

Ex post evaluatie ten behoeve van verkeersveiligheidsbeleid

Datum 11 mei 2010
Status definitief

Colofon

Titel	Ex post evaluatie ten behoeve van verkeersveiligheid
Uitgegeven door	Dienst Verkeer en Scheepvaart
Informatie	DVS loket Telefoon 088 7982555 Email: dvsloket@rws.nl
	Contactpersoon DVS: drs. W. Vermeulen
Uitgevoerd door	Centrum voor Beleidsevaluatie, B&A groep
Auteur	D. Hanemaayer
Datum	20 mei 2010
Copyright	Rijkswaterstaat Dienst Verkeer en Scheepvaart Delft 2010

Disclaimer

De Dienst Verkeer en Scheepvaart (DVS), en degenen die aan deze publicatie hebben meegewerkt, hebben de in deze publicatie opgenomen gegevens zorgvuldig verzameld naar de laatste stand van wetenschap en techniek. Desondanks kunnen er onjuistheden in deze publicatie voorkomen. Rijkswaterstaat sluit iedere aansprakelijkheid uit voor schade die uit het gebruik van de hierin opgenomen gegevens mocht voortvloeien.

Inhoud

1	Beleidskaders	5
1.1	Context	5
1.2	Functies	6
1.3	Methodologie	7
1.4	Literatuur	10
1.5	Causaliteit	11
1.6	Conclusies	13
2	Evaluatie verkeersveiligheid	14
2.1	Evaluatie in het mobiliteitsbeleid	14
2.2	Evaluatie van verkeersveiligheidsbeleid	14
3	Toepassingen van ex post evaluatie in het verkeersveiligheidsbeleid	17
3.1	Procesevaluatie	17
3.2	Evaluatie doelbereik – ‘meten’	18
3.3	Evaluatie doeltreffendheid – ‘verklaren’	18
3.4	Evaluatie doelmatigheid	21
4	Model ex post evaluatie	22
4.1	Evaluatie light	30
	Bijlage	32

1 Beleidskaders

1.1 Context

Er was een periode waarin evaluatieonderzoek niet of nauwelijks onderdeel uitmaakte van het overheidsbeleid. Swanborn¹ positioneert het begin van een evaluatietraditie bij de Commissie voor de Ontwikkeling van de Beleidsanalyse die in de jaren zeventig actief was. De COBA was (onder andere) op zoek naar rationaliteit in het overheidsbeleid: het formuleren van eenduidige doelstellingen van overheidsbeleid. Op basis daarvan zouden strategie en instrumentarium eenvoudig af te leiden zijn. De COBA moest echter ervaren dat het beleid niet voldeed aan de eisen van rationaliteit en logica².

Het Platform Beleidsanalyse was in de jaren tachtig de opvolger van de COBA. De COBA heeft – aldus Swanborn - vanuit het ministerie van Financiën de rol en betekenis van beleidsevaluatie onder de aandacht gebracht, met primaire aandacht voor de technisch-industriële sector en met een sterke financieel-economische oriëntatie. Kennelijk was een en ander onvoldoende om het evaluatie-onderzoek onder de aandacht te brengen. Swanborn verwijst naar rapporten van de Commissie Hoofdstructuur Rijksdienst (1980) en de Nota Gedrags- en Maatschappijwetenschappen (1984). In beide rapportages wordt geconstateerd dat het evaluatieonderzoek nog niet tot ontwikkeling is gekomen.

De Commissie Programma Evaluatie functioneerde, op basis van een Kamer-motie, in de periode 1986 – 1991, met als afsluiting het Eerste nationaal Evaluatiecongres (1991). Tijdens die periode besteedde ook de Algemene Rekenkamer³ (voor het eerst) aandacht aan departementale beleidsevaluaties. Daarnaast wijst Swanborn op het Congres van de Vereniging van Bestuurskunde in 1987 over Macht en onmacht van bestuurlijke evaluaties.

Swanborn constateert eind jaren negentig – ten tijde van het verschijnen van de eerste druk van het boek – dat '... het evaluatie-onderzoek in Nederland een van de krachtigste sectoren (is) van sociaal-wetenschappelijk onderzoek'⁴, gelet op de inbedding van het evaluatieonderzoek in het beleid.

Na deze constatering is de betekenis van evaluatieonderzoek binnen de rijksoverheid alleen maar toegenomen, met het ministerie van Financiën en de Algemene Rekenkamer als drijvende krachten. Op lokaal en regionaal niveau is een dergelijke tendens minder herkenbaar, hoewel halverwege het lopende decennium wel bij gemeenten, provincies en waterschappen rekenkamers zijn ingericht.

Op rijksniveau is het (ex post) evaluatieonderzoek verankerd in de Comptabiliteitswet 2001, de Regeling Prestatiegegevens en Evaluatieonderzoek (RPE) en vaak in sectorwetten.

¹ Onderstaand historisch overzicht is ontleend aan: P.G. Swanborn (2004), *Evalueren. Het ontwerpen, begeleiden en evalueren van interventies: een methodische basis voor evaluatieonderzoek*. Amsterdam: Boom.

² N.A.A. Braakman (1993), *De sociogenese van de instrumenten van het beleid*. Elders betitelen Van Heffen en Klok de COBA als 'berucht': O. van Heffen en P.J. Klok (1998), *Beleidswetenschap als modernistisch hoogtepunt*. *Bestuurskunde* 7(2).

³ Verslag van de Algemene Rekenkamer over 1990 (1991). (TK 22032, 1-2)

⁴ Swanborn (2004), p. 37.

Voor de rijksbegroting vraagt de Comptabiliteitswet om gegevens inzake doeltreffendheid en doelmatigheid (voor het beleid en voor de apparaatskosten), al dan niet verkregen uit evaluatieonderzoek (artikel 5,3). Verder worden in artikel 20 de ministers verantwoordelijk gesteld voor de doeltreffendheid en doelmatigheid van het beleid en voor het periodiek onderzoek van doeltreffendheid en doelmatigheid. De Algemene Rekenkamer dient over deze onderzoeken tijdig op de hoogte te worden gesteld. In artikel 21 worden de ministers verantwoordelijk gesteld voor de doelmatigheid van de bedrijfsvoering, en voor het uitvoeren van onderzoek naar de bedrijfsvoering. Een en ander dient ook een plaats te krijgen in de verantwoordingsstaat in de departementale jaarverslagen (artikel 54, 3). De doelmatigheid van het financieel beheer, het materieelbeheer en de administraties is onderwerp van onderzoek door de departementale accountantsdienst (artikel 66, 2).

De Comptabiliteitswet bepaalt voorts dat de Algemene Rekenkamer onderzoek doet naar de doeltreffendheid en de doelmatigheid van het gevoerde beleid en de doelmatigheid van het financieel en materieel beheer, de administraties en de organisatie (artikel 85).

Artikel 20 van de Comptabiliteitswet is uitgewerkt in de Regeling periodiek evaluatieonderzoek en beleidsinformatie 2006 (RPE, inmiddels opgenomen in de Regeling Rijksbegrotingsvoorschriften 2008). De RPE is in 2006 aangepast op grond van de evaluatie van het VBTB-proces (2004) waaruit naar voren kwam dat het denken in termen van – realisatie van – beleidsdoelen nog onvoldoende van de grond was gekomen. Nieuw in de RPE 2006 ten opzichte van de RPE 2001 is de beleidsdoorlichting: de evaluatie van een integraal beleidsterrein (begrotingsartikel), aan de hand van een aantal ' simpele doch fundamentele'⁵ doorlichtingsvragen⁶, en met een rol voor onafhankelijken in de uitvoering. De Tussenevaluatie beleidsdoorlichting⁷ heeft onder andere uitgewezen dat de uitgevoerde beleidsdoorlichtingen weliswaar aan de in de RPE gestelde eisen voldoen, maar zwakten laten zien op meting van de netto beleidseffectiviteit en dat de causaliteit tussen beleidsmaatregelen en effecten vaak niet aannemelijk is gemaakt.

1.2 Functies

In het algemeen worden aan de ex post evaluatie twee functies toegekend: verantwoorden en leren. Verantwoording wil zeggen dat via een ex post evaluatie het bestuur de volksvertegenwoordiging (en anderen) informeert over uitvoering en resultaten van beleid; de verantwoording kan ook betrekking hebben op de middelenbesteding en de verhouding tussen kosten en baten. De lerende evaluatie heeft als functie om positieve en negatieve ervaringen in ontwikkeling en uitvoering van beleid te inventariseren, teneinde positieve punten in de toekomst te benutten en negatieve punten te voorkomen / vermijden. Het kan gaan over alle aspecten van het beleid zoals beleidsvoorbereiding, instrumentenkeuze, organisatie van de uitvoering. Lessen kunnen betrekking hebben op het geëvalueerde beleidsdossier maar kunnen ook betrokken worden in andere dossiers. Hoewel lastig hard te maken, is in het algemeen is de ervaring dat de combinatie van beide functies moeizaam is op het punt van de opstelling en bijdragen van degenen die verantwoordelijk zijn (waren) voor het dossier.

⁵ Regeling periodiek evaluatieonderzoek en beleidsinformatie 2006. Staatscourant 28 april 2006.

⁶ Zie paragraaf 1.2.

⁷ Tussenevaluatie beleidsdoorlichting. Den Haag: APE, 2008

De verantwoordingsfunctie leidt vaak tot een defensieve houding van de beleidsverantwoordelijken, de lerende functie vraagt om een open houding van de beleidsverantwoordelijken.

1.3 Methodologie

In deze paragraaf wordt beschreven wat in de regelgeving en in de praktijk op rijksniveau de aanpakken zijn voor ex post evaluaties. Aan de orde komen de aanpak volgens de RPE, de aanpak volgens de Algemene Rekenkamer en thema's in de (wetenschappelijke) literatuur. Ten slotte worden enkele opmerkingen over doelmatigheidsonderzoek gemaakt.

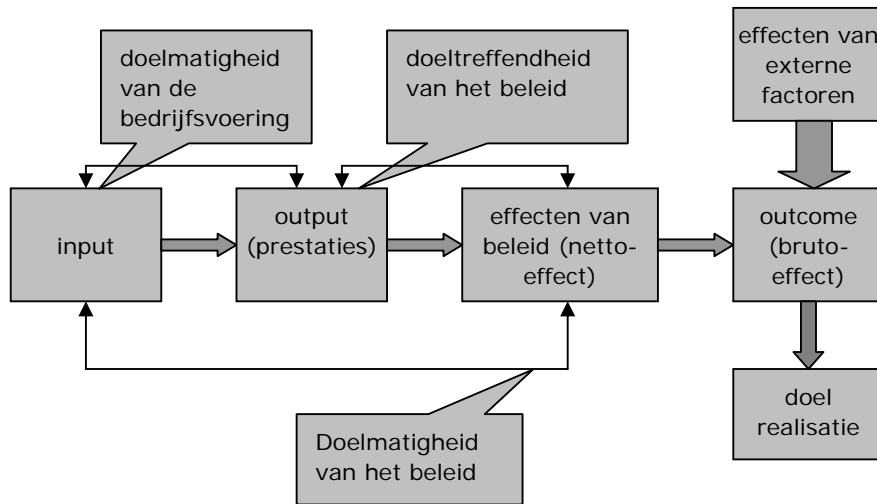
RPE

Effectenonderzoek ex-post wordt gedefinieerd als ' Het op systematische en methodologisch verantwoorde wijze onderzoeken en beoordelen van de effecten van beleid in de maatschappij of overheidsorganisatie'. Voor de systematisering biedt de RPE twee ingrediënten.

In de eerste plaats schrijft de RPE voor dat in doorlichtingsonderzoek – als bijzondere vorm van effectenonderzoek ex post – de volgende vragen beantwoord moeten worden:

- . Wat is het probleem dat aanleiding is (geweest) voor het beleid? Is dit probleem nog actueel?
- . Wat is de oorzaak van het probleem?
- . Waarom rekent de overheid het tot haar verantwoordelijkheid om het probleem op te lossen?
- . Waarom ligt de verantwoordelijkheid op rijksniveau (en niet op decentraal of EU-niveau?) Hoe is de verantwoordelijkheid vormgegeven en waarom?
- . Welke doelstelling heeft de overheid geformuleerd voor de oplossing van het probleem?
- . Welke instrumenten worden ingezet? Hoe is de samenhang tussen de instrumenten? Is er sprake van overlap? Wat is bekend over de uitvoering van het beleid en over de doelmatigheid van de bedrijfsvoering?
- . Wat is het effect van de instrumenten op de geformuleerde doelstellingen (oplossing van het probleem)? Hebben instrumenten op andere beleidsterreinen ook belangrijke effecten op de geformuleerde doelstellingen? Wat zijn belangrijke positieve en negatieve neveneffecten?
- . Hoe is de hoogte bepaald van de budgetten die zijn ingezet? Wat is hiervan de onderbouwing?

In de tweede plaats biedt het RPE een aantal kernbegrippen die in onderstaande figuur zijn geordend:

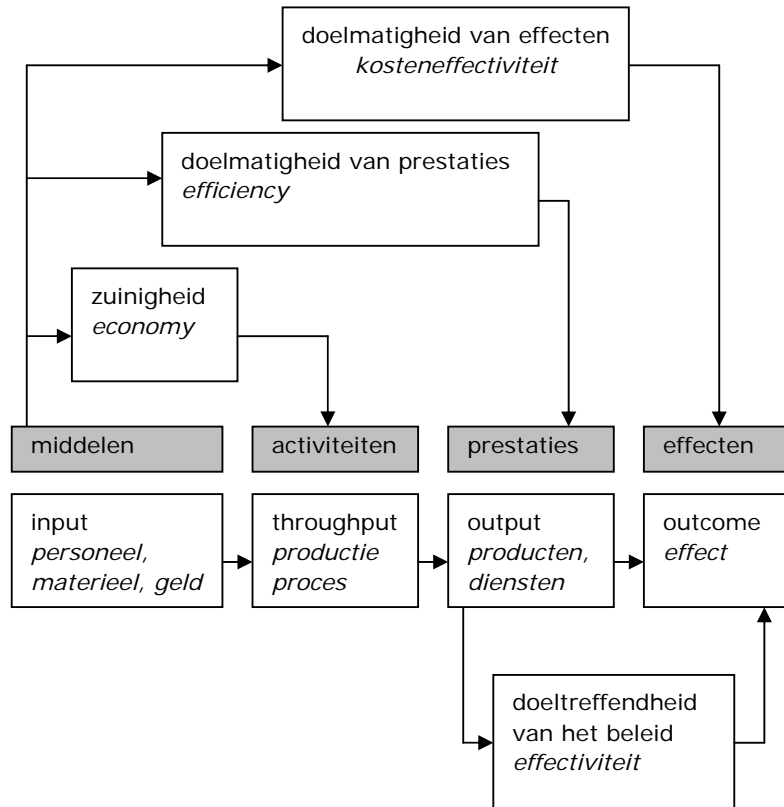


Figuur 1: Begrippenkader

Algemene Rekenkamer

De Algemene Rekenkamer⁸ heeft een model opgesteld voor de uitvoering van onderzoek naar doelmatigheid en doeltreffendheid, hetgeen overeenkomt met ex post evaluatie. De kern van het model is in het schema op de volgende bladzijde samengevat.

⁸ Algemene Rekenkamer (2005), Handleiding onderzoek naar doelmatigheid en doeltreffendheid.



Figuur 2: Productieproces in de publieke sector (bron: Ministerie van Financiën, 2001)

Het model is ontleend aan een eerdere versie van de RPE.

De ARK onderscheidt vier typen ex post evaluaties:

- . Doelrealisatie: de na uitvoering van het beleid ontstane maatschappelijke situatie wordt vergeleken met hetgeen met het beleid beoogd wordt. Niet wordt nagegaan of het beleid de oorzaak is van de ontstane situatie (causaliteit). In het verband van onderzoek naar doelrealisatie wordt ook verwezen naar doorwerkingsonderzoek: onderzoek in de context van decentrale beleidsuitvoering (provincies, gemeenten, waterschappen) naar de vraag of de voorwaarden van rijkswege voor de beleidsuitvoering hebben gewerkt.
- . Doelgroepbereik: indien onderzoek naar de einddoelen van beleid lastig is, is onderzoek naar doelgroepbereik een goede second-best optie. Het betreft onderzoek naar tussendoelen. Vaak is beleid gericht op gedragsverandering; het bereik van de doelgroep door het beleid is dan een noodzakelijke voorwaarde om tot gedragsverandering – en maatschappelijke effecten – te komen.

- . Doeltreffendheid: dit is onderzoek naar de vraag '... in hoeverre dankzij het regeringsbeleid de gewenste beleidsresultaten worden bereikt'⁹. 'Beleidsresultaten' omvat output en outcome. Deze vorm van onderzoek is complex, zeker wanneer het rijksbeleid onderdeel is van een lange beleidsketen en wanneer er externe invloeden zijn.
- . Doelmatigheid: in doelmatigheidsonderzoek worden de kosten beoordeeld die gemaakt zijn ter uitvoering van beleid. Het kan betrekking hebben op doelmatigheid van beleid en doelmatigheid van bedrijfsvoering.

1.4 Literatuur

In de (wetenschappelijke) literatuur worden verscheidene stromingen onderscheiden.

Hospes¹⁰ onderscheidt drie benaderingen voor de ex post evaluatie.

. Evidence-oriented approach

In deze benadering ligt de nadruk op het meten van veranderingen die aan het beleid kunnen worden toegeschreven; de benadering is sterk kwantitatief georiënteerd. De beleidsdoelen worden als uitgangspunt genomen. Als voorbeeld worden de IOB-onderzoeken van het ministerie van Buitenlandse Zaken aangehaald. Op deze benadering – die o.i. in de Nederlandse praktijk van beleidsevaluatie¹¹ in pure vorm nog maar weinig voorkomt – is veel kritiek geleverd¹². Deze benadering biedt geen inzicht in het waarom van slagen of falen van beleid (verklaringen). Deze benadering sluit niet aan op de context waarin het maken en uitvoering van beleid inmiddels geen overheidsmonopolie meer is maar waarin organisaties en burgers een rol spelen. En inzake de beleidsdoelstellingen – startpunt in de beleidsevaluatie – geldt dat deze vaak onduidelijk en onmeetbaar zijn, dat deze in de loop van de tijd vaak verschuiven, en dat zeker in een multi-actor context ook doelstellingen zelf onderdeel van evaluatie zouden moeten zijn.

. Realistic evaluation

In reactie op de beperkingen van de evidence-oriented benadering hebben in 1997 Pawson en Tilly¹³ een pleidooi gevoerd voor de realistic evaluation. Het gaat in deze benadering er vooral om om inzicht te krijgen in de verklaringen waarom beleid / maatregelen wel of niet werkzaam zijn. In de realistic evaluation wordt gezocht naar de agents, de werkzame bestanddelen: wat vindt er in de werkelijkheid plaats tussen input en output? Kernbegrippen zijn Context – Mechanismen – Outcome. De context heeft betrekking op de omstandigheden (en externe factoren) waaronder een programma of beleid wordt uitgevoerd; mechanismen hebben betrekking op veelal gedragseffecten als gevolg van de programma- of beleidsinterventies; de outcome is het resultaat van het programma. De realistic evaluation 'kruipt' in de uitvoering, terwijl de evidence-oriented aanpak daar langsheen gaat. Het gaat om een theorie-gedreven vorm van evaluatie: de evaluatie wordt gevoerd op basis van

⁹ ARK 2005, p. 5

¹⁰ O. Hospes (2008), Evaluation evolution? www.thebrokeronline.eu

¹¹ Te onderscheiden van programma-evaluatie in bijvoorbeeld onderwijs en welzijn.

¹² Zie o.a. E.H. Klijn en J. Koppejan (2001), Besluitvorming en management in netwerken: een multi-actor perspectief op sturing. In T. Abma en R. In 't Veld, Handboek Beleidswetenschap. Amsterdam: Boom. En: T. Abma (2001), Variaties in beleidsevaluaties. In T. Abma en R. In 't Veld, Handboek Beleidswetenschap. Amsterdam: Boom.

¹³ Pawson en Tilly (1997), Realistic Evaluation.

een beleids- of programmatheorie over de werking van maatregelen; in de evaluatie wordt deze theorie getoetst.

Realistic evaluation wordt in het algemeen op programmaniveau toegepast, minder op het niveau van beleidsevaluaties. Binnen Nederland is Research voor Beleid 'ambassadeur' van deze benadering.¹⁴

Realistic evaluation sluit nauw aan bij de benadering van de Theory of Change, waarin de beleidstheorie als het te evalueren object gezien wordt.

. Complexity evaluation¹⁵

In reactie op de realistic evaluation is de 'complexity evaluation'¹⁶ ontstaan. Als gevaar van de realistic evaluation wordt gewezen op het risico dat de context als niet meer dan een externe factor beschouwd wordt omdat in de praktijk veelal alle aandacht uitgaat naar de invloed van een programma op de leden van de doelgroep. Terwijl in complexe beleidssituaties deze externe omstandigheden volwaardig het verloop en de uitkomsten van beleid meebepalen. Daarnaast dient in de aanpak van een evaluatie oog te zijn voor verschuivingen in doelen en strategie tijdens de looptijd van een programma of beleidsuitvoering. En er moet oog zijn voor het feit dat op veel terreinen beleid gevoerd wordt door verscheidene departementen; voor veel beleidsopgaven geldt een gedeelde betrokkenheid en verantwoordelijkheid. Dit vraagt om interactief beleid; evaluatie van interactief beleid vereist een multidisciplinaire aanpak. Meting van oorzaak en gevolg is in complexe beleidssituaties moeilijk; de optiek in dergelijke evaluaties zou moeten verschuiven naar het leren vanuit ervaringen en de ontwikkeling van alternatieve, responsieve stijlen van beleid maken.

. Onderhandelend evalueren

Gemeenschappelijk in de drie bovenstaande benaderingen is dat het de evaluator is die een oordeel over het beleid velt. Guba en Lincoln¹⁷ voerden in 1989 een pleidooi voor de vierde generatie evaluatie: onderhandelend evalueren¹⁸. Kernpunten in een onderhandelende evaluatie zijn 'claims', 'concerns' en 'issues' van belanghebbenden; het zijn deze CCI's die in de evaluatie centraal staan, niet de doelstellingen van beleid. De evaluatie wordt uitgevoerd in een sterk iteratief proces met alle belanghebbenden, waarbij de evaluator vooral de rol krijgt van facilitator. Onderhandelend evalueren is gericht op de toekomst, op basis van lessen uit het verleden.

1.5 Causaliteit

Kernvraag in vrijwel iedere ex post evaluatie is de causaliteit tussen beleid en realiteit. In de wetenschappelijke literatuur wordt deze doorgaans benaderd als kwantitatieve opgave, zoals de benadering in Swanborn (2004)¹⁹:

¹⁴ Een voorbeeld binnen Nederland is E. Sol (2008), Op zoek naar werkzame bestanddelen: een zoektocht naar de heilige graal. Amsterdam: UVA – Hugo Sinzheimer Instituut.

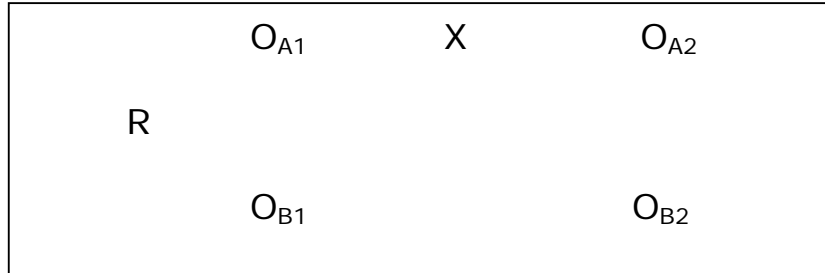
¹⁵ Ook wel aangeduid als pluralistische evaluatie.

¹⁶ Barnes M, et al. (2003) Evidence, understanding and complexity: Evaluation in non-linear systems. Evaluation 9(3): 265–284. SAGE Publications. Het pleidooi vertoont gelijkenissen met het pleidooi van Ringeling (Het imago van de overheid. 1993) voor de multipelen evaluatie: beoordeling van waarde en kwaliteit vanuit verschillende perspectieven.

¹⁷ Guba, E. & Y. Lincoln, Fourth Generation Evaluation, Beverly Hills, 1989.

¹⁸ T. Abma (2000), Onderhandelend evalueren. 'Fourth generation evaluation' van Egon Guba en Yvonna Lincoln. Bestuurskunde 9, nr. 8 .

¹⁹ Swanborn, op. cit., pp. 233 e.v.



Figuur 3: Het standaardontwerp: ontwerp met controlegroep(en) en voor- en nameting

De R staat voor random selectie van onderzoeksgroepen, de X voor de te evalueren (beleids)interventie en de O voor onderzoeksgroepen. Waarbij O_A de groep is op wie de interventie is toegepast en O_B de controlegroep, en waarbij O_1 de onderzoeksgroep is voor de interventie en O_2 na de interventie. In deze benadering wijzen statistische maten uit in welke mate onderzoeksgroep O_A is beïnvloed door interventie X.

Deze aanpak is vaak niet goed van toepassing op beleidssituaties, bijvoorbeeld omdat een nulmeting niet beschikbaar is of omdat onderzoeksgroepen A en B niet onderscheiden kunnen worden. Doorgaans zijn dergelijke kwantitatieve benaderingen wel geschikt voor medisch of psychologisch onderzoek.

Wel zijn er methoden ontwikkeld om binnen het kader van vergelijking van kwantitatieve gegevens tot zo nauwkeurig mogelijke metingen te komen. TNO/SEO²⁰ noemen vier typen methoden om bij het ontbreken van de counterfactual toch effecten van beleidsmaatregelen vast te kunnen stellen:

- . Het sociale experiment: op een aantal leden van de doelgroep wordt een beleidsmaatregel toegepast, op een ander deel niet. Voor deze selectie gelden procedures. De verschillen tussen beide groepen, na toepassing van de maatregel, zijn het gevolg van de maatregel.
- . Natuurlijke of quasi-experimenten: het verschil met het sociale experiment is dat de selectie van de beide onderzoeksgroepen niet volgens procedures tot stand komt maar door een (externe) omstandigheid.
- . Cross sectie variatie: hierbij wordt gebruik gemaakt van een onderscheiding in de tijd binnen de doelgroep, bijvoorbeeld wanneer een regeling (tijdelijk) leeftijds specifiek is.
- . Variatie in de tijd: het betreft vergelijking in de tijd van gegevens op persoonsniveau: voor en na de maatregel.

Een andere – meer inhoudelijke - benadering wordt geboden in de realistic evaluation, waarvan een kernpunt is dat in de evaluatie processen op micro-niveau ‘ontleed’ worden en waarin gezocht wordt naar werkzame bestanddelen.

²⁰ TNO/SEO (2008), Handboek meetmethoden voor effectiviteit van activerend arbeidsmarktbeleid op persoonsniveau. Een handvat voor beleidsmakers en onderzoekers voor het vaststellen van de (netto)effectiviteit van activerend arbeidsmarktbeleid.

Te allen tijde zal gelden dat een spijkerhard bewijs voor de betekenis van overheidsmaatregelen niet te achterhalen valt, maar dat vaak - langs kwantitatieve en/of kwalitatieve weg - wel een zodanig inzicht ontstaat dat met redelijke zekerheid conclusies kunnen worden getrokken over de betekenis van een beleidsinterventie.

Doelmatigheidsonderzoek

In de praktijk van het evaluatieonderzoek in publieke sectoren valt op dat over het karakter van doelmatigheidsonderzoek uiteenlopende denkbeelden bestaan. Zowel in de RPE als in de ARK-handleiding zijn twee vormen van doelmatigheid onderscheiden: doelmatigheid van de bedrijfsvoering (input <> output / prestaties) en doelmatigheid van het beleid (input / prestaties <> effecten van beleid). Daarbij duidt de ARK doelmatigheid aan als een relatief begrip: er zijn geen absolute normen, doelmatigheid wordt beoordeeld aan de hand van vergelijkingen in de tijd of met andere objecten.

In de praktijk van de beleidsevaluatie wordt doelmatigheid meestal anders omschreven, namelijk als de kwaliteit van het beleidsproces, verbonden aan indicatoren als tijdigheid, daadkracht en verhoudingen tussen beleidspartners. Daarnaast is er ook de interpretatie van doelmatigheid als de mate waarin het rijk heeft bijgedragen aan het doelbereik (dus in feite de doeltreffendheid). Analyses van specifiek de efficiency van de middelenbesteding, zoals aangeduid in de RPE en door de ARK worden in beleidsevaluaties zelden gevraagd.

1.6 Conclusies

Bovenstaande tekst bevat ingrediënten voor de ontwikkeling van de aanpak van ex post evaluaties op het gebied van verkeersveiligheid. In beginsel bevatten de rijkskaders goede aangrijpingspunten in de begrippen doeltreffendheid en doelmatigheid. Daarnaast bieden ook andere benaderingen aangrijpingspunten, vooral de realistic evaluation, omdat daarin gezocht wordt naar gedragscomponenten die bij verkeersveiligheidsmaatregelen van groot belang zijn.

2 Evaluatie verkeersveiligheid

2.1 Evaluatie in het mobiliteitsbeleid

Onderzoek van het Kennisinstituut voor Mobiliteitsbeleid heeft uitgewezen dat ex post evaluaties naar mobiliteitseffecten bij infrastructurele maatregelen zelden uitgevoerd worden²¹. Hiervoor zijn vier redenen:

- . Politici zijn meer geïnteresseerd in de toekomst dan in het verleden, en zijn al helemaal niet geïnteresseerd in mogelijk negatieve uitkomsten
- . Professionals lopen bij voorkeur niet het risico van reputatieschade bij eventuele negatieve uitkomsten, zeker omdat in de voorbereidingsfase professionals (te) optimistisch zijn over de resultaten van een project.
- . Capaciteitsgebrek en personeelsverloop in de organisatie die het project tot stand heeft gebracht
- . Methodologische problemen om de causaliteit te bepalen en keuzeproblemen inzake de timing van de ex post evaluatie.

Het KiM hecht aan ex post evaluaties van infra-projecten. Ex post evaluaties (en meta-evaluaties) leveren bouwstenen voor nieuwe infra-projecten. Met het oog hierop doet het KiM een aantal voorstellen om de bezwaren tegen of belemmeringen voor ex post evaluaties te ondervangen. Onder meer wordt gesuggereerd een lichte vorm van ex post evaluatie toe te passen.

2.2 Evaluatie van verkeersveiligheidsbeleid

Deze paragraaf gaat over evaluatie van verkeersveiligheidsbeleid en –maatregelen. De paragraaf bevat een aantal opvallende waarnemingen over de wijze waarop de evaluatie functie wordt ingezet binnen het verkeersveiligheidsbeleid.

De basis van deze paragraaf wordt gevormd door korte beschrijvingen van een aantal ex post evaluaties; deze zijn opgenomen in de bijlage. Het gaat om toegepast onderzoek van concrete maatregelen. De selectie is betrekkelijk willekeurig: een aantal evaluaties is verstrekt door DVS, een aantal is ontleend aan Tien jaar Duurzaam Veilig en een aantal is gegoogled onder de trefwoorden evaluatie en verkeersveiligheid. Deze paragraaf is gebaseerd op een kleine steekproef, dus bij de onderstaande beelden past enig voorbehoud.

Onderzoekprogrammering

Duurzaam veilig is de visie van de SWOV en sinds 1998 via het convenant tussen rijk, provincies en gemeenten een belangrijke bouwsteen in het verkeersveiligheidsbeleid. Uit de tussenbalans (SWOV 2009) blijkt dat op veel punten ex post evaluaties worden uitgevoerd. Maar tegelijkertijd is niet de indruk ontstaan van een systematische aanpak van de evaluatie; het geheel maakt een wat fragmentarische indruk. Niet altijd lijkt basisinformatie beschikbaar te zijn, niet op alle programmaonderdelen wordt (met diepgang) geëvalueerd (bijvoorbeeld educatie). En ook niet wordt in de tussenbalans duidelijk hoe uitkomsten van ex post evaluaties doorwerken in het beleid cq. de inrichting van het programma.

²¹ Kennisinstituut voor Mobiliteitsbeleid (2009), Na het knippen van het lint. Het ex post evalueren van infrastructuur. Den Haag

Het in het KiM-rapport opgeroepen beeld over ex post evaluaties (zie paragraaf 2.1.: veel ex ante, weinig ex post) is voor het verkeersveiligheidsbeleid niet herkenbaar: ex post evaluaties worden in redelijke aantallen uitgevoerd terwijl ex ante evaluaties minder worden uitgevoerd, althans in de vorm van publiek toegankelijke publicaties.

Evaluatietypen

In termen van de RPE en de ARK is het meeste onderzoek te classificeren als doeltreffendheidsonderzoek, waarbij soms de nadruk ligt op (alleen) doelbereik of doelgroepbereik.

Doeltreffendheidsonderzoek start bij beleidsdoelstellingen. In de meeste verkeersveiligheidsevaluaties ontbreekt een duidelijke formulering van de te evalueren beleidsdoelen; waarmee overigens niet gezegd is dat die er niet waren; het kan ook zijn dat in de ex post evaluatie weinig aandacht is besteed aan de doelen. Dit geldt in mindere mate voor de ex post evaluatie van infra-maatregelen, en meerdere mate voor maatregelen in de sfeer van voorlichting en educatie. Het betekent dat in de evaluaties het beoogde doelbereik strikt genomen niet kan worden getoetst.

Wat verder opvalt is dat er in vrijwel geen enkele ex post evaluatie sprake is van een expliciet beschreven beleidstheorie (al dan niet op basis van een ex ante evaluatie of een impact assessment) waar resultaten mee kunnen worden vergeleken. En waarmee de mate van doelbereik (deels) zou kunnen worden verklaard. De beleidstheorie geeft de veronderstelde causale ketens van maatregel naar maatschappelijk effect. Aan de hand hiervan kan na analyse worden verklaard waarom maatregelen wel of niet tot de beoogde effecten leiden.

Meer dan eens missen conclusies scherpheid en helderheid inzake doelbereik cq. beoordeling van hetgeen bereikt is (mede bij gebrek aan duidelijke doelstellingen).

In de ex post beleidsevaluaties die in dit kader zijn gezien, wordt geen aandacht besteed aan doelmatigheid. Kosten van maatregelen komen niet aan de orde, evenmin als kosten-baten verhoudingen. De evaluaties zijn gericht op effecten en/of procesgang, niet op de kostenkant.

Methodiek

De opzet van de meeste evaluaties sluit niet aan bij het causaliteitsmodel zoals gegeven in het eerste hoofdstuk (paragraaf 1.3. onder causaliteit). Dit komt mede omdat quasi-experimentele opzetten niet makkelijk gerealiseerd kunnen worden.

Wat betreft de gegevens in ex post evaluaties wordt in de meeste onderzoeken goed gebruik gemaakt van triangulatie, bijvoorbeeld de koppeling van statistische

gegevens, soms gedragsobservaties, expertinterviews en gebruikersonderzoek (focusgroepen, enquêtes). Gebruik van uiteenlopende methoden zijn nodig voor het typen vraagstelling in verkeersveiligheidsevaluaties vanwege de complexe samenhangen (interactie mens, voertuig, weg). Wel is de vraag of gedragsobservaties een grotere rol zouden kunnen spelen in ex post evaluaties (in lijn met de benadering van de realistic evaluation – zie hoofdstuk 1).

Ex post evaluaties zijn overwegend gericht op het te evalueren systeem (mens – weg – voertuig). Zelden komen externe factoren aan de orde, zeker niet in de vorm waarin gestructureerd gezocht wordt naar externe omstandigheden en hun betekenis. Dat geldt ook voor de meta-evaluatie (Tien jaar Duurzaam Veilig).

3 Toepassingen van ex post evaluatie in verkeersveiligheidsbeleid

In dit hoofdstuk wordt uiteengezet wat de toepassingsmogelijkheden en betekenissen kunnen zijn van ex post evaluatie in het verkeersveiligheidsbeleid. Voor het beleid wordt het Strategisch Plan Verkeersveiligheidsbeleid 2008-2010²² als uitgangspunt genomen. Het strategisch plan vormt het ambitieuze ('Wereldkampioen blijven') kader voor de lange termijn verbetering van de verkeersveiligheid. De aanpak rust op drie pijlers: aanpak duurzaam veilig, samenwerking, integrale aanpak.

Ten behoeve van verkeersveiligheidsbeleid wordt veel onderzoek uitgevoerd. In het plan wordt verwezen naar lopende programma's bij ondermeer SWOV, CROW, TNO, planbureaus, Kenniskamer V&W, Trimbos, universiteiten, en anderen. In het kader van deze notitie is niet nagegaan in hoeverre ex post evaluatie hierin aan de orde is.

Over ex post evaluatie wordt in het plan opgemerkt dat er '... nog wel behoefte (is) aan kennis over de werkelijke effecten (ex post) van maatregelen' (p. 45); dit is ook opgenomen in de generieke maatregelen (p. 46).

Ofschoon in dit hoofdstuk de nadruk ligt op ex post evaluatie van maatregelen, is de scope wel wat breder. De vraag die hier beantwoord wordt, is wat ex post evaluatie kan bijdragen aan verkeersveiligheidsbeleid.

3.1 Procesevaluatie

Verkeersveiligheidsbeleid is een complex werkkterrein met veel actoren en betrokkenen. Het feit dat samenwerking één van de pijlers is, duidt hierop. Naarmate de omvang en/of intensiteit van samenwerkingsrelaties (op programma- of projectniveau) toeneemt, kan de vraag naar de uitvoering van de samenwerking pregnanter spelen.

Evaluatie van de samenwerking (als verbijzondering van procesevaluatie) kan soms zinvol zijn, zeker als er indicaties zijn dat deze niet soepel verloopt en/of niet tot de beoogde resultaten leidt. De procesevaluatie is dan interventie-gericht. Een procesevaluatie zal duidelijk maken:

- . Hoe feitelijk de samenwerking verloopt
- . Wat obstakels zijn om tot een soepele samenwerking te komen

In een procesevaluatie kan in de eerste plaats vastgesteld worden wat de feitelijkheid in de samenwerking is, hoe deelnemers deze percipiëren en hoe (on)productief de samenwerking is. Vaststelling van de feiten is nuttig omdat rond en binnen samenwerkingsverbanden gemakkelijk beelden kunnen ontstaan die niet of maar gedeeltelijk met de werkelijkheid overeenstemmen. Het zal duidelijk worden of er reden tot interventie is.

Evaluatie van de oorzaken van gebrekkige samenwerking vraagt om verdiepend onderzoek op aspecten als de inrichting van de uitvoeringsorganisatie, de communicatielijnen, de taakverdeling, de randvoorwaarden, en voorts de

²² Ministerie van Verkeer en Waterstaat (2008), Strategisch Plan Verkeersveiligheid 2008 – 2010. Van, voor en door iedereen.

transparantie over dit alles voor de deelnemers. Het onderzoek gaat over feiten en percepties; in principe komen probleemanalyses en oplossingsrichtingen tegelijkertijd aan de orde. Het heeft kenmerken van de in hoofdstuk 1 genoemde aanpak van 'onderhandelend evalueren'.

De procesevaluatie – zoals hier uitgewerkt – is dus een interventie in een lopend beleidsproces, gericht op verheldering en verbetering: ex durante, terugkijkend maar ook vooruitkijkend. In deze vorm wordt de evaluatie alleen uitgevoerd wanneer daar aanleiding toe is. Daarnaast komt het ook voor dat een procesevaluatie na afloop van een beleidproces uitgevoerd wordt. Dan is de functie minder specifiek en gericht op het in algemene zin borgen van de lessen die ten behoeve van toekomstige situaties uit de samenwerking getrokken kunnen worden.

3.2 Evaluatie doelbereik – 'meten'

Evaluatie van het doelbereik – zoals omschreven door de ARK (zie hoofdstuk 1) – houdt in dat vastgesteld wordt of het beoogde doel wel, deels of niet behaald is. Dit gegeven heeft twee functies: (1) verantwoording aan de volksvertegenwoordiging en anderen over de resultaten van beleid en (2) signalering of voortgezet beleid cq. aanpassing van het beleid wenselijk is. In de doelbereikevaluatie wordt alleen het doelbereik gemeten; in de doeltreffendheidsevaluatie - zie de volgende paragraaf - wordt verklaard waarom het doelbereik is zoals het is.

De evaluatie van het doelbereik vindt zowel na afloop van een beleidsproces plaats als tijdens. Dan wordt doorgaans gesproken over monitoring. Hierbij kan overigens opgemerkt worden dat beleid doorgaans een continue karakter heeft waarbij het verschil tussen 'tijdens' en 'na' vaak niet scherp gemarkeerd kan worden.

Of doelbereik is vast te stellen is afhankelijk van de duidelijkheid van het beoogde doel. Dat kan te wensen overlaten waardoor de mate van doelbereik niet goed is vast te stellen. Beleidsformuleringen als 'verbetering van ...' of 'stimuleren van ...' maken het niet altijd mogelijk om te bepalen of de gewenste situatie bereikt is.

Doelbereik kan gemeten worden op verschillende schaalniveaus. Bijvoorbeeld: met gemeenten kan afgesproken worden om in bepaalde situaties kruispunten te vervangen door rotondes. Doelbereik kan gemeten worden in termen van het aantal gerealiseerde rotondes en in termen van de daling van ongevallen en slachtoffers op de betreffende locaties.

3.3 Evaluatie doeltreffendheid – 'verklaren'

De evaluatie van de doeltreffendheid (de mate waarin beleidsmaatregelen hebben bijgedragen aan een maatschappelijk effect) gaat een forse stap verder dan de evaluatie van het doelbereik. De kern van de doeltreffendheidsevaluatie is om de mate van doelbereik te verklaren.²³

²³ Overigens: in de praktijk van de ex post evaluatie worden doelbereik en doeltreffendheid vaak in één project onderzocht.

De functies van doeltreffendheidsonderzoek zijn:

- . Het geven van verantwoording aan de volksvertegenwoordiging en anderen over de bereikte resultaten en de invloed daarop van beleidsmaatregelen (Identiek aan de evaluatie van het doelbereik)
- . Het trekken van lessen uit ervaringen, met het oog op (gewenste) verbetering van de maatregel
- . Het trekken van lessen uit ervaringen, met het oog op continuering respectievelijk verdere implementatie van de maatregel

Er zijn altijd twee en meestal drie factoren die maken dat het niet goed mogelijk is om een 100 procent zeker antwoord te geven op de vraag naar de mate waarin een beleidsmaatregel heeft bijgedragen aan een maatschappelijk effect. Daarbij is het uitgangspunt dat de inzet van een beleidsmaatregel per definitie is het beïnvloeden van het gedrag van verkeersdeelnemers, hetzij rechtstreeks (bijvoorbeeld het dragen van autogordels), hetzij indirect (bijvoorbeeld het toezicht op het dragen van autogordels).

De eerste factor is dat menselijk gedrag wel waargenomen kan worden (zie hierboven: evaluatie doelbereik) maar niet zondermeer verklaard kan worden. Het feitelijke verkeersgedrag in relatie met een beleidsmaatregel kan waargenomen en beschreven worden, maar het is niet goed mogelijk om met volledige zekerheid te zeggen dat gedrag of gedragsverandering het gevolg is van de beleidsmaatregel.

De tweede, verbonden, factor is dat beleidsmaatregelen genomen worden in een dynamische omgeving. Er zullen altijd (externe) factoren en ontwikkelingen zijn die, naast de beleidsmaatregel, ook van invloed zijn op het gedrag dat onderwerp is van de betreffende beleidsmaatregel.

De derde factor is dat doorgaans vanuit het beleid het gedrag wordt beïnvloed met een pakket aan maatregelen (bijvoorbeeld de koppeling tussen voorlichting en toezicht). In geval van een combinatie van maatregelen is extractie van het effect van één maatregel niet of lastig.

In de beleidspraktijk wordt deze wezenlijke beperking in de mogelijkheden van doeltreffendheidsanalyses meer dan eens gebruikt als 'excuus' om er maar helemaal niet aan te beginnen. Dat is ten onrechte. Er zijn verscheidene mogelijkheden om via onderzoek zodanig 'in de buurt' te komen van de door beleidsmaatregelen teweeg gebrachte effecten, dat in veel gevallen deze inzichten voor het beleid bruikbaar zijn.

Voorwaarde om zo dicht mogelijk in de buurt te komen is het evaleren aan de hand van een mix aan gegevens en methoden, ook wel aangeduid als triangulatie:

- . Gegevens over doelbereik: maatschappelijke doelen (zoals aantal slachtoffers / ongevallen)
- . Gegevens over doelbereik op tussendoelen (voor zover gesteld)
- . Gegevens over (geleverde) beleidsprestaties
- . Referentiegegevens over doelbereik, te verkrijgen uit experimenten, buitenlandse ervaringen en ervaringen in andere sectoren, een en ander: voor zover van toepassing
- . Gegevens over doelgroepbereik (zie ook ARK): de bekendheid van verkeersdeelnemers met de beleidsmaatregel
- . Gegevens inzake gedrag van verkeersdeelnemers op basis van gedragsobservaties
- . Gegevens inzake gedrag van verkeersdeelnemers op basis van gerapporteerd gedrag (survey)

- . Indrukken van verkeersdeelnemers (focusgroepen, T11, etc.)
- . Indrukken van deskundigen: ervaringsdeskundigen zoals rijinstructeurs en verkeersveiligheidsexperts (zoals onderzoekers)
- . Externe factoren: een systematische analyse van factoren die naast de beleidsmaatregel van invloed kunnen zijn op het gerealiseerde effect, opgevolgd door gegevensverzameling hierover.

Deze werkwijze is eerder in deze notitie (paragraaf 1.3) aangeduid als de complexity evaluation ofwel pluralistische evaluatie. De inrichting van de evaluatie zal van geval tot geval verschillen, afhankelijk van beschikbaarheid en verwerfbaarheid van gegevens en afhankelijk van de randvoorwaarden voor de evaluatie. Gevolg van deze benadering is wel dat evaluaties omvangrijk en kostbaar kunnen worden. Het is aan de opdrachtgever om te bepalen hoe kosten en baten zich verhouden. Voor dit type evaluaties zijn geen standaard-analyseschema's voorhanden; van geval tot geval moet de beste aanpak gekozen worden, met als randvoorwaarden: gegevens over maatschappelijke effecten vormen het startpunt van de analyse en alle beschikbare gegevens moeten in de analyse worden betrokken. Het is aan de evaluator om aan de lezer glashelder te maken hoe de analyse is verlopen.

Een pluralistische evaluatie is zeer gebaat bij een duidelijke beleidstheorie die de redenering bevat hoe maatschappelijke effecten bewerkstelligd kunnen worden met als redeneerstappen bijvoorbeeld: maatschappelijk doel < tussendoelen < beleidsprestaties < instrumenten, en daarnaast de lijn van mogelijke externe factoren.

In een light versie²⁴ van de doeltreffendheidsevaluatie wordt een beperkt deel van de bovenstaande gegevensbronnen gebruikt. Welke dat zijn, verschilt van geval tot geval. Wel blijft ook in de light versie van belang dat met uiteenlopende typen gegevens gewerkt wordt: één van de bouwstenen van de doeltreffendheidsevaluatie is in welke mate verschillende soorten gegevens dezelfde richting uitwijzen.

De basisgegevens zullen in alle gevallen zijn het doelbereik op maatschappelijke effect en tussendoelen. En daarnaast doelgroepoordelen (bijvoorbeeld in de vorm van focusgroepen) en (in sommige gevallen) gedragsobservaties. Ook in de light versie is aandacht voor externe effecten – al is maar in bird's eye view – geboden.

Het bovenstaande leidt tot het beeld dat doeltreffendheidsevaluaties uitvoerbaar zijn, al is nooit met 100 procent zekerheid of nauwkeurigheid te zeggen wat de invloed van een maatregel is. Maar redelijke en plausibele schattingen kunnen doorgaans wel gemaakt worden. Naar typen maatregelen bezien verschillen de mogelijkheden om een duidelijk beeld te krijgen van de maatschappelijke effecten. De effectiviteit is beter op te sporen naarmate:

- de gevolgen van de beleidsmaatregel direct waarneembaar zijn, bijvoorbeeld de bromfietshelm versus het bromfietsexamen (de gevolgen kunnen direct geobserveerd worden)
- de aantallen slachtoffers die aan een beleidsmaatregel gerelateerd zijn groter zijn (bij relatief grote aantallen slachtoffers is de kans op merkbare getalsmatige veranderingen groter dan bij kleine aantallen dan bij kleine aantallen)
- maatregelen naar lokatie beperkt zijn, bijvoorbeeld spitsstroken versus snelheidslimieten (locationeel beperkte maatregelen maken zeer gericht onderzoek mogelijk)
- de beleidsevaluatie beter is voorbereid: beleidstheorie, nulmeting, en dergelijke

²⁴ Zie KIM-rapport

3.4 Evaluatie doelmatigheid

Evaluatie van de doelmatigheid is in de kern de toetsing tussen enerzijds de effecten van een maatregel en anderzijds de kosten. Waarbij onder kosten niet alleen geld, maar ook personele inzet verstaan wordt. Het element van de effecten is hierboven uiteengezet. De doelmatigheidsevaluatie is in zoverre overzichtelijk dat de effecten en de kosten met elkaar worden vergeleken. Er zijn drie typen mogelijk voor de beoordelende vergelijking, mede afhankelijk van het doel van de evaluatie:

- . De verhouding tussen kosten en baten kan berekend worden langs de lijn van monetarisering van de baten (het aantal vermeden slachtoffers / ongevallen, uitgedrukt in een geldbedrag).
- . De vergelijkende analyse van de kosten en baten tussen verschillende (typen) beleidsmaatregelen (hier hoeven de baten niet te worden gemonetariseerd)
- . De vergelijkende analyse van de kosten en de baten van een maatregel in de tijd (in de situatie dat een instrument zelf in ontwikkeling kan zijn, bijvoorbeeld voorlichting, handhaving, educatie)

De functies van doelmatigheidsonderzoek zijn:

- . Het geven van verantwoording aan de volksvertegenwoordiging en anderen over de bereikte resultaten en de daartoe bestede middelen
- . Het trekken van lessen uit ervaringen, met het oog op de keuze van die beleidsinstrumenten met de beste kosten-baten verhouding

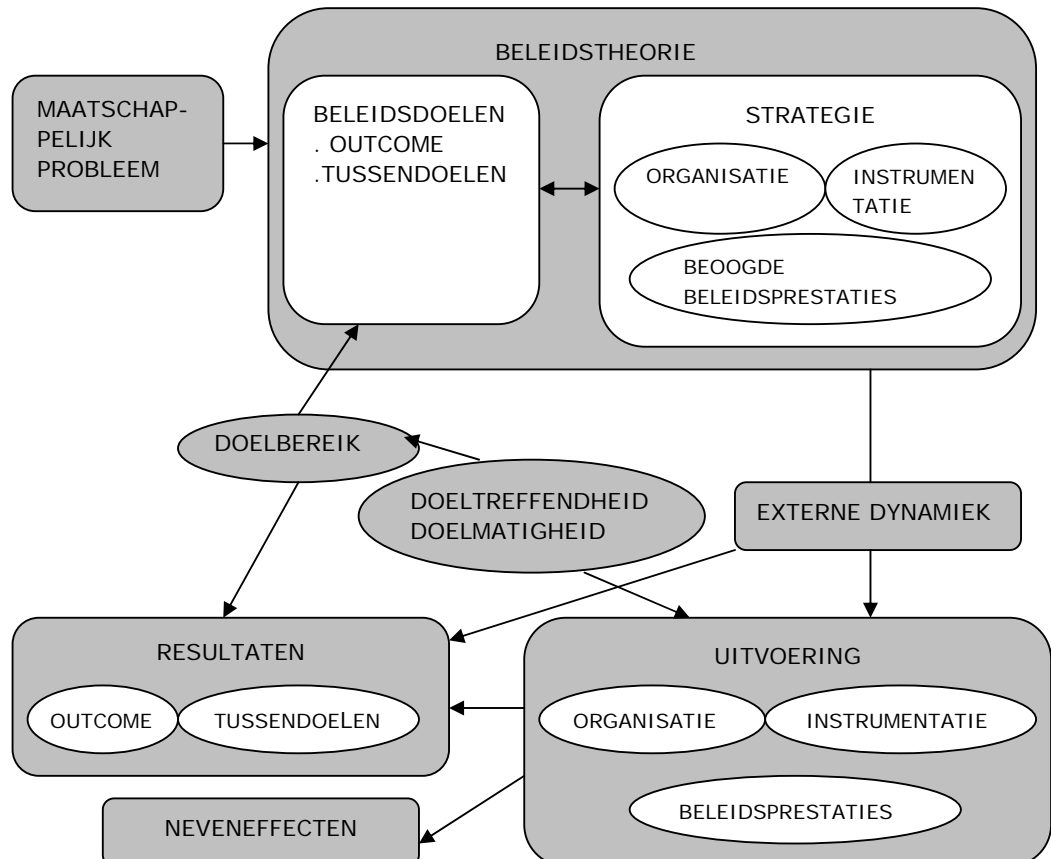
4 Model ex post evaluatie

Dit hoofdstuk bevat een universeel bruikbaar model voor de ex post evaluatie. Het model bevat alle elementen van een beleidsprogramma, van het maatschappelijke probleem tot de resultaten van het beleid. De basis van het model is beleidslogica, dat wil zeggen logische samenhangen tussen probleem, doelen, strategie, uitvoering en resultaten.

Het model is gebaseerd op een rationeel uitgangspunt: meting en verklaring van effecten zoveel mogelijk op basis van feiten en cijfers, waarbij beleidsdoelen het beoordelingskader vormen. Deze uitgangspunten staan tegenover benaderingen als een op interactie gebaseerde aanpak van de evaluatie, oordeelsvorming door betrokkenen, en evaluatie als onderhandelingsproces. Op enkele punten worden elementen van andere benaderingen opgenomen, maar het uitgangspunt blijft een rationalistische wijze van evalueren.

De praktijk is dat beleid lang niet altijd volgens deze rationaliteit of logica tot stand komt. Het model is dan ook niet bedoeld als een norm voor hoe beleid gevoerd zou moeten; dat zou een naïeve benadering zin. Het model is bedoeld om – terugkijkend – de samenhangen in het beleid (of het ontbreken daarvan) op te sporen en om langs die weg de uitkomsten van het beleid te meten en te verklaren waarom het beleid tot stand gebracht heeft wat tot stand gebracht is.

Het model laat zich als volgt verbeelden:



Hieronder worden de elementen uit het model beschreven.

. *Maatschappelijk probleem*

De evaluatie start met een goede beschrijving van het maatschappelijk probleem dat tot een oplossing gebracht moet worden. In de beschrijving van het maatschappelijk probleem wordt aangegeven:

- . wat het probleem is
- . voor wie het een probleem is
- . (indien van toepassing) waar het een probleem is
- . wat de omvang van het probleem is
- . waarom het probleem bestaat, wat de oorzaken zijn

Uitvoering van de ex post evaluatie

In principe is in de documentatie, opgesteld bij de start van het beleidsproces, beschreven wat het maatschappelijk probleem is. In de ex post evaluatie hoeft hier dus geen eigen onderzoek naar verricht te worden. Alleen als het beleid zonder duidelijke probleemstelling van start is gegaan, is enige reconstructie nodig, bijvoorbeeld via interviews met personen die bij de beleidsontwikkeling betrokken waren.

. *Beleidstheorie*

De beleidstheorie bevat de beleidsreactie op het maatschappelijke probleem. Het omvat doelstellingen en strategie: de situatie die het beleid wil bereiken plus de weg daar naar toe, gebaseerd op een analytische en samenhangende redenering van hoe het doel bereikt kan worden.

o Doelstellingen

De doelstellingen geven aan wat de door het beleid gewenste maatschappelijke situatie is. Doelstellingen zeggen iets over de situatie in de maatschappij, niet iets over het beleid. Bij voorkeur zijn doelstellingen exact geformuleerd, niet alleen om doelbereik meetbaar en beoordeelbaar te maken maar ook om – in het beleidsproces - alle beleidspartners duidelijk te maken wat de te bereiken situatie zou moeten zijn. Meetbaar wil zeggen dat uit de doelformulering duidelijk blijkt welke de indicatoren (kwantitatief of kwalitatief) zijn om de resultaten van het beleid aan af te meten. Beoordeelbaar wil zeggen dat duidelijk is welke score op een indicatoren gewenst wordt.

Doelen zijn geformuleerd in maatschappelijke termen. Meestal zal er ook sprake zijn van tussendoelen. Tussendoelen kunnen gezien worden als de voorwaarden om de beoogde doelen (outcome) te bereiken, ook wel genoemd de implementatieketen. Vanuit het oogpunt van de ex post evaluatie kunnen tussendoelen handzaam zijn in geval de outcome lastig meetbaar is, of in geval de causaliteit tussen beleidsmaatregelen en outcome moeilijk te traceren is. In dat geval vormen de tussen doelen verbindende schakels.

Onderdeel van de ontwikkeling van de beleidstheorie is beantwoording van de vraag of gegevens beschikbaar zijn of kunnen komen om doelbereik te meten.

o Strategie

De strategie bevat de weg om tot realisatie van doelstellingen te komen. Het bevat drie elementen: de organisatie van het beleid, het in te zetten instrumentarium en de beoogde beleidsprestaties. De organisatie van het beleid heeft betrekking op de verdeling van verantwoordelijkheden, taken en rollen in de uitvoering. Het instrumentarium heeft betrekking op de inzet van wetgeving, geld, communicatie, voorzieningen, en dergelijke. Het instrumentarium omvat de in te zetten instrumenten: geld, regelgeving, voorlichting, voorzieningen, en dergelijke. De (beoogde) beleidsprestaties zijn de concrete opbrengsten van de beleidsacties (output).

De strategie zal gebaseerd zijn op een analyse van de oorzaken van het maatschappelijk probleem, leidend tot inzichten over hoe oorzaken weggenomen of geredresseerd kunnen worden.

Uitvoering van de ex post evaluatie

Opstelling van de beleidstheorie (doelen en strategie) is uiteraard geen onderdeel van de ex post evaluatie. De doelen en de strategie hebben in de evaluatie wel de functie van toetsingskader, en zijn als zodanig onderdeel van de evaluatie.

In principe is in de documentatie opgesteld bij de start van het beleidsproces beschreven welke de doelen en strategie zijn. In de ex post evaluatie dient hier dus geen zelfstandig onderzoek naar verricht te worden. Alleen als het beleid zonder duidelijke doelen of duidelijke strategie van start is gegaan, is reconstructie nodig, bijvoorbeeld via interviews met personen die bij de beleidsontwikkeling betrokken waren.

Wel kan een analyse van de beleidstheorie zinvol zijn, omdat daaruit zal blijken wat de sterke en zwakke punten zijn. Deze bieden aanknopingspunten voor het vervolg van de evaluatie. Het onderstaande blok bevat bij wijze van voorbeeld de vragen die over de beleidstheorie gesteld kunnen worden.

Terzijde: meer dan eens zou het een wat te rigide aanpak zijn om alleen te toetsen aan de oorspronkelijke doelen en strategie. In ieder beleidsproces treedt voortschrijdend inzicht op. In de aanpak van de ex post evaluatie moet hiervoor ruimte zijn.

Lijst van vragen ter analyse van de beleidstheorie²⁵

1. Hoe ziet de programmatheorie van het project Overlastdonatie eruit?
 - a. Welk probleem of welke problemen dient het Instrument Overlastdonatie op te lossen? Welke doelen zijn hiervoor geformuleerd en op welke doelgroepen hebben deze betrekking?
 - b. Wat zijn de veronderstelde oorzaken van de gesignaleerde problemen?
 - c. Wat zijn de veronderstelde doel-middelrelaties waarmee de gewenste doelen worden gerealiseerd?
 - d. Welke contextuele factoren beïnvloeden de hiervoor genoemde relaties?
2. Is de programmatheorie van het project Overlastdonatie geldig en beleidsmatig realistisch?
 - a. Zijn de problemen, doelen en doelgroepen voldoende geëxpliciteerd?
 - b. Zijn de veronderstelde causale en finale relaties logisch consistent?
 - c. Zijn de veronderstelde causale en finale relaties empirisch houdbaar?
 - d. Zijn de centrale variabelen voldoende manipuleerbaar?
 - e. Zijn er (ongewenste) neveneffecten te verwachten als gevolg van dit project (bijv. verplaatsingseffecten)?

. *Resultaten*

Het blok 'resultaten' heeft betrekking op hetgeen bereikt is: realisatie van outcome en tussendoelen. Dit wordt gemeten in dezelfde termen als opgenomen in de beleidstheorie.

Doorgaans zijn outcome en tussendoelen goed of redelijk (met redelijke mate van betrouwbaarheid) meetbaar. Maar soms doen zich situaties voor waar dat een dermate omvangrijke dataverzameling vraagt die redelijkerwijs niet uitvoerbaar is. Dit kan het geval zijn wanneer het beleid betreft dat verbonden is aan kleine slachtofferaantallen waarvoor een omvangrijke onderzoekspanning nodig is om verandering in deze (kleine) aantallen te kunnen meten. Dan kan een second best oplossing zijn om experts en anderszins betrokkenen te raadplegen.

²⁵ B. Rovers en R. Boers (2009), Plan- en procesevaluatie van Project Opverlastdonatie. Den Haag, WODC

Uitvoering van de ex post evaluatie

In het ideale geval is bij aanvang van de evaluatie ervoor gezorgd dat resultaatgegevens verzameld (kunnen) worden, en heeft een nulmeting plaatsgevonden. Is dat het geval, dan volstaat in het kader van de ex post evaluatie een een-meting. Is dat niet het geval, dan dient binnen de ex-post evaluatie ook de methodiek voor de resultaat-meting ontwikkeld te worden. Gegevensbronnen ten behoeve van de resultaatmeting zijn bijvoorbeeld bestaande statistische gegevens, surveys, tellingen, observaties, interviews, etc.

In geval een nul-meting ontbreekt zal in de een-meting een schatting gemaakt moeten worden van de 'delta', de verandering die is opgetreden in de periode sinds de start van het beleidsproces.

. *Neveneffecten*

In de ex post evaluatie moet oog zijn voor neveneffecten die kunnen optreden bij de uitvoering van beleid. Dit is een open vraag, dat wil zeggen dat de antwoorden meestal niet in een evaluatiekader kunnen worden voorgeprogrammeerd. Wel kan de beleidstheorie vermoedens van mogelijke neveneffecten bevatten.

Uitvoering van de ex post evaluatie

De meest voor de hand liggende mogelijkheid om neveneffecten op het spoor te komen, is om hierover te spreken met personen die in de uitvoering een rol spelen; doorgaans hebben zij daar een goed inzicht in.

Het zal van geval tot geval verschillen in hoeverre apart onderzoek nodig is naar de werking en waardering van neveneffecten. Vaak zullen kwalitatieve indicaties (van experts, belanghebbenden, etc.) volstaan.

. *Uitvoering*

Onder de noemer 'uitvoering' wordt in de ex post evaluatie nagegaan hoe de uitvoering verlopen is. Dit omvat dezelfde elementen als de strategie: organisatie, instrumenten en beleidsprestaties.

Evaluatie van de organisatie wordt ook wel aangeduid als procesevaluatie. Er wordt nagegaan of en hoe beleidspartners hun verantwoordelijkheden, taken en rollen vervuld hebben, of de voornemens in de strategie zijn uitgevoerd (dan wel of daar beredeneerd van is afgeweken), hoe de onderlinge communicatie verlopen is, en dergelijke.

Evaluatie van het instrumentarium houdt in dat geïnventariseerd wordt welke instrumenten zijn ingezet, door wie, op welke schaal, en tegen welke kosten. De beleidsprestaties worden met specificaties geïnventariseerd. De specificaties verschillen van geval van geval; het kan bijvoorbeeld gaan om aangepaste regelgeving, uitvoering van een voorlichtingscampagne, aanleg van een spitsstrook, et cetera.

Uitvoering van de ex post evaluatie

In een ex post evaluatie gaat van een groot deel van het werk zitten in het verwerven van inzicht in de uitvoering.

Het onderdeel organisatie kan bestaan uit:

- . raadpleging documentatie: vergaderverslagen, voortgangsrapportages, en dergelijke
- . interviews met participanten in de uitvoering

Het onderzoek naar de organisatie van de uitvoering vindt plaats aan de hand van een lijst van algemene aandachtspunten, zoals transparantie (in taken, rollen, verantwoordelijkheden en procesgang), daadkracht, tempo, mandatering, leiderschap, en dergelijke, toegespitst op de betreffende situatie. Alleen bij evidente problemen in de organisatie van de uitvoering kan het nodig zijn diepergaand onderzoek te doen, bijvoorbeeld door feitenreconstructie aan de hand van een gedetailleerde tijdlijn met gebeurtenissen; meestal voldoet een globaal beeld vanuit kwalitatieve interviews.

Het onderdeel instrumentatie omvat de inventarisatie van de inzet van het instrumentarium. Bij de uitvoerders wordt (via documentatie en interviews) nagegaan welke activiteiten zijn uitgevoerd. Dit wordt naast het programma gelegd waarmee nagegaan wordt wat wel / niet volgens plan is uitgevoerd; interviews zullen duidelijk maken wat de oorzaken zijn van afwijkingen tussen programma en uitvoering.

De beleidsprestaties zullen bekend zijn bij de uitvoerders, en daar doorgaans op eenvoudige wijze te verkrijgen zijn.

. Externe dynamiek

Tijdens de periode van de beleidsuitvoering kunnen zich externe omstandigheden voordoen die op de uitvoering en het resultaat van invloed kunnen zijn, in positieve en negatieve zin. Het is van belang om in de ex post evaluatie na te gaan welke externe dynamieken zijn opgetreden omdat dit van invloed kan zijn op het uitvoeringsproces, op de geleverde beleidsprestaties en op de resultaten. Analyse van de externe dynamiek is onderdeel van de verklaring van het bereikte resultaat (zie verderop).

De kwestie van de externe dynamiek zit in het hart van de vraagstelling naar de specifieke bijdrage van het beleid in de oplossing van een maatschappelijk probleem.

Immers, gemeten resultaten kunnen het gevolg zijn van beleid maar ook van andere ontwikkelingen en factoren. Bij de bepaling van de doeltreffendheid dienen deze zo goed mogelijk uit elkaar te worden gehouden. Het is dan ook raadzaam om op een systematische manier te inventariseren welke drivers van invloed zijn geweest op de ontwikkeling van het probleem, en om deze invloed te specificeren (zoals in welke mate en bij welke doelgroepen).

Uitvoering van de ex post evaluatie

Het opsporen van belangrijke externe invloeden is deels onderwerp van gesprek met uitvoerders maar zeker ook eigen onderzoek van het evaluatieteam. Het bepalen van de mate en aard van de invloed kan om apart onderzoek vragen, ofschoon daar in de praktijk lang niet altijd de ruimte voor is.

. *Doelbereik*

Op dit punt in de evaluatie gaat het om de vraag of het maatschappelijk probleem al dan niet gedeeltelijk of geheel is opgelost. Het gaat hier dus nog niet om de vraag in welke mate het beleid hieraan heeft bijgedragen; immers, ook hetgeen hier aangeduid is als externe dynamiek kan bijdragen aan de probleemoplossing. Het is van belang om beide invalshoeken goed te onderscheiden. De opvatting is mogelijk dat het in het beleid primair gaat om de vraag of een probleem is opgelost. En secundair of dat wel, niet of in enige mate het gevolg is van de beleidsmaatregelen. Het zijn onderscheiden vraagstellingen, die in de ex post evaluatie onderscheiden worden behandeld.

Uitvoering van de ex post evaluatie

Onder het doelbereik worden doelen en resultaten vergeleken. Daarmee wordt vastgesteld in welke mate de doelen bereikt zijn.

. *Doeltreffendheid*

Doeltreffendheid gaat over de mate waarin het beleid heeft bijgedragen aan het doelbereik. Dit staat naast de mate waarin externe dynamiek daaraan heeft bijgedragen.

Globaal gesproken zijn er, zoals eerder beschreven twee hoofdlijnen om causaliteit te onderzoeken:

- De eerste lijn is die van de quasi-experimentele opzet. Daarin wordt de situatie waarin het beleid wordt toegepast vergeleken met de situatie waarin dat niet het geval is. Veelal is dat de situatie voordat het beleid tot uitvoering kwam (vgl. de 'nulmeting'), maar het kan bijvoorbeeld ook gaan om gelijktijdige situaties op andere lokaties (zoals spitsstroken t.o.v. vergelijkbare wegvakken zonder

spitsstrook, of brommers wel of niet op de rijbaan). Het gaat veelal om survey-onderzoek waarbij met behulp van statistische technieken verschillen gemeten worden. Deze verschillen kunnen benoemd worden als de gevolgen van het beleid.

- De tweede lijn is die van de analyse van de beleidstheorie, ofwel de analyse van de verschillende stappen in de implementatieketen. Anders dan de quasi-experimentele analyse is deze benadering vooral erop gericht uit te zoeken wat precies de werkzame bestanddelen in het beleid zijn die tot de gewenste resultaten leiden. En anders dan de quasi-experimentele opzet kan hier gebruik gemaakt worden van elders opgedane inzichten.

De beide methoden kunnen elkaar goed aanvullen. Gezien het feit dat veelal de precieze effecten van het beleid lastig te isoleren zijn, kan een combinatie van methoden zeker leiden tot verduidelijking van de bijdrage van het beleid. In een combinatie van methoden zal de werkzaamheid van het beleid het beste benaderd worden. Anderzijds stijgen de kosten van de evaluatie door inzet van meer methoden. Het doel en de gebruikswaarde van de evaluatie zal uitwijzen welke aanpak de voorkeur verdient.

De quasi-experimentele methode is vergaand gestandaardiseerd. Een vast stramien voor de doeltreffendheidsanalyse op basis van de beleidstheorie is niet te geven. Wel zijn in het bovenstaande enkele bouwstenen genoemd die in de analyse betrokken worden:

- > de doelbereikmeting(en)
- > de uitvoering van het beleidsprogramma
- > de gerealiseerde beleidsprestaties
- > de bereikte tussendoelen
- > de externe dynamiek

Het basisrecept is om met tenminste deze gegevens in de hand de onderdelen van de implementatieketen op juistheid te onderzoeken.

Uitvoering van de ex post evaluatie

De doeltreffendheidsanalyse wordt uitgevoerd in de vorm van desk research; de gegevens zijn reeds in de bovenstaande onderdelen van het evaluatieonderzoek verzameld.

o Doelmatigheid

Doelmatigheid gaat over de vraag hoe het zit met de verhouding kosten – baten, en iets simpeler geformuleerd: had hetzelfde resultaat ook met inzet van minder middelen behaald kunnen worden. Het bereikte resultaat is hierboven reeds duidelijk geworden. Daarnaast moet – voor zover nog niet gebeurd - geïnventariseerd worden wat de inzet geweest is van geld, tijd en andere middelen.

De doelmatigheidsvraag kan zelden met empirisch onderzoek beantwoord worden; immers, de werkelijkheid kan niet meer veranderd worden. Dit betekent dat de doelmatigheidsvraag alleen maar beantwoord kan worden via een denk-exercitie van de evaluator, al dan niet in samenspraak met uitvoerders. In deze denk-

exercitie kan enig houvast gevonden worden in het onderzoek naar de uitvoering en in de analyse van de externe dynamiek. Als gebleken zou zijn dat van een bepaald instrument weinig werkingskracht is uitgegaan, dan kan het zijn dat de inzet hiervan niet doelmatig was. Als gebleken zou zijn dat externe dynamieken veel hebben bijgedragen aan het doelbereik, dan kan dit wellicht leiden naar de conclusie dat hiervan intensiever gebruik gemaakt had kunnen zijn. Voor het overige is het onderzoek naar de efficiency een theoretische analyse van mogelijke andere (goedkopere, minder tijdrovende) inzet van instrumenten en organisatie. Daartoe beschouwt de evaluator instrumenten en organisatievormen, en doordenkt voor zover mogelijk wat de mogelijke gevolgen daarvan geweest zouden kunnen zijn. En of deze mogelijkheden geleid zouden hebben tot vergelijkbare beleidsresultaten.

Voor de goede orde: een dergelijke aanpak van de efficiency-vraag komt in de praktijk van de beleidsevaluatie zelden voor. Doorgaans beperkt de beantwoording van de doelmatigheidsvraag tot een onderzoek naar de souplesse van het uitvoeringsproces.

Uitvoering van de ex post evaluatie

Ten eerste: ten behoeve van de beantwoording van de doelmatigheidsvraag worden gegevens verzameld over de input: kosten en wellicht personele inzet.

Ten tweede: uitkomsten uit het doeltreffendheidsonderzoek kunnen een-op-een omgezet worden in doelmatigheidsconclusies, voorzover er sprake is van beperkte doeltreffendheid.

Ten derde: de evaluator stelt hypothetische redeneringen op over mogelijk doelmatiger aanpakken, en toetst deze bij experts en betrokkenen.

4.1 Evaluatie light

Tot slot: het bovenstaande is het model voor een integrale ex post evaluatie. Niet altijd is er aanleiding of gelegenheid om volgens deze relatief kosten- en arbeidsintensieve aanpak te werk te gaan.

Een light optie kan zijn om de beslissing voor een integrale evaluatie af te laten hangen van de meting van het doelbereik. Wanneer het doelbereik redelijk in de buurt zit van hetgeen beoogd werd, dan kan dat een goede reden zijn om het verklarende deel van de evaluatie achterwege te laten.

Een overweging bij de beslissing om wel of niet een integrale evaluatie uit te laten voeren, kan zijn of het beleidstraject leerzame elementen bevat, leerzaam voor toekomstig beleid. Als dat niet het geval is, is een uitgebreide evaluatie naar de werkzame bestanddelen van het beleid minder zinvol.

Als er indicaties zijn dat het beleidsproces haperend is verlopen, kan dat een reden zijn het proces te onderzoeken. Indien daar geen indicaties voor zijn, is voor een procesanalyse minder aanleiding.

BIJLAGE

Een aantal ex post evaluaties over verkeersveiligheid

Deze bijlage bevat een overzicht van evaluaties van verkeersveiligheidsbeleid in de afgelopen jaren. De selectie is vrij willekeurig: een deel van de rapporten is verstrekt door RWS/DVS, een deel is ontleend aan het rapport Tien jaar Duurzaam Veilig en een deel is gegoogled op 'evaluatie' en 'verkeersveiligheid'. Per rapport worden kort besproken: vraagstelling, methodieken en conclusies, gevolgd door enkele opmerkingen. Deze opmerkingen worden gemaakt vanuit de kernvraag van dit onderzoek: inzicht in doeltreffendheid, doelmatigheid, causaliteit en plausibiliteit van verkeersveiligheidsmaatregelen.

. *W.A.M. Weijermans en I.N.L.G. van Schagen (2009), Tien jaar Duurzaam Veilig. Leidschendam: SWOV.*

In dit rapport – een meta-evaluatie – wordt de balans opgemaakt van het in 1998 gestarte programma Duurzaam Veilig. Achtereenvolgens worden besproken: de achtergronden en visie, de realisatie van (voorgenomen) maatregelen naar typen, effecten op slachtofferaantallen (op basis van ex post evaluaties), en de maatschappelijke kosten-baten verhouding.

. *Evaluatie verkeersveiligheid spitsstroken (Arcadis, 2007)*

Vraagstelling

Spitsstroken zijn ingegeven door de wens op korte termijn knelpunten in de verkeersafwikkeling die zich gedurende een deel van de dag voordoen, op te lossen. Voor spitsstroken gelden randvoorwaarden in de sfeer van de verkeersveiligheid, waartoe ook voorzieningen (vluchthavens, signalering) getroffen worden. Doel van de evaluatie is om na te gaan of aan deze randvoorwaarden wordt voldaan. De evaluatie dient een zo goed mogelijk beeld op te leveren van

- o De verkeersveiligheid op trajecten met spitsstroken (objectief, subjectief)
- o De belangrijkste verklarende factoren van de verkeersveiligheid
- o De mogelijkheden om de verkeersveiligheidspositief te beïnvloeden.

Thema's zijn objectieve verkeersveiligheid, subjectieve verkeersveiligheid, verkeerskundige inrichting en procedures.

Methodiek

Evaluatie op basis van vier typen gegevens: ongevalldata, interviews, focusgroepen, visuele inspecties

Conclusies

Vanwege de verbeterde doorstroming is de verkeersveiligheid na opening toegenomen: sterke afname van het aantal kop-staartbotsingen. De verbeterde doorstroming is het gevolg van de capaciteitsuitbreiding, niet van het fenomeen spitsstrook.

De daling van de aantallen ongelukken op spitsstrooktrajecten loopt min of meer parallel aan de landelijke daling van de aantallen ongelukken. 'Dit duidt erop dat een deel van de afname aan ongevallen en slachtoffers op de spitsstrooktrajecten naar verwachting positief wordt beïnvloed door het landelijke beleid. Met andere woorden, op deze trajecten zou waarschijnlijk ook een daling hebben plaatsgevonden, als er geen spitsstrook zou zijn aangelegd.'
Waar op spitsstroken de intensiteit tot file-niveau is gestegen, wordt het oorspronkelijke effect op de verkeersveiligheid teniet gedaan.

De evaluatie leidt tot tal van verbeterpunten in inrichting en beheer van de spitsstroken.

Commentaar

Op zichzelf valt op deze evaluatie weinig aan te merken.

Maar als het gaat om de causaliteit vallen twee passages op. In de eerste plaats is niet helemaal te begrijpen wat wordt bedoeld met de passage dat vermindering van de aantallen ongelukken is toe te schrijven aan capaciteitsverbreding en niet van het fenomeen spitsstrook. In de onderzochte gevallen vindt capaciteitsuitbreiding plaats in de vorm van een spitsstrook; en anderzijds is een spitsstrook die geen capaciteitsuitbreiding levert, niet goed voorstelbaar.

In de tweede plaats is onduidelijk hoe de – oorzakelijke – relatie ligt tussen de ontwikkeling van de aantallen ongelukken op de spitsstrooktrajecten en de landelijke daling. In de geciteerde passage wordt impliciet een relatie gelegd tussen ontwikkeling van aantallen ongelukken het beleid; de vraag is of die relatie bewezen is. Daarnaast is niet onderbouwd dat ook zonder spitsstrook het aantal ongevallen op deze trajecten zou zijn afgenomen. Om tot zo'n uitspraak te kunnen komen is detailonderzoek nodig over aard en locaties van ongevallen. Hierbij kan naast onderzoek van ongevalgegevens ook observaties op locatie (voor- en nameting) behulpzaam zijn.

. M. Galesloot (2004 – concept), *Evaluatie A5 – een jaar na opening*. Ministerie V&W

Vraagstelling

De vraagstelling voor de evaluatie is: Heeft de aanleg van de A5 tot gevolg gehad dat het beoogde doel bereikt is?

Het beoogde doel is verbetering van de bereikbaarheid van Schiphol en ontlasting van knooppunt Badhoevedorp.

Methodiek

- MTR- en MARE-data
- Visuele tellingen bij aansluitingen en verbindingbogen.
- Tellingen provincie Noord-Holland
- Kentekenenquête

Conclusies

De bereikbaarheid van Schiphol is toegenomen: de file ten zuiden van Schiphol is aanvankelijk verdwenen. Er is daar na verloop van tijd wel 'nieuw' verkeer bijgekomen. Op de A9 is er wel de avondfile, maar deze is naar het noorden opgeschoven en slaat dus minder terug op knooppunt Badhoevedorp. Het knooppunt Badhoevedorp is flink ontlast omdat 20.000 voertuigen per dag zijn opgeschoven naar de A5.

Commentaar

Dit is een duidelijke ex post evaluatie, deugdelijk in opzet, rapportage en conclusies. Het enige gemis is een duidelijk beeld van hetgeen met de A5 beoogd werd; het kan zijn dat er inzake de A5 een duidelijker toekomstbeeld (ofwel effectprognoses) was dan de in de evaluatie in algemene bewoordingen gegeven doelstelling. Zo is de vraag of de aangroei met nieuw verkeer op de A4 zuid van Schiphol voorzien was.

. P. Broeren, M. van Westing, W. Groen, M. van den Bos, Evaluatie proef samenvoegstrook A12 Westervoort, 2007

Vraagstelling

In 2005/2006 vond een proef plaats met een samenvoegstrook: invoegen zonder blokmarkering.

Doel van de evaluatie was het onderzoeken van het rijgedrag.

Methodiek

Literatuuronderzoek; videoanalyses; focusgroepen met weggebruikers.

Conclusies

De samenvoegstrook is een goed alternatief: de verkeersveiligheid lijkt niet te verslechteren, invoegen verloopt soepeler. Effect op doorstroming is hier niet goed te meten vanwege beperkt verkeersaanbod.

Commentaar

Cijfers over ongevallen / slachtoffers ontbreken maar wellicht is daar geen aanleiding toe. De videoregistraties waren wellicht wat aan de dunne kant.

. Evaluatie N50 Hattemerbroek – Kampen Zuid. Rijkswaterstaat Noord-Nederland 2007

Vraagstelling

Object van evaluatie is de herinrichting van een deel van de N50 tot een 2+1 profiel, aangelegd met het oog op de verbetering van de verkeersveiligheid en het rijcomfort (inhaalbaarheid); er is een fysieke rijbaanscheiding aangebracht. De evaluatie is bedoeld om na te gaan of deze doelen zijn verwezenlijkt.

Methoden

- Ongevallengegevens
- Enquête weggebruikers
- Focusgroepen

Conclusies

Er is sprake van 'slechts gedeeltelijke' verbetering van de verkeersveiligheid. De ernst van de ongevallen is afgenomen (geen frontale botsingen / geen dodelijke afloop) maar het aantal lichte ongevallen is toegenomen. Dit wordt toegeschreven aan een VRI op een stroomweg. De verkeersveiligheidsbeleving van weggebruikers is duidelijk toegenomen, evenals het rijcomfort. Met als kanttekeningen het einde van de inhaalstrook en de VRI. De conclusie is dat het 2+1 profiel in principe navolging verdient, met aandachtspunten voor de specifieke inrichting en met de restrictie dat onduidelijk is hoe het profiel functioneert bij hoge verkeersintensiteiten.

Commentaar

De algemene conclusie op grond van de feiten – minder ernstige maar wel meer ongelukken – wordt niet beredeneerd. De enigszins impliciete aanbeveling om het profiel op grotere schaal toe te passen, is wellicht wat voorbarig. Het is denkbaar om eerst na te gaan hoe een verbeterde inrichting van dit wegvak functioneert.

. Evaluatie 80 km/uur-maatregel A13 Overschie. Doorstroming en Verkeersveiligheid. Rijkswaterstaat 2003

Vraagstelling

De 80 km grens is ingevoerd met het oog op geluidsoverlast en luchtkwaliteit. De maatregel heeft ook effecten op doorstroming en verkeersveiligheid. Het doel van de evaluatie is het verwerven van inzicht in effecten inzake doorstroming en verkeersveiligheid, met het oog op toepassing van de maatregel elders.

Methoden

Verkeers- en ongevalsgegevens
- Doorstroming: voor en nametingen
- Interviews (?) weggebruikers

Conclusies

Goede naleving – als gevolg van de strenge handhaving - leidt tot een rustiger verkeersbeeld: minder snelheidsverschillen. De effecten op files zijn wisselend naar locatie en rijrichting. De verkeersveiligheid is verbeterd: minder ongevallen en slachtoffers, vooral op het maatregelvak. Wel wordt opgemerkt dat de effecten op verkeersveiligheid onzeker zijn gezien de korte periode waarin de maatregel van kracht is.

Commentaar

De evaluatie geeft een duidelijk beeld van de ontwikkeling van de verkeersveiligheid en van de invloed van de maatregel. De claim dat naleving (mede) het gevolg is van strenge handhaving is aannemelijk maar wordt op zichzelf niet bewezen.

. L. Bergevoet e.a. (2008), Potentiële deelname nieuw bromfietsexamen. Den Haag: ministerie V&W

Vraagstelling

Vanwege het hoge aantal verkeersslachtoffers onder jeugdige bromfietzers wordt overwogen een praktijkgedeelte toe te voegen aan het bromfietsexamen; daartoe behoort eventueel ook een groepsdiscussie.

In het onderzoek wordt de potentiële markt voor het bromfietsexamen in beeld gebracht.

Methodiek

Online enquête onder 475 leerlingen van 15 – 17 jaar en focusgesprekken.

Conclusies

Er is een nauw verband tussen de deelnamebereidheid en de kosten. De perceptie van ongeveer tweederde van de doelgroep is dat het praktijkexamen niet nuttig is; eenderde vindt het een goede ontwikkeling vanwege verwachte afname van het aantal ongelukken..

Commentaar

Dit is een ex ante evaluatie en past op zichzelf niet helemaal in dit rapport. Het onderzoek biedt een beeld van de omvang van de te verwachten deelname – met marges - aan het examen, maar gaat verder niet in op de verkeersveiligheidskundige betekenis hiervan. Maar dat was ook niet de bedoeling; de gegevens waren bedoeld voor de planning van het CBR.

. W. de Haan (2008), *Evaluatie rijwizertest.nl*. Delft: RWS/DVS

Vraagstelling

De rijwizertest is voortgekomen uit de beleidswens om automobilisten enige vorm van scholing te geven. Dit geschiedt in de vorm van voorlichting, een zelftest en een daarop gebaseerd advies.

De evaluatie bevat een analyse van het invulgedrag, meningen van respondenten over de test en de resultaten van de online publiciteit.

Methodiek

Gegevens invulling test, vragen over de testlijst (pm online werving ?)

Conclusies

In de onderzoekperiode zijn 8777 mensen aan de test begonnen waarvan er 3819 de test geheel hebben doorlopen. Ruim een derde deel van de respondenten vindt dat de vragenlijst heeft aangezet tot nadenken over het eigen rijgedrag. Tweederde is het eens met de eigen scores, minder bij jongeren. De helft zal advies opvolgen.

Commentaar

In het rapport wordt niet duidelijk gemaakt wat de functie van het onderzoek is. Los van informatie daarover valt te constateren dat het rapport weinig inzicht biedt in de betekenis van de rijwizertest op het rijgedrag: de gedragseffecten op deelnemers, en de betekenis van het aantal invullers.

. J. Tijsen en W. Vermeulen (2009), *Nameting Rijwizertest. Evaluatie van de effecten van de Rijwizertest op langere termijn*. Delft: RWS/DVS.

Vraagstelling

In de nameting wordt getoetst of de Rijwizertest bij respondenten heeft geleid tot een groter bewustzijn van verkeersveiligheid en eventueel aanpassingen in het rijgedrag. Kernbegrippen in de nameting zijn retentie, attitude en gepercipieerd gedrag.

Methodiek

Online enquête deelnemers

Conclusies

In de periode 100308 – 311008 hebben 1204 automobilisten van de Rijwizertest gebruik gemaakt.

Weerstand wordt opgeroepen wanneer de test uitmondt in een toegespitst advies op basis van een lage score. De resultaten doen vermoeden dat het stellen van vragen in de test meer effect op het gedrag heeft dan het geven van een gericht advies.

Een duidelijke meerderheid van de respondenten geeft aan dat het gelukt is het rijgedrag te veranderen, het gaat om aspecten als bewustzijn van onveilige verkeerssituaties, het vermijden van onveilige verkeerssituaties, minder verkeersovertredingen maken en herkenning van onveilige situaties.

De Rijwizertest biedt mogelijkheden om automobilisten tot ander, bewuster en beter gedrag aan te zetten.

Commentaar

De nameting is een vervolg op de eerste meting waarin veel dieper wordt ingegaan op de invloeden op denken en gedrag van de automobilist, op het niveau van het gepercipieerde gedrag.

Het feitelijke gedrag is geen onderwerp van onderzoek (vereist uiteraard een geheel ander design) en de (geringe) omvang van het gebruik van de Rijwizertest wordt niet geduid.

. *D.A.M. Twisk (1998), Evaluatie bromfietscertificaat. Een onderzoek naar de korte termijneffecten van de invoering van het theoriecertificaat. Leidschendam: SWOV*

Vraagstelling

Met het bromfiets-theoriecertificaat, ingevoerd in 1996, wordt beoogd dat met name jonge brom- en snorfietzers beter voorbereid aan het verkeer deelnemen. De evaluatie had als doel om zicht te krijgen op bedoelde en onbedoelde (neven)effecten.

Methodiek

Documenten, interviews uitvoerders, secundaire analyses bestaande gegevens, survey onder scholieren

Conclusies

Er wordt een sterke en substantiële verbetering van kennis van en inzicht in het verkeer waargenomen.

De slaagkans hangt vooral samen met intelligentie en opleiding. De slaagkans ligt het laagst bij de categorie meisjes in het VBO.

Het certificaatbezit is hoog; er wordt weinig zonder certificaat gereden.

Er is geen effect gemeten op het bromfietsbezit.

Er is weinig animo voor een praktijkcursus als vervolg op het theoriecertificaat.

Commentaar

De gebruikswaarde van de evaluatie is niet precies omschreven. Los daarvan is het algemene beeld dat deze evaluatie weinig inzicht geeft in de betekenis van het bromfiets-theoriecertificaat voor verkeersgedrag en verkeersveiligheid. De enige indicatie is dat gemiddeld het niveau van kennis en inzicht duidelijk is toegenomen.

. *Educatieve Maatregel Gedrag en verkeer. Procesevaluatie – korte termijn. Evaluatie instroom. DHV / Min V&W –DVS, 2009.*

Vraagstelling

De doelgroep van de EMG zijn rijbewijsbezitters die in het verkeer zijn opgevangen door risicovol en gevaarzettend rijgedrag. Het doel van de EMG is hen te motiveren tot veilig en sociaal verkeersgedrag en het faciliteren tot veranderingen in hun verkeersgedrag.

De procesevaluatie heeft tot doel inzicht te krijgen in het uitvoeringsproces van de EMG.

Methodiek

- Interviews met vertegenwoordigers van de organisaties die in het proces handelend optreden.
- Kwantitatieve gegevens CBR over de verschillende stappen in het proces (niet volledig; politiegegevens ontbreken nog).

Conclusies

De uitvoering verschilt per politieregio. Interpretatie van de feiten kan van persoon tot persoon (politieagent) behoorlijk verschillen. Aspecten waarop gewogen wordt of iemand doorgestuurd moet worden naar EMG zijn of men de straf passend vindt ten opzichte van de ernst van het rijgedrag, of er sprake is van dubbel straffen en of men het opmaken van de mededeling teveel gedoe vindt.

Het voorziene niveau van 1500 mededelingen per jaar wordt gehaald. Bijna de helft is 18 – 24 jaar; daarnaast is er een grote groep ervaren bestuurders met een hoog kilometrage, het CBR werkt zoals bedoeld.

Commentaar

Deze evaluatie is grotendeels een procesevaluatie; enkele resultaatgegevens komen wel in beeld zoals het aantal mededelingen dat overeenkomt met hetgeen voorzien is.

De waarnemingen over de diversiteit in het uitvoeringsproces binnen de politieregio's worden niet van een duidelijke kwalificatie voorzien.

. *XTNT (2006), Procesevaluatie Leerlijnen. Utrecht.*

Vraagstelling

Onder de noemer Leerlijnen zijn vijf lesdossiers ontwikkeld, zodanig ontworpen dat leerlingen gedwongen worden over hun eigen gedrag na te denken. Het doel van Leerlijnen is dat middelbare scholen gedurende minimaal een leerjaar aandacht besteden aan verkeerseducatie en de Leerlijnen opnemen in hun jaarplanning. De procesevaluatie heeft tot doel om inzicht te krijgen in het onderwijspakket en de werkwijze van CNF bij de lessen over verkeersveiligheid en de waardering hiervoor door leerlingen en docenten.

Methodiek

Groepsgesprekken met leerlingen en interviews met docenten.

Conclusies

Leerlingen beoordelen een projectdag over verkeersveiligheid als leuk – de meeste herinneren zich de dag een half jaar later nog – , niet als leerzaam want ze doen geen nieuwe kennis op. Het lesdossier wordt als saai beoordeeld. Docenten zijn positief over de combinatie van het ophalen en toepassen van kennis goed, plus de toepassing van vaardigheden. De projectdag is een eenmalige activiteit, zonder vervolg.

De procesevaluatie geeft geen inzicht in het leerrendement.

Commentaar

De aanbevelingen zijn niet geadresseerd. In het bijzonder de aanbevelingen over de verankering in de onderwijsprogramma's komen nogal gratis over.

. *Thuiskomen in 2007. Een overzicht van de monitoringsresultaten van de verkeersveiligheidscampagnes in 2003 – 2007. Rijkswaterstaat 2007.*

Vraagstelling

Doelstelling van het MPCV is 'versterking van de samenhang, consistentie en effectiviteit van de campagnes over veilig verkeersgedrag, teneinde een bijdrage te leveren aan permanente verbetering van de verkeersveiligheid in termen van slachtoffereductie. In het MPCV wordt een integrale inzet gepleegd van

gedragsbeïnvloedende instrumenten (voorlichting, handhaving) en is er continuïteit over langere periode op vijf thema's.

Vier campagnes worden in de rapportage betrokken: autogordels, Bob, Rij met je hart en fietsverlichting.

Methodiek

- Met publieksenquêtes en gedragswaarnemingen worden effecten van campagne-inspanningen gemeten.
- Periodieke onderzoeken naar houding en gedrag van weggebruikers.

Er wordt gewezen op het causaliteitsprobleem: 'Vele onzekerheden belemmeren het om inzichtelijk te maken welk deel van een geconstateerd effect (...) is toe te schrijven aan voorlichting, aan handhaving of aan andere factoren (...).'

Conclusies

Het campagnebereik is hoog, de herinnering wisselend en de boodschap komt bij tweederde van de doelgroep over. De risico-perceptie is vrij hoog, de subjectieve pakkans is gestegen maar wordt niet als hoog gepercipieerd. De houding is stabiel respectievelijk toegenomen (in het bijzonder bij het gordelgebruik). Het zelfgerapporteerde gedrag is tijdens de campagneperiode verbeterd. Geobserveerd gedrag wijst uit dat het gordelgebruik is gedaald, het alcoholgebruik in de weekendnachten gedaald en het gebruik van fietsverlichting gestegen. Indicatieve berekeningen van de ontwikkeling van slachtofferaantallen laten bij de drie bovenstaande campagnes laten dalingen zien van de aantallen doden (150) en ziekenhuisgewonden (2500), met de kanttekening dat er geen zekerheid is over wat aan de campagnes kan worden toegewezen. Een en ander leidt tot de conclusie dat de campagneaanpak een gunstige invloed heeft op de verkeersveiligheid, in termen van gedragsverandering en in termen van aantallen slachtoffers. En daarmee is het campagnedoel bereikt.

Commentaar

Een opmerking terzijde: in het monitoringsmodel staat onder 'output (ex post)' 'Opgave van de ingezette materiële middelen in kwantitatieve zin'. Zowel de koppeling tussen output en ex post als de beschrijving daarvan zijn ongebruikelijk. Wat opvalt is dat in de analyses een rechtstreeks verband gelegd wordt tussen het programma (voorlichting en handhaving) enerzijds en gedragsaspecten en slachtofferaantallen anderzijds²⁶. Ofschoon deze relatie plausibel is, past wel de methodologische kanttekening dat geen (zichtbare) poging gedaan is te zoeken naar andere factoren die hierop van invloed hebben kunnen zijn.

. M. de Jager (2008), *Onderzoek uitvoering maatregel beginnend bestuurder*. Den Haag: B&A.

Vraagstelling

Doel van de Maatregel beginnend bestuurder (2002) is om het aantal ongelukken door beginnend bestuurders te doen afnemen. De uitvoering van de regeling ligt in handen van politie, lokaal OM, BVOM en CBR.

Doel van de evaluatie was om inzicht te krijgen in de wijze waarop de door BVOM opgestelde werkinstructie wordt toegepast.

²⁶ Met de reserve dat ik geen inzicht heb in de indicatieve berekeningen.

Methodiek

- Analyselijijn: Weten – willen – kunnen – doen
- Casestudies (interviews korpsen en arrondissementen)
- Telefonische interviews
- Panel leidinggevenden

Conclusies

Politie en lokaal OM zijn bekend met de maatregel, wel is er enige onduidelijkheid over de uitvoering van de werkinstructie. De verkeershandhavingsteam zijn veel beter op de hoogte dan de wijk- en noodhulpteams. De maatregel wordt door de uitvoerende organisaties als zeer zinvol ervaren, maar de prioriteitstelling wisselt per korps. Medewerkers van arrondissementen hebben meer dan eens onvoldoende tijd om de werkinstructie te volgen. Wijkagenten worden onvoldoende geattendeerd op de toepassing van de maatregel. In de praktijk ontbreekt het pv vaak de vermelding beginnend bestuurder.

Commentaar

Het betreft de evaluatie van werkprocessen. De evaluatie geeft een goed beeld van de uitvoering en van de knelpunten daarin.

. *W.P. Vlakveld en H.L. Stipdonk (2009), Eerste verkenning naar de effectiviteit van het beginnersrijbewijs in Nederland. Leidschendam: SWOV.*

Vraagstelling

Het beginnersrijbewijs annex puntensysteem geldt sinds 2002.

In de evaluatie is nagegaan of na de invoering van het beginnersrijbewijs het aantal ernstige ongevallen relatief sterk is afgenomen bij de houders van een beginnersrijbewijs.

Methodiek

- Ongevallengegevens; kansberekeningen

Conclusies

Het aantal ernstige ongevallen bij jonge beginners is niet sterker afgenomen dan bij oudere rijders. Het beginnersrijbewijs heeft dus waarschijnlijk geen gunstig effect gehad op de ongevallenbetrokkenheid. Het incasseren van punten heeft, op basis van kwantitatieve gegevens, geen afschrikwekkende werking leidend tot ander verkeersgedrag.

. *C.C. Schoon, e.a. (2008), De toedracht van dodehoekongevallen en maatregelen voor de korte en lange termijn. Leidschendam: SWOV*

Vraagstelling

Ondanks maatregelen (dodehoek spiegel) neemt het aantal slachtoffers niet af. Doel van het onderzoek is om de toedracht van dit type ongevallen te achterhalen.

Methodiek

Ongevalanalyses, processen verbaal, enquêtes overlevenden.

Conclusies

Drie toedrachten: (hoge vrachtauto's van voor 2007 bieden geen goed zicht op fietsers rechts; chauffeurs maken niet goed gebruik van de spiegels rechts; voor

fietsers geldt geen gedragscode. Daarnaast is er het type ongeval van een haakse kruising waarbij de chauffeur een van rechts komende fietser over het hoofd ziet (voorrangskruisingen, rotondes). Ook wordt een lange termijn perspectief geschetst: scheiding van vrachtverkeer en fietsers.

Commentaar

Dit is geen beleidsevaluatie, maar wel een voorbeeld van een diagnostisch onderzoek naar de oorzaak van ongevallen. In lijn met de benadering in de realistische evaluatie wordt in detail geanalyseerd hoe onderlukken plaatsvinden. Terzijde: inzake de gedragscode geldt de vraag welke verwachtingen hieraan verbonden zijn; de werking van de gedragscode wordt niet onderbouwd.

. *SWOV-Factsheet Rijopleiding in Stappen (RIS). 2008*

Vraagstelling

De RIS is een rijopleiding gestructureerd in vier modules met duidelijk leerdoelen. De vraag die aan het factsheet ten grondslag ligt, is wat dit type opleiding bijdraagt aan de verkeersveiligheid.

Methodiek

Deze is te meten langs de weg van de opleidingsinhoud, de slagingspercentages, de ongevalsbetrokkenheid en veilig gedrag. Elke van deze vier heeft bezwaren en beperkingen.

Conclusies

De inhoud van de opleiding voldoet aan de meeste van de BASIC-criteria. De slagingspercentages liggen hoog. Rijproeven wijzen uit dat RIS-opgeleiden – licht – beter rijden dan niet-RIS-opgeleiden. Uit de gegevens kan niet de conclusie getrokken worden dat de RIS-opleiding leidt tot veiliger verkeersgedrag.

Commentaar

Dit onderzoek laat zien hoe complex het bij verkeersveiligheid is om de relatie te leggen tussen een maatregel en de uiteindelijke effecten.

. *Thuiskomen in 2006. Een overzicht van de monitoringsresultaten van de verkeersveiligheidscampagnes in 2003-2006. AVV 2007*

Vraagstelling

Onderdeel van deze rapportage is de evaluatie van de campagne gordels, gericht op verbetering van het gordelgebruik voor- en achterin.

Methoden

Surveys, gegevens bekeuringen

Conclusies

Het aantal bekeuringen steeg van 2002 tot 2004 en daalde vervolgens in 2005 en vooral in 2006.

Het bereik van de campagne is hoog (90 procent), de geholpen boodschapoverdracht ligt op 60 a 70 procent. Risicoperceptie en subjectieve pakkans zijn in 2004 – 2006 stabiel. Het waargenomen gedrag is over de

onderzoekperiode gestegen (met schommelingen). Dit alles wordt toegeschreven aan '... hoogstwaarschijnlijk (...) de gecombineerde inzet van voorlichting en handhaving'.

Commentaar

Op zichzelf is het niet onwaarschijnlijk dat in combinatie voorlichting en handhaving hebben bijgedragen aan verbetering van het gordelgebruik. Maar in het onderzoek blijkt uit niets op welke manier daar vanuit de handhaving specifiek aan bijgedragen is.

Daarnaast maakt dit onderzoek duidelijk hoe lastig macro-gegevens te interpreteren zijn: er zijn geen zeer duidelijke verbanden tussen het verloop over de vier-jaars periode op de verschillende indicatoren zoals kennis over de campagne, houding, risicoperceptie, subjectieve pakkans, zelfgerapporteerd en feitelijk gedrag en bekeuringen.

