvoor bestaan andere rapporten. Door DVS en TNO is recentelijk de versie 2010 van de "Leidraad evaluaties benutting" tot stand gekomen. Dit rapport omvat een uitgebreidere beschrijving van methoden en technieken inzake de evaluatie van projecten. Deze rapportage is gebaseerd op Plan-Do-Check-Act cyclus en het V-model dat is uitgewerkt door het FESTA-consortium. Op onderdelen is sprake van een belangrijke overlap met deze rapportage. Belangrijk verschil is dat deze rapportage zich richt het evalueren van maatregelen die zijn genomen om beleidsdoelen te realiseren. De DVS-TNO-methodiek is meer gericht op de evaluatie van projecten en instrumenten.

Ook zal verdere vernieuwing van onderzoeksmethoden moeten plaatsvinden Voorgesteld wordt om ook niet uitputtend te zijn ten aanzien van de eisen. We stellen voor om te focussen op een aantal verbeterpunten ten opzichte van ex-post evaluatierapportages zoals tot op heden veelal worden uitgebracht.

## **2. Kwaliteitseisen en -criteria ex-post onderzoek**

### **2.1 Inleiding**

Dit hoofdstuk beschrijft inhoudelijke eisen die aan de ex-post beleidsevaluatie van benuttingsmaatregelen worden gesteld. Onderscheiden worden verschillende stadia in het evaluatieproces, te weten de probleemstelling, de uitwerking van de gehanteerde beleidstheorie, de hoofdlijnen van de onderzoeksopzet, de dataverzameling, de analysemethoden en de rapportage.

Ten aanzien van elk van deze fasen kunnen aandachtspunten worden geformuleerd ten aanzien van de kwaliteit van de uitvoering. Deze worden in de vorm van checklists geformuleerd.

Omdat deze notitie zich richt op ex-post evaluatie van benuttingsbeleid in brede zin, zijn de checklists algemeen geformuleerd, zodat ze voor een brede range van vraagstukken kunnen worden gebruikt. Daarnaast moet worden opgemerkt dat steeds de insteek gekozen wordt van een beleidsevaluatie van projecten en programma's. De checklist is dan ook uitgewerkt vanuit de gedachte dat beleidsdoelen centraal staan bij de evaluaties. Uit de meta-evaluatie is gebleken dat de meeste evaluaties uitgevoerd worden vanuit specifieke projectmatige overwegingen en niet gekoppeld worden aan beleidsdoelen. Met deze richtlijnen wordt beoogd aan uitvoerders van de evaluaties duidelijkheid te verschaffen over de eisen die vanuit beleid aan evaluaties worden gesteld.

### **2.2 Probleemstelling van het onderzoek**

De probleemstelling van onderzoek bestaat uit vier onderdelen: aanleiding, doelstelling, context en vraagstelling. Deze onderdelen worden meestal in onderlinge wisselwerking uitgewerkt.



**Aanleiding**

De aanleiding is meestal een probleem, waarbij sprake is van een verschil tussen de door beleid beoogde situatie en de feitelijke of verwachte situatie. Bij ex-post evaluatie van een maatregel of een programma wordt aangegeven om welke maatregel het gaat en wat de te verwachten bijdrage is aan de oplossing van het probleem.

**Doelstelling van de evaluatie**

Bij ex-post evaluatie onderzoek vervult de doelstelling een belangrijke functie, namelijk van het expliciet maken van wat wel en niet van de onderzoeker wordt verwacht, bijvoorbeeld: alleen informatie, of ook aanbevelingen, of zelfs ook hulp bij de implementatie daarvan?

**Context**

Wie een goede vraagstelling wil ontwikkelen, moet zich in de context van het probleem verdiepen, waarbij aandacht wordt besteed aan kenmerken van de aan de orde zijnde problemen, de genomen maatregelen en andere relevante ontwikkelingen.

**Vraagstelling<sup>1</sup>.**

Dit betreft de vragen waarop het onderzoek antwoorden moet geven. Vaak is sprake van één of enkele hoofdvragen, al dan niet onderverdeeld in enkele deelvragen.

NB: Wie een hele reeks specifieke vragen opstelt, is waarschijnlijk al bezig met het maken van een onderzoeksinstrument, en niet met een onderzoeksvraagstelling. Tot de probleemstelling behoren de hoofdvragen die voortvloeien uit de aanleiding en doel van het onderzoek.

Ten aanzien van deze fase kunnen een aantal kwaliteitscriteria worden benoemd waaraan goed evaluatieonderzoek moet voldoen.

---

<sup>1</sup> Soms worden ook hypothesen dwingend voorgeschreven. Naar onze mening is dit alleen het geval als expliciet bepaalde veronderstellingen moeten worden getoetst. Het is vaak beter om deze in de vorm van onderzoeksvragen te formuleren; dat laat meer ruimte aan onderzoekers om dieper op de materie in te gaan

#### **Checklist kwaliteitseisen probleemstelling**

- Is de aanleiding tot het evaluatieonderzoek helder omschreven? Is het onderzoek gericht op vaststellen van het doelbereik, de doeltreffendheid van de maatregel en/of de doelmatigheid van de maatregel vanuit beleidsdoelen bezien?
- Zijn de beleidsdoelstellingen die men met de te bestuderen maatregelen wil bereiken voldoende helder en operationeel beschreven? (doelvariabelen, normen, onderscheid subgroepen van betroffenen(equity), planperiode etc.)
- Is het doel van het evaluatieonderzoek helder omschreven in beleidstermen? Gaat het om verwerven van inzichten? Wat gebeurt er met de verworven inzichten? Worden aanbevelingen verwacht ten aanzien van bijstellingen van de maatregel, verdere uitrol en dergelijke?
- Zijn de relevante aspecten uit de context van de evaluatie goed omschreven? Wordt een beleidsmatige insteek gekozen bij beschrijving van de context?
- Zijn de hoofdvraagstellingen zodanig geformuleerd dat het doel van onderzoek wordt gerealiseerd?
- Is het probleem voldoende ingeperkt zodat het hanteerbaar is voor de onderzoeker?

De implicatie van deze kwaliteitseisen is dat benuttingsprojecten beschreven worden in termen van de specifieke beleidsdoelen (bereikbaarheid, veiligheid en milieu) waaraan het project een bijdrage zal leveren. Deze beleidsdoelen en de verwachte bijdrage van het te evalueren project daaraan staan centraal bij beleidsevaluaties.

Daarnaast wordt reeds bij de probleemstelling helder aangegeven wat het beoogde gebruik van de onderzoeksresultaten is. Deze probleem- en doelstelling verschilt ten opzichte van de gangbare praktijk waarin verkeerskundige problemen vaak de probleemstelling omvatten en verkeerskundige (tussen)doelen vaak als hoofddoelen worden geformuleerd van evaluaties. Dergelijke evaluaties richten zich veelal op het vaststellen van effecten op de intermediaire variabelen (capaciteit, prestaties, etc) in plaats van op de beleidsdoelen.

Er zijn ook andere typen evaluaties, zoals de evaluatie van de werking van de maatregel. Dergelijke evaluaties kunnen gericht zijn op verdere optimalisatie van maatregelen. Dergelijke evaluaties zijn vooral bedoeld om wegbeheerders handvatten te bieden om de inzet van maatregelen te verbeteren.

Een heldere uitwerking van aanleiding, doel- en vraagstelling is behulpzaam om de onderzoekers te richten op het te adresseren probleem en te komen met bruikbare conclusies en aanbevelingen. Het belang van het zorgvuldig uitwerken van de probleemstelling in de beginfase van het onderzoek kan niet genoeg worden benadrukt, omdat het van zeer groot belang is voor de uiteindelijke kwaliteit van het onderzoek. Zoals eerder gezegd betekent dit ook dat de beleidsmakers hun beleidsdoelen zo specifiek mogelijk moeten aangeven. Het onderzoeksdoel moet daarom zo nauwkeurig en zo concreet mogelijk gedefinieerd zijn. Het type probleemstelling geeft implicaties voor de uitwerking van het onderzoeksontwerp en de te volgen onderzoeksstrategie.

## 2.3 Beleidstheorie

De centrale vraagstelling van beleidsevaluatieonderzoek is het vaststellen van de effectiviteit en/of de efficiency van de (combinatie van) maatregelen. Daartoe moeten de effecten van de maatregelen worden vastgesteld door deze te scheiden van andere factoren die ook van invloed zijn op de beleidsdoelen. De beleidstheorie wordt gebruikt als nadere uitwerking van de onderzoeksvragen. Belangrijke elementen daarin zijn:

- De beoogde effecten en eventuele neveneffecten. Dit is de zogenaamde outcome. Beoogde effecten zijn altijd afgeleid van de beleidsdoelen zoals weergegeven in de NoMo dan wel het beleidskader benutten. Hierbij gaat het om de effecten op bereikbaarheid (reistijden, betrouwbaarheid van de reistijd, robuustheid), veiligheid en milieu;
- De maatregel en relevante kenmerken van de maatregel die het gedrag van weggebruikers beïnvloeden. Dit is de output van het beleid;
- Uiteenlopende factoren die tussen de maatregel en de relevante beleidseffecten liggen (de intermediaire factoren);
- Versturende (confounding) factoren die van invloed zijn op de beleidsdoelen, maar die niet samenhangen met de maatregel;
- De inputs in termen van middelen en menskracht, van belang voor bepalend doelmatigheid van het beleid.

Figuur 1.1 geeft een eerste voorbeeld van een conceptueel model waarin de beleidstheorie is vastgesteld.

De beleidsdoelen van benutten zijn geformuleerd in de NoMo en het Beleidskader Benutten. Het kan zijn dat voor bijzondere projecten specifieke beleidsdoelen zijn geformuleerd, in al dan niet algemene termen. Mocht daar sprake van zijn dan is een specifieke doelanalyse van belang.

Soms is niet duidelijk wat men wil bereiken óf niet duidelijk waarom bepaalde instrumenten worden ingezet voor bepaalde doelstellingen. Bij zulk beleid is het zaak om – voorafgaand aan een evaluatieonderzoek – de beleidstheorie te reconstrueren. Bij dit ‘achteraf opstellen’ van een beleidstheorie wordt de bijdrage die een beleidsvoerder impliciet verwacht van zijn instrumenten en werkzaamheden om een bepaald probleem aan te pakken, expliciet of zichtbaar gemaakt. Hiervoor kan een doelenboom worden opgesteld. Merk op dat doelen wel in SMART-termen moeten worden geformuleerd om evalueerbaar te zijn.

De overige factoren kunnen worden benoemd en met elkaar in verband worden gebracht in een conceptueel model of een analyseschema. Daarin worden alle belangrijke factoren benoemd die van invloed zijn op de beleidsdoelen. Het beleidsprobleem kan helder in kaart worden gebracht door de relatie tussen instrumenten, tussendoelen en einddoel en door de relatie tussen externe factoren, doelen en neveneffecten in kaart te brengen. Bij evaluatieonderzoek naar de effectiviteit van beleid staat de relatie tussen een instrument (instrumentele variabele) en het beoogde effect (doelvariabele) centraal. Echter, op deze relatie werken veelal andere factoren in. In een evaluatieonderzoek zal zelden sprake zijn van een enkelvoudige causale relatie tussen instrumentele en een doelvariabele. Vaak verloopt het bereiken van effecten langs zogeheten ‘interveniërende variabelen’: tussengelegene effecten die onmisbare schakels vormen naar het

uiteindelijke effect. Ook kan van andere variabelen dan de instrumentele variabele een zelfstandige invloed uitgaan op de doelvariabele (rivaliserende variabele; externe factoren).

#### **Checklist kwaliteit uitwerking beleidstheorie**

1. Zijn de beleidsdoelen opgenomen als de te onderzoeken doelvariabelen. Zijn eventuele neveneffecten volledig weergegeven, bijvoorbeeld ten aanzien van specifieke doelgroepen;
2. Is de maatregel nauwkeurig weergegeven. Daarbij gaat het niet alleen om het instrument zelf, maar ook om de werking van het instrument;
3. Zijn interveniërende factoren benoemd en goed uitgewerkt. Hierbij gaat het om verkeerskundige en gedragsmatige factoren die tussen de maatregel en de doelen in staan;
4. Zijn externe factoren benoemd die de doelen beïnvloeden. Denk hierbij aan de factoren die van invloed zijn op de verkeersvraag (evenementen, seizoensfluctuaties en dergelijke) en op de capaciteit (andere maatregelen, tijdelijke capaciteitsfluctuaties, etc);
5. Zijn verbanden tussen de verschillende factoren goed benoemd. Denk hierbij ook aan samenhangen tussen de maatregelen en externe factoren.

## **2.4 Keuze van de onderzoeksopzet (hoofdpijnen)**

Doel van de onderzoeksopzet is het zoveel als mogelijk afbakenen van de invloed van het gevoerde beleid van andere factoren die van invloed zijn op de mate van doelbereiking. Daarvoor zijn drie stappen nodig:

1. **Uitwerking van de scope van de evaluatie.** Hierbij gaat het bijvoorbeeld om het vaststellen van het onderzoeks- en invloedsgebied, alsmede om de tijdsperiode die de evaluatie beslaat;
2. **Operationalisatie van de beleidstheorie.** Hierbij gaat het om de uitwerking van de beleidstheorie in concreet meetbare indicatoren;
3. **Het vaststellen van de onderzoeksopzet** waarmee de effecten van beleidsmaatregelen op de doelen kunnen worden vastgesteld. In termen van het conceptuele schema betekent dit het vaststellen van de relaties tussen de blokken. Merk op dat als nulsituatie de ontwikkeling zonder de betreffende maatregel wordt verondersteld.

### **Stap 1: Scope van de studie**

Bij de bepaling van de scope van de studie is het van belang aandacht te besteden aan de volgende aspecten:

1. De afbakening van het beschouwde netwerk. Daarbij wordt onderscheid gemaakt tussen het maatregelgebied en het invloedsgebied. Het invloedsgebied wordt gezien als dat deel van het netwerk waarin mogelijk effecten van de maatregel te verwachten zijn. Omdat veel weggebruikers over grotere afstanden reizen is het van belang het invloedsgebied voldoende groot te kiezen om een goed inzicht in de effecten van maatregelen te verkrijgen;
2. De onderzoeksperiode. Daarbij is van belang dat een periode wordt gekozen die voldoende groot is om de (beoogde en neven) effecten van de maatregelen te kunnen vaststellen. Bij een voor- en nasituatie geldt dit zowel voor de voorsituatie(s)

- als de nasituatie(s). Eventueel kan gewerkt worden met tussenmetingen. De onderzoeksperiode is van belang vanuit twee optieken. Ten eerste is deze van belang om rekening te houden met statistische fluctuaties in het verkeer. Ten tweede zijn er maatregelen die enige tijd vragen voordat de effecten volledig zichtbaar zijn;
3. Omstandigheden. Hierbij is de vraag of de maatregel niet alleen in reguliere situaties effecten sorteert of dat ook aandacht besteed moet worden aan bijzondere situaties (incidenten, evenementen, weer, etc).

## **Stap 2: Operationalisatie van de beleidstheorie**

In deze stap worden de elementen uit de beleidstheorie uitgedrukt in concreet meetbare indicatoren. Het belangrijkste zijn natuurlijk de beleidsdoelen. Maar ook de andere factoren moeten in meetbare termen worden uitgedrukt. Zo zijn de condities van belang om de uitkomsten te kunnen beoordelen. Daarbij gaat het om het vaststellen van specifieke kenmerken van het netwerk, de infrastructuur en omstandigheden die van invloed kunnen zijn op de beleidsdoelen en de effecten van de maatregel.

In de bijlage is een voorzet gegeven voor deze uitwerking.

## **Stap 3: Onderzoekopzet**

Er bestaat een aantal typen onderzoekopzetten waarmee de effectiviteit van beleid kan worden nagegaan. Mede afhankelijk van de beschikbaarheid van gegevens kan uit een van deze typen onderzoekopzetten een keuze worden gemaakt:

- een opzet gebaseerd op een voor- en nameting;
- een opzet gebaseerd op een met- en zonder (maatregel) meting, en;
- een opzet gebaseerd op gebruik van een controlegroep zonder de maatregel.

Opgemerkt kan worden dat de keuze voor één van bovenstaande onderzoekopzetten niet wordt bepaald door alleen strikt methodologische overwegingen (als dat wel het geval zou zijn, verdient de middelste opzet de voorkeur) omdat deze met de minste assumpties kan worden geevalueerd. De gegevens waarover men kan beschikken, zijn vaak bepalend voor de keuze van een bepaalde opzet. Soms zal men alleen in staat zijn tot een nameting, in andere gevallen beschikt men over omvangrijke tijdreeksen.

Informatie over controlegroepen is zeer zeldzaam: vaak gaat hier een echte experimentele onderzoekopzet aan vooraf (waarmee deze opzet veelmeer een ex ante karakter zal hebben). Daarnaast zullen in veel gevallen de beschikbare tijd, het onderzoeksbudget, de tijdsdruk om het probleem op te lossen en de kosten van onderzoek van grote invloed zijn. Ieder evaluatieonderzoek vergt maatwerk.

#### **Checklist Onderzoeksopzet (validiteit)**

1. Is sprake van een weloverwogen en heldere uitwerking van de scope van het onderzoek? Is sprake van voldoende omvang van het invloedsgebied? Is de onderzoeksperiode voldoende groot? Wordt aandacht besteed aan niet-reguliere situaties?
2. Worden de indicatoren in het beleidsmodel goed gemeten met de benoemde variabelen? Worden alle relevante grootheden gemeten?
3. Mag verwacht worden dat de gekozen onderzoeksopzet kan leiden tot geldige uitspraken over de effecten van de maatregelen en de condities waaronder deze optreden, zonder dat de uitkomsten kunnen worden verklaard door andere factoren en maatregelen (doeltreffendheid van de maatregelen).
4. Indien de doelmatigheid van de maatregel aan de orde is, worden de inputs in termen van geld en menskracht goed gemeten?

Vooraf het derde punt is geen eenvoudige opgave. Vaak worden maatregelen in combinatie genomen. Men kan dan onderzoek doen naar de effecten van het gehele programma. Echter, soms is het ook wenselijk om uitspraken over deelaspecten te verkrijgen. Indien bijvoorbeeld tegelijkertijd kleine infrastructurele maatregelen worden genomen met de invoering van Gecoördineerd Netwerkbreed Verkeersmanagement, dan kunnen de effecten van beide niet zonder meer worden gescheiden.

## **2.5 Dataverzameling**

De volgende stap betreft het verzamelen van de data waarmee de ontwikkelingen in de variabelen (geoperationaliseerde indicatoren) worden gemeten. Er zijn een groot aantal dataverzamelmethode beschikbaar. In de leidraad Evaluatie Benuttingsmaatregelen worden een groot aantal verschillende methoden beschreven.

Ten aanzien van deze richtlijnen is vooral het aspect van datakwaliteit van belang. Datakwaliteit is de mate waarin data geschikt is voor het doel waarvoor ze gebruikt wordt, nl het doen van onderbouwde uitspraken over de beleidseffecten van maatregelen.

De belangrijkste aspecten van datakwaliteit zijn:

- Case en item compleetheid. Mate waarin alle cases en items aanwezig zijn die volgens de inclusiecriteria aanwezig zouden moeten zijn;
- Validiteit (accuraatheid). De mate waarin data items meten wat ze zouden moeten meten.
- Variabiliteit (accuraatheid). De mate van consistentie van het verzamelen van een data item met onderscheid tussen between-observer and within-observer variabiliteit
- Tijdigheid. De mate waarin de data tijdig/up-to-date zijn.
- Representatie. De mate waarin de data in overeenkomstig formaat en in geschikte taal, symbolen en units is vastgelegd en de definities duidelijk zijn.
- Beschikbaarheid en toegankelijkheid. De mate waarin de data snel en betrouwbaar beschikbaar zijn.

Om deze criteria te toetsen is een expliciete kwaliteitscontrole van de beschikbaar gestelde gegevens nodig. Deze wordt vaak impliciet uitgevoerd, maar vanuit evaluatieoptiek is het wenselijk dat in de rapportage dit onderdeel expliciet aan de orde wordt gesteld. Indien sprake is van knelpunten worden deze gerapporteerd en wordt aangegeven hoe met eventuele knelpunten wordt omgegaan en wat de consequenties zijn voor de uitkomsten.

Bijzondere aandacht behoeft het onderzoek naar de kwaliteit van kwalitatieve gegevens die middels interviews worden verzameld

#### **Aandachtspunten bij enquêtes en interviews**

Het is belangrijk veel aandacht te besteden aan de vragenlijsten die worden gebruikt bij interviews en enquêtes. Deze dienen gebaseerd te zijn op de probleem- en doelstelling van het onderzoek en de beleidstheorie. Een adequate meting voldoet aan de volgende vereisten:

- het kiezen van goede indicatoren (geldigheid);
- het maken van een operationele definitie die: a) geen systematische afwijkingen vertoont ten opzichte van wat je wilt weten (validiteit) en b) geen onsystematische afwijkingen vertoont (betrouwbaarheid)
- Een beslissing over de vraagformulering. Houdt voor ogen dat alle vragen beantwoordbaar moeten zijn. Construeer begrijpbare, toepasbare vragen die beschikken over het juiste referentiekader van de respondenten.
- Een beslissing over de antwoordmogelijkheden. Kies tussen open, half open/half gesloten of gesloten vragen en bepaal welke antwoordmogelijkheden op welke manier worden aangeboden.
- Vaststellen van de vraagvolgorde: Begin met eenvoudig te beantwoorden vragen, zorg voor een logische opbouw, verplaats moeilijke vragen naar achter, let op mogelijke contexteffecten en eindig met een algemene open vraag.
- Ontwerp een lay-out voor de vragenlijst. Hou hierbij rekening met de methode van dataverzameling (persoonlijk, telefonisch, schriftelijk of web).
- Voer een pretest en een pilot onderzoek uit. Tijdens een pretest kan de vragenlijst door deskundigen worden beoordeeld of kunnen bepaalde onderdelen van de vragenlijst getest worden met behulp van kwalitatieve, cognitieve methoden. Tijdens een pilot onderzoek wordt het hele onderzoeksplan op kleine schaal uitgetest, van de selectie van onderzoekselementen tot en met de analyse van de verzamelde gegevens.

Op basis van het voorgaande kunnen de volgende kwaliteitchecks bij dataverzameling worden uitgevoerd:

### Checklist kwaliteit gegevens

1. Wordt duidelijk aangegeven hoe de onderzoeksgegevens worden verzameld (opzet en onderzoeksmethoden)? Berust het onderzoek bijvoorbeeld op primaire bronnen/veldwerk (systematische verzamelde gegevens direct bij doelgroep) of op secundaire bronnen/ bureauwerk (literatuurstudies, enquêtes onder deskundigen)?
2. Zijn de eenheden op een adequate wijze geselecteerd en verzameld, gezien het type uitspraken dat men in het onderzoek wil doen (bijvoorbeeld op willekeurige wijze getrokken en een steekproef van voldoende omvang)?
3. Zijn de kenmerken van de onderzochte eenheden op een adequate wijze gemeten en op een valide en achteraf controleerbare wijze vastgesteld, d.w.z. betrouwbaar (bij herhaling door anderen hetzelfde resultaat?) en valide (is onderzocht wat beoogd was te onderzoeken?)?
4. Wordt voldaan aan de gestelde eisen van datakwaliteit? Indien sprake is van knelpunten, wordt aangegeven wat consequenties zijn voor de uitkomsten?
5. Zijn kwalitatieve gegevens middels interviews en enquêtes op de gewenste wijze uitgevoerd?

## 2.6 Analyse van de data

Bij het analyseren van gegevens draait het om het beantwoorden van de onderzoeksvragen: kunnen op basis van de gegevens conclusies worden getrokken over de doelmatigheid en doeltreffendheid van beleid?

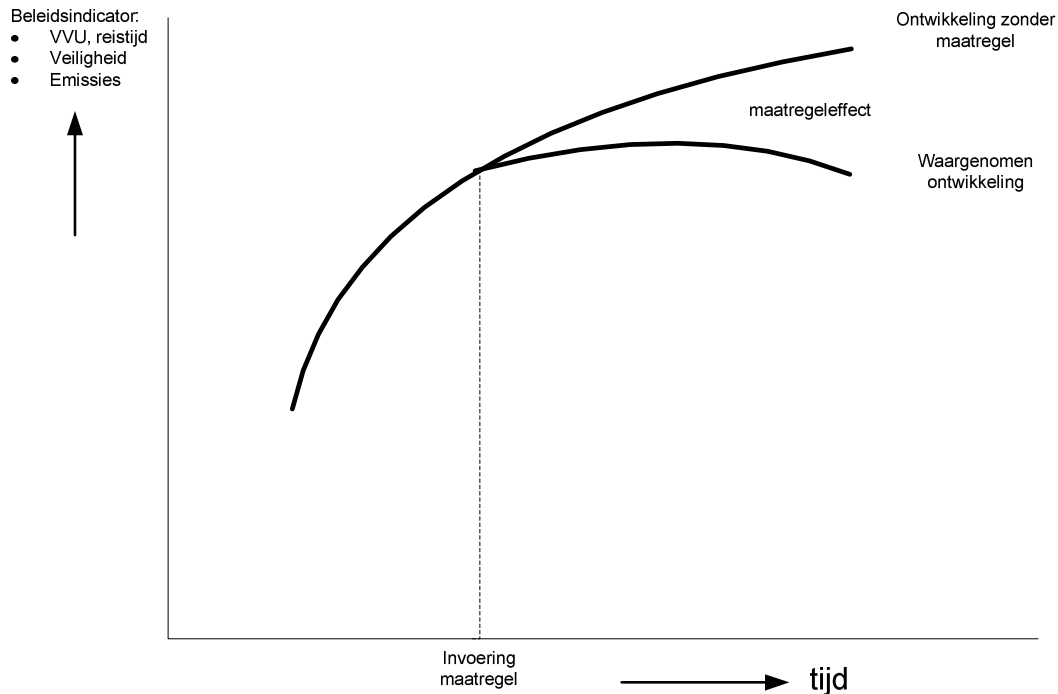
De analyse van gegevens vindt meestal plaats in een aantal stappen:

1. Vaststellen van het bereiken van de doelen. Hierbij wordt nagegaan welke ontwikkelingen optreden in de variabelen die de beleidsdoelen representeren. Zo wordt nagegaan welke ontwikkelingen in de bereikbaarheid (VVU, reistijden, etc), de veiligheid dan wel in milieuindicatoren zijn opgetreden. Soms worden dergelijke indicatoren nog uitgesplitst naar specifieke situaties en doelgroepen relevant voor een goed begrip van de ontwikkelingen. Indien sprake is van een controlegroep worden de ontwikkelingen zowel bij de proef als bij de controlegroep vastgesteld;
2. Vaststellen van de effectiviteit van de maatregel. Hierbij gaat het om het vaststellen van het effect van de maatregel door deze te scheiden van andere ontwikkelingen;
3. (indien van toepassing) het vaststellen van de doelmatigheid waarbij het gaat om de kosteneffectiviteit van de genomen maatregelen.

In figuur 2.1 zijn de belangrijkste elementen van de analyse schematisch weergegeven in het geval van een evaluatie op basis van een voor- en nastudie.



*Figuur 2.1: Ontwikkelingen in de maatregelindicator met en zonder beleidsmaatregel*



De figuur geeft weer dat middels metingen vastgesteld kan worden welke ontwikkelingen zich hebben voorgedaan in de beleidsindicatoren. Indien sprake is geweest van kwantitatieve beleidsdoelen kan op basis daarvan worden vastgesteld of de gestelde (outcome) doelen zijn gerealiseerd. Indien deze niet zijn gerealiseerd kan intensivering van de beleidsinspanningen wenselijk zijn.

In de volgende stap van de analyse wordt nagegaan wat het effect van de beleidsmaatregel is geweest bij de geconstateerde ontwikkeling. Daartoe wordt op basis van het ex-post onderzoek vastgesteld hoe de (hypothetische) ontwikkeling zou zijn geweest zonder de betreffende maatregel. Het verschil tussen beide geeft de beleidseffecten weer. Merk op dat tal van andere factoren ook van invloed kunnen zijn geweest op de beleidsindicator. Middels veelal multi-variate statistische analyses worden de beleidseffecten gescheiden van de effecten van 'overige' factoren zoals benoemd in het conceptuele model.

Om deze analyses uit te voeren zijn verschillende analysemethoden beschikbaar. Gedacht kan worden aan:

- **Regressie of andere uni/multivariate analyse.** Bij deze methode worden de ontwikkelingen in de beleidsindicatoren 'verklaard' uit de beleidsmaatregelen en andere factoren zoals benoemd in het conceptueel model. Deze methode wordt gehanteerd bij de analyses zoals die door het KIM worden uitgevoerd. Bij de methode kunnen veel specifieke veronderstellingen worden getoetst. Er wordt relatief direct ingegaan op de beleidsindicatoren. Merk op dat zelfs eenvoudige voor- en nametingen varianten zijn van deze benadering;
- **Modelmatige analyse.** Bij deze methode worden verkeersmodellen gebruikt om de effecten van maatregelen vast te stellen. Dit is een meer indirecte methode. Uit het evaluatieonderzoek worden middels analyses modelparameters afgeleid die eventueel

modelparameters uit ex-ante analyses vervangen. Vervolgens worden de ex-post effecten afgeleid door het model te draaien. Bij deze methode worden meer veronderstellingen gehanteerd dan bij de eerste methode, maar kunnen (voor reguliere situaties) netwerkeffecten eenvoudiger worden onderzocht en voorspeld.

- **Kwalitatieve methoden.** Door uiteenlopende oorzaken kunnen de hiervoor genoemde kwantitatieve onderzoeksmethoden niet bruikbaar zijn. Om dan toch uitspraken over beleidseffecten te kunnen doen kan een expertgroep van dienst zijn. Beleidseffecten worden dan bepaald op basis van zogenaamde 'intersubjectieve overeenstemming'. Daarbij is wel wenselijk om beschikbare gegevens maximaal te gebruiken om zoveel mogelijk informatie op tafel te hebben.

Ook ten aanzien van de analysefase kunnen kwaliteitseisen worden benoemd. Daarbij is de vraag of de analyseresultaten inzake het doelbereik, de effectiviteit van de maatregel en de doelmatigheid een goed beeld geven van de werkelijke ontwikkelingen en effecten. In het onderstaande staan de belangrijkste kwaliteitscriteria ten aanzien van de analysefase. Centraal daarbij staan de begrippen validiteit en betrouwbaarheid. Validiteit zegt iets over de inhoud: wordt er gemeten wat de bedoeling is. Betrouwbaarheid heeft te maken met de stabiliteit en significantie van het onderzoeksresultaat. Wanneer het onderzoek zou worden herhaald, komen dan dezelfde resultaten naar voren? Kortom, bij validiteit is de vraag naar systematische onderzoeksfouten aan de orde, bij betrouwbaarheid gaat het vooral om toevallige of statistische fouten.

Vaak wordt bij validiteit onderscheid gemaakt tussen interne en externe validiteit. Bij interne validiteit spelen vragen een rol die betrekking hebben op de sterkte van de verbanden tussen de maatregel en de beleidseffecten. Tevens is daarbij de vraag van belang of de maatregel een 'causaal' effect heeft op de beleidsdoelen. Externe validiteit heeft betrekking op de vraag of het effect ook in andere situaties van toepassing is en welke condities daarbij bepalend zijn.

#### **Checklist kwaliteit analyses**

1. Geven de analyses een betrouwbaar en valide beeld van de ontwikkelingen in de door de maatregel te beïnvloeden beleidsindicatoren? Worden op basis van deze uitkomsten valide conclusies getrokken over de ontwikkeling in de beleidsindicatoren?
2. Geven de analyses een valide beeld van de effecten van de beleidsmaatregel in termen van de beleidsindicatoren? Worden de geconstateerde effecten niet vertekend door het optreden van verstoringen? Zijn de condities voor het optreden van de effecten goed benoemd?
3. Zijn er uitkomsten geformuleerd die gebruikt kunnen worden als vuistregels bij ex ante bepaling van effecten in nieuwe situaties?
4. Wordt de werking van de maatregel volledig en helder beschreven, waarbij in het bijzonder aandacht geschonken wordt aan de werking van de maatregel richting de weggebruikers en eventuele knelpunten daarin;
5. Worden gedegen relaties gelegd en getoetst tussen de verschillende onderdelen in het model?

## **2.7 Rapportage**

Nadat de gegevens zijn verzameld en analyse van de gegevens heeft plaatsgevonden, is het mogelijk om conclusies te trekken en aanbevelingen te doen. Deze stap betreft de rapportage van de uitkomsten van evaluatieonderzoek. De daadwerkelijke benutting van – de uitkomsten van – evaluatieonderzoek, maar ook van reguliere prestatiegegevens, geldt als een kritische succesfactor voor goed evaluatieonderzoek. Het doen van evaluatieonderzoek en het hebben van prestatiegegevens zijn immers geen doelen op zich. Het gaat er uiteindelijk om dat de aldus gegenereerde informatie wordt benut en zo de kwaliteit van de besluitvormings- en verantwoordingsprocessen helpt verbeteren. Het doel van de vierde stap is zo ervoor te zorgen dat de uitkomsten van evaluatieonderzoek ex-post kunnen worden benut door politieke en bestuurlijke besluitvormers.

## Checklist kwaliteit rapportage

### *Bevindingen/conclusies*

- Worden de onderzoeksprobleemstelling en centrale vraagstelling beantwoord door de conclusies en aanbevelingen?
- Sluiten de bevindingen, conclusies en eventuele aanbevelingen op elkaar aan? (Is er sprake van interne consistentie?)

### *Gebruiksdoel*

- Is het gebruiksdoel van het evaluatieonderzoek omschreven (als onderdeel van de doelstelling van het onderzoek) en sluit het aan op de informatiebehoefte van de ambtelijke en politieke besluitvormers?
- Sluiten de opdrachtformulering en de onderzoeksvragen aan bij het gebruiksdoel van de evaluatie (d.w.z. dat de wijze waarop de opdracht tot het onderzoek (inclusief gegevensverzameling, gegevensanalyse, etc.) is geformuleerd 'logisch' voortvloeit uit het doel waarvoor het onderzoek wordt verricht)?
- Sluiten de uitkomsten en conclusies/aanbevelingen aan bij het gebruiksdoel van de evaluatie (d.w.z. zijn de conclusies/aanbevelingen toegesneden op het doel waarvoor het onderzoek wordt uitgevoerd)?
- Zijn de uitkomsten van het evaluatieonderzoek tijdig beschikbaar: sluiten ze aan op de bestuurlijk-politieke besluitvorming?

### *Presentatie*

- Worden de uitkomsten van het evaluatieonderzoek op een heldere en toegankelijke wijze gepresenteerd? (is het evaluatierapport goed leesbaar?)
- Zijn er beleidsaanbevelingen gedaan?
- Zijn de aanbevelingen concreet genoeg geformuleerd en technisch uitvoerbaar? Houden de aanbevelingen rekening met de politiek/maatschappelijke/financiële haalbaarheid?
- Worden de budgettaire gevolgen van de aanbevelingen aangegeven?
- Bevat het evaluatierapport een samenvatting en is deze representatief van aard? (geeft de samenvatting de hoofdlijnen van het rapport op evenwichtige wijze weer?)

### 3. Organisatie en uitvoering van evaluatie

Evaluëren, zowel wanneer sprake is van verantwoorden, als van leren, vraagt om een specifieke houding. De evaluaties worden geacht consistent te zijn met de volgende leidende principes:

- *Onafhankelijkheid*. De oordelen van de onderzoeker zijn niet beïnvloed door externe druk of een belangenconflict. Ook de schijn van verstrengeling dient vermeden te worden. Dit is een belangrijk principe bij evaluatie.
- *Objectiviteit*. De evaluatie is gebaseerd op verifieerbare feiten. Merk op dat het daarbij zowel om kwantitatieve feiten kan gaan als om meer kwalitatieve informatie. De onderzoeker streeft ernaar dat de gegevensbasis waarop de evaluatie is geënt geen onjuistheden bevat. De weergave van feiten dient duidelijk en herkenbaar gescheiden te worden van oordeelsvorming.
- *Organisatie*. Het evaluatieonderzoek dient zodanig te worden opgezet dat de onderzoekers zelfstandig onderzoek kunnen uitvoeren, conclusies kunnen trekken en aanbevelingen kunnen doen. De projectleider van het evaluatieonderzoek heeft geen belangen bij de uitkomsten van het onderzoek. Onderzoekers kunnen beschikken over alle informatie benodigd om een zelfstandig oordeel te kunnen vormen. De begeleidingsgroep is gericht op verhoging van de gebruikswaarde en de wetenschappelijke kwaliteit van het onderzoek. Vooraf worden duidelijke afspraken gemaakt over rollen en taken van alle betrokkenen;
- *Transparantie*. De evaluatie dient op een voor de belanghebbenden en derden navolgbare wijze te worden uitgevoerd. Vraagstelling, gegevensbasis, gebruikte onderzoeksmethoden, bevindingen en conclusies dienen in het rapport op een heldere, toegankelijke wijze te worden weergegeven.
- *Projectmatige kenmerken*. Het project is zodanig opgezet dat duidelijke afspraken worden gemaakt over de tijdsplanning, de fasering, informatiestromen tussen betrokkenen, externe communicatie en risico-beheersing;
- *Kosteneffectiviteit*. De evaluatie wordt uitgevoerd tegen de laagst mogelijke kosten met in achtneming van de overige leidende principes.

#### Checklist Organisatie en uitvoering

- Het evaluatieproject is zodanig opgezet dat aan genoemde leidende principes wordt voldaan