

Benchmarking van verkeersveiligheid: uitwerking van een methode voor gemeenten

Charlotte Bax

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid¹

Letty Aarts

Stichting Wetenschappelijk Onderzoek Verkeersveiligheid²

Bij decentrale overheden zoals gemeenten, bestaat een groeiende behoefte aan methoden om beleid te kunnen evalueren en verantwoorden en effectieve maatregelen kostenefficiënt uit te voeren. Benchmarking is een van die methoden: vergelijken van de eigen prestaties met die van vergelijkbare anderen en leren en implementeren van succesvolle werkwijzen die uit die vergelijking naar voren komen. Benchmarking wordt bij gemeenten op allerlei terreinen toegepast, maar nog niet of nauwelijks op het terrein van de verkeersveiligheid. Wel is in het verkeersveiligheidsonderzoek aandacht besteed aan benchmarking, maar dan vooral gericht op het vergelijken van landen. In dit artikel wordt benchmarking, de invulling en uitgangspunten van deze methode volgens de bestuurskundige literatuur vergeleken met die in de verkeersveiligheidsliteratuur. De belangrijkste conclusie is dat de vooral op het proces met doelgroep gerichte benadering vanuit de bestuurskunde een mooie synthese kan vormen met de meer op de technische op evidentie gerichte uitwerking in de verkeersveiligheidsliteratuur. Uitgaande van deze synthese wordt tot een stappenplan voor een gemeentelijke benchmark van verkeersveiligheid gekomen, alsmede indicatoren die daarbij gebruikt kunnen worden. Daarbij wordt niet alleen naar beleidsuitkomsten in termen van slachtoffers gekeken, maar ook naar omstandigheden, beleidsprocessen en maatregelen die door gemeenten worden getroffen.

Trefwoorden: benchmarking, gemeenten, leren, methode, verantwoording, verkeersveiligheid.

¹ Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, E: charlotte.bax@swov.nl

² Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, E: letty.aarts@swov.nl

1. Inleiding

Economische crisis, onder druk staande budgetten en een toenemende decentralisatie zijn recent belangrijke elementen van de beleidscontext. Deze context zorgt er voor dat de afgelopen jaren evaluatie en verantwoording van beleidskeuzen en inspanningen om effectiever en efficiënter beleid te plegen steeds belangrijker zijn geworden voor decentrale overheden (zie bijvoorbeeld Groot et al., 2006; Pommer et al., 2013; Ten Tije et al., 2010; Boogers et al., 2008). Ook op het terrein van de verkeersveiligheid is dat het geval.

De evaluatie- en verantwoordingsdruk geldt niet alleen binnen de overheid, bijvoorbeeld verantwoording van ambtenaren naar wethouder of gedeputeerde, of van wethouder richting gemeenteraad, maar ook naar buiten toe. Mondige burgers en pers (van den Brink, 2002) vragen verantwoording van overheden over beleidskeuzen, inzet van budgetten en beleidsprestaties en de snelle opkomst van lokale rekenkamers laat een toenemende interesse in evaluatie zien. Decentrale overheden worden bijvoorbeeld ter verantwoording geroepen door burgers over verkeersongevallen en over locaties die in de ogen van burgers gevaarlijk zijn³ en lokale rekenkamers lichten het gemeentelijk verkeersveiligheidsbeleid door (zie bijvoorbeeld Rekenkamercommissie Gemeente Den Haag, 2006).

Ook de druk om beleid efficiënter en effectiever te maken vloeit voort uit de genoemde beleidscontext. Als maatregelen goedkoper kunnen worden uitgevoerd en als er geleerd kan worden van anderen die goede ervaringen hebben opgedaan met een bepaalde aanpak, dan kan dit uitermate interessant zijn voor andere decentrale overheden. Het voorkomt dat overheden met identieke vraagstukken ieder voor zich het wiel opnieuw uit gaan vinden, en daaraan meer geld en tijd kwijt zijn dan noodzakelijk is. En als uit evaluaties blijkt dat effectieve maatregelen tegen lagere kosten kunnen worden uitgevoerd, dan is dat waardevolle informatie die overheden kunnen meenemen in hun eigen beleid.

Bij decentrale overheden zoals gemeenten en provincies, bestaat hierdoor behoefte aan methoden om hun beleid zowel intern als extern te kunnen verantwoorden en effectieve maatregelen kostenefficiënt uit te voeren (Groot et al., 2006; Pommer et al., 2013; Ten Tije et al., 2010). Hiervoor kunnen zowel ex-post en ex-ante evaluaties en maatschappelijke kosten-batenanalyses worden gebruikt (zie bijvoorbeeld Annema et al., 2012). Een van de methoden die hiertoe mogelijkheden biedt is 'benchmarking', een methode waarbij beleidsprestaties van een overheid worden geijkt aan die van andere, vergelijkbare overheden en te leren van succesvolle werkwijzen die daaruit naar voren komen: (zie bijvoorbeeld Groot et al., 2006; Pommer et al., 2013; Ten Tije et al., 2010).

Decentrale overheden passen benchmarks al op diverse beleidsterreinen toe, zoals:

- openbare verlichting (IGOV);
- riolering (RIONED);
- afvalscheiding (Cyclus Management en SenterNovem);
- parkeren (Platform Parkeren);
- benchmark publiekszaken (KING via www.waarstaatjegemeente.nl).

Een voorbeeld van een terrein waar benchmarking door gemeenten nog niet op grote schaal plaatsvindt, is verkeersveiligheid. In het verkeersveiligheidsveld bestaat benchmarking wel, maar is het tot nu toe vooral het domein van wetenschappers: publicaties over benchmarking zijn over het algemeen vrij technisch van aard en gericht op een (vaak internationaal) wetenschappelijk publiek (zie bijvoorbeeld Bax et al., 2012; Breen, 2000; Eksler, 2010; Hermans et

³ Zo biedt de Fietsersbond bijvoorbeeld op haar site tips om gemeenten aansprakelijk te stellen na een ongeval als gevolg van gebrekkige infrastructuur (<http://www.fietsersbond.nl/de-feiten/wetten-en-recht/aansprakelijkheid-gemeente#.VRl1ruHttHw>)

al., 2009a; 2009b; Wegman & Oppe, 2010). Een breed gedragen verkeersveiligheidsbenchmark voor beleidsmakers, analoog aan de eisen zoals bijvoorbeeld het Kwaliteitsinstituut Nederlandse Gemeenten (KING) aan benchmarks stelt, bestaat nog niet. Deze kwaliteitseisen houden onder meer in dat samen met gemeenten naar gedegen verklaringen wordt gezocht waarom de een minder presteert dan vergelijkbare anderen, in een stimulerende leeromgeving, waarbij de gemeenten bovendien deelnemen vanuit bestuurlijk commitment (zie bijvoorbeeld KING, 2011).

Dit artikel gebruikt verkeersveiligheid als voorbeeld hoe benchmarking op het terrein van verkeer en vervoer kan worden toegepast in gemeenten en pleit dan ook voor het toepassen van benchmarking op het terrein van de verkeersveiligheid om beleidsmakers te ondersteunen in de uitvoering van hun werk. Daartoe wordt benchmarking eerst nader belicht vanuit zowel een bestuurskundig perspectief als vanuit het technische vakspecifieke perspectief van verkeersveiligheidsonderzoek. Ook staat dit artikel stil bij de vraag wat benchmarking oplevert. Op basis van de sterke en zwakkere punten van de verschillende benaderingen, komt dit artikel tot een synthese die bruikbare handvatten kan bieden voor benchmarking in het algemeen en voor benchmarking van verkeersveiligheid bij decentrale overheden in het bijzonder. Tot slot schetst dit artikel, als voorbeeld van de uitwerking van een benchmark, de contouren voor een verkeersveiligheidsbenchmark voor decentrale overheden, met name gericht op gemeenten. Dit beoogt uiteindelijk bij te dragen aan het pallet aan methoden om tot een veiliger verkeer te komen.

2. De basis van benchmarking: definities en effectiviteit

2.1 Definities van benchmarking

Definities van benchmarking zijn in allerlei soorten te vinden. Het Engelse woord 'benchmark' is letterlijk te vertalen als: criterium, referentiemaat, referentiepunt of ijkpunt.

In bestuurskundig onderzoek is benchmarking een gebruikelijke methode om beleid van twee actoren, bijvoorbeeld gemeenten, met elkaar te vergelijken. Bestuurskundige literatuur omschrijft benchmarking dan ook meestal als een vergelijkende methode die leidt tot verbetering van de eigen organisatie of het eigen beleid. Zo definieert Camp (1989) benchmarking als "systematisch onderzoek naar de prestaties en de onderliggende processen en methoden van een of meer leidende referentieorganisaties op een bepaald gebied, en de vergelijking van de eigen prestaties en werkmethode met deze goede voorbeelden, met het doel om de eigen prestaties te plaatsen en te verbeteren". Ook anderen definiëren benchmarking op soortgelijke wijze (zie bijvoorbeeld Kozak, 2004; Moriarty & Smallman, 2009). Alle definities hebben een aantal begrippen gemeen. Ze spreken allemaal van:

- de eigen prestaties vergelijken met goede praktijkvoorbeelden;
- leren van deze goede praktijkvoorbeelden; en
- verbeteringen doorvoeren in de eigen praktijk.

Korsten et al. (2013) noemen dat kortweg "spiegelen, spieken en sprankelen".

Binnen de verkeersveiligheidsliteratuur bestaat ook een definitie van benchmarking. Deze wijkt op enkele cruciale punten af van de definities uit de bestuurskunde. In het internationale verkeersveiligheidsveld blijkt het begrip 'benchmarking' vooral gebruikt te worden om een *vergelijkend onderzoek* aan te duiden. Zo definiëren Wegman & Oppe (2010) benchmarking als "een proces waarin landen of regio's verschillende aspecten van hun prestaties evalueren in relatie tot vergelijkbare anderen die tot 'de beste van de groep' behoren". Opvallend aan deze definitie is dat de nadruk wordt gelegd op het vergelijkingsaspect van benchmarken, maar dat niet wordt gesproken over leren en verbeteren, het aspect van benchmarking waar in de bestuurskundige literatuur juist nadruk op ligt.

Soms wordt het begrip 'benchmarken' ook gebruikt om de prestaties van een organisatie te vergelijken met een externe norm, een (beleids)doelstelling of taakstelling. Uitkomsten van beleid kunnen bijvoorbeeld vergeleken worden met de landelijke taakstellingen, of met een ander absoluut criterium als 'nul doden'. Zolang benchmarking plaatsvindt zonder dat de gebenchmarkte organisaties daaraan een actieve bijdrage leveren, ervan leren en leerpunten implementeren, spreekt de bestuurskundige literatuur echter niet van het begrip 'benchmarking' maar van *prestatiemeting* (Korsten, 2006). Een of meer prestatiemetingen zijn wel een essentieel onderdeel van benchmarking.

2.2 Werkt benchmarking?

Diverse auteurs (Dattakumar & Jagadeesh, 2003; Moriarty & Smallman, 2009; Yasin, 2002) concluderen dat de literatuur over benchmarking 'overwhelmingly pragmatic' is. Er is dan ook weinig onderzoek te vinden naar de effectiviteit van benchmarking. Eén van de weinige uitzonderingen hierop is onderzoek uitgevoerd in Nederland onder 24 waterschappen (Van Helden & Tillema, 2005). De onderzoekers van deze studie stellen dat in hun casussen 18 van de 24 waterschappen effectiever zijn gaan werken op het gebied van afvalwaterzuivering na toepassing van een benchmark. Het is echter niet duidelijk of deze effectiviteitsslag daadwerkelijk te danken was aan toepassing van een benchmark, of aan de min of meer tegelijkertijd vanuit de overheid opgelegde strengere effectiviteitseisen. Over een effect op meer efficiëntie zijn geen duidelijke conclusies te trekken.

Meer in het algemeen zijn er wel signalen dat organisaties die benchmarking toepassen over het algemeen positief zijn over deze methode. Zo laat onderzoek van Coopers & Lybrand (1994, in Francis & Holloway, 2007) in een grootschalige enquête zien dat 75% van de geënquêteerde grote bedrijven hun benchmarkproject als positief beoordeelden. Een ander voorbeeld komt van de in Nederland uitgevoerde rioolbenchmark in het kader van het Bestuursakkoord Waterketen uit 2007, om de doeltreffendheid en doelmatigheid van het (water)beleid van gemeenten vast te kunnen stellen. In een actualisering van het akkoord in 2011 is vastgesteld dat deze benchmarks nuttig zijn gebleken (zie www.benchmarkrioleringszorg.nl) en daarom nu herhaaldelijk plaatsvinden. In onderzoek van de TU Twente naar benchmarking van (digitale) dienstverlening bij gemeenten (Ten Tije et al., 2010) bleek dat gemeenten voordelen van benchmarking zien. Wel blijken ze zich ook bewust van mogelijke ongewenste neveneffecten zoals 'window dressing' en feiten en cijfers laten prevaleren boven echt leren van elkaar. Gemeenten zouden vooral graag melding maken van goede prestaties en successen en niet of nauwelijks van hun minder fortuinlijke activiteiten.

Wat ook opvalt is dat bij benchmarking vaak niet expliciet wordt gekeken naar *verklaringen* voor de geconstateerde verschillen tussen organisaties. Impliciet wordt ervan uitgegaan dat de oorzaak van verschillen tussen beter en slechter presterende organisaties ligt in een andere bedrijfsvoering en dat alles van die bedrijfsvoering toe te schrijven is aan de geconstateerde verschillen in succes. Om tot concrete verbeterpunten te komen, kan het echter helpen op zoek te gaan naar verklaringen voor de verschillen, zodat de verbeterpunten specifiek kunnen worden toegesneden op de verschillende organisaties.

3. Benchmarking vanuit bestuurskundige invalshoek

Benchmarking is een begrip dat in eerste instantie vooral in een bedrijfskundige en vervolgens bestuurskundige context is bestudeerd en uitgewerkt. Benchmarking als concept en instrument is ontwikkeld door bedrijven in Japan. Het bedrijf Xerox maakte het begrip begin jaren tachtig ook buiten Japan bekend (zie Huggins, 2010; Noordhoek, 2013). Benchmarking heeft zich aanvankelijk verder ontwikkeld binnen het bedrijfsleven, maar is later, onder invloed van de

New Public Management-beweging (de filosofie uit de jaren '80 om het private managementmodel toe te passen op de publieke sector), ook bij overheden in zwang geraakt (Francis & Holloway, 2007; Korsten, 2006; Yasin, 2002). Twee belangrijke redenen voor benchmarking zijn inzicht krijgen in de eigen prestaties en het afleggen van verantwoording aan de samenleving of aan een financier (Korsten, 2006). Vanuit bestuurskundige context werpt dat vragen op zoals:

- Wat voor soorten benchmarks bestaan er?
- Welke eisen worden gesteld aan benchmarking?
- In hoeverre werkt benchmarking?

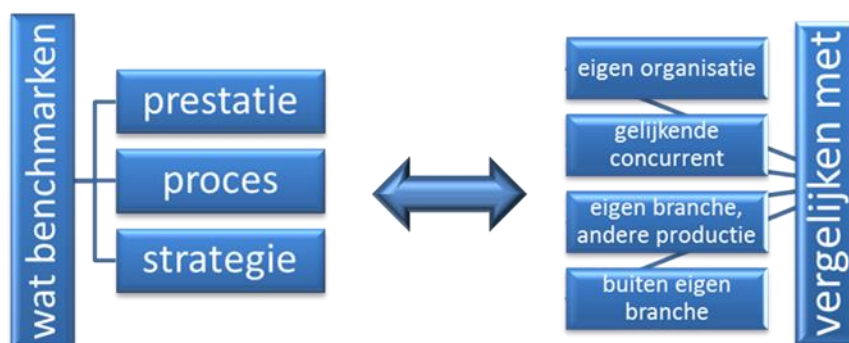
3.1 Soorten benchmarks

In de bestuurskundige literatuur worden drie typen benchmarks onderscheiden, elk met een ander doel (Bhutta & Huq, 1999; Kyrö, 2003; Moriarty & Smallman, 2009):

- Prestatiebenchmarking: dit betreft benchmarken van prestatie-indicatoren, wat wil zeggen dat vergeleken en geleerd wordt van de maatregelen die worden getroffen (output) en de effecten die dat heeft opgeleverd (outcome).
- Procesbenchmarking: dit betreft benchmarken van werkmethoden en werkprocessen (throughput) om de werkprocessen in de eigen organisatie te verbeteren door van anderen te leren.
- Strategische benchmarking: dit betreft benchmarken van de beleidsuitgangspunten van een organisatie (throughput) om de strategische koers van de eigen organisatie te veranderen.

Daarnaast onderscheiden de genoemde auteurs vier mogelijke vergelijkingsgroepen waar een benchmarking zich op kan richten (zie ook *Figuur 1*):

- Interne vergelijking: vergelijking tussen onderdelen van één organisatie;
- Competitieve vergelijking: vergelijking tussen de eigen organisatie en een als 'beste' scorende concurrent/organisatie met dezelfde taken of producten;
- Functionele vergelijking: vergelijking met organisaties uit de eigen branche, maar met andere taken of producten;
- Algemene vergelijking: vergelijking met goede praktijkvoorbeelden buiten de eigen branche, met organisaties die sterk kunnen verschillen van de eigen organisatie.



Figuur 1: Soorten benchmarks.

Niet elk benchmarkdoel (vergelijken, verbeteren of strategie veranderen) verhoudt zich even goed met de vier vergelijkingsmogelijkheden (intern, concurrent, eigen branche, buiten de branche). Zo zal een interne benchmark zich niet lenen voor een heroverweging van de strategische doelen van een organisatie, maar een competitieve benchmark juist wel. Bhutta &

Huq (1999) geven een overzicht in van geschikte en minder geschikte vergelijkingspartners afhankelijk van het benchmarkdoel (Tabel 1).

Tabel 1: Typen benchmarks waarvoor de verschillende vergelijkingsmogelijkheden meer of minder geschikt zijn volgens Bhutta & Huq (1999).

	Eigen organisatie	Gelijkende concurrent	Eigen branche, ander product	Buiten eigen branche
Prestatie	+/-	+	+/-	-
Proces	+/-	-	+	+
Strategie	-	+	-	-

3.2 Eisen aan benchmarking

Verschillende auteurs hebben algemene methoden voor benchmarks beschreven, stappen die daarvoor moeten worden doorlopen, en vereisten waaraan benchmarks zouden moeten voldoen. Bhutta & Huq (1999) vatten dit samen in de volgende vijf stappen (zie Figuur 2):



Figuur 2: Stappen in benchmarkproces, aangepast op basis van Bhutta & Huq (1999).

Hoewel veel eisen aan een goede benchmark contextafhankelijk en specifiek voor het gekozen beleidsveld zijn, kunnen bovenstaande stappen toegepast worden op benchmarks in diverse beleidsterreinen. Het betreft hier enerzijds de technische aspecten van benchmarking (stap 1 en 4) waarbij de interne en externe validiteit in het oog moet worden gehouden: meten we het juiste op de juiste manier en zijn organisaties wel echt vergelijkbaar? Anderzijds betreft het de interactieve kanten van benchmarking (stap 2, 3 en 5), waarbij het uitgangspunt is dat betrokkenheid van organisaties de kans op doorwerking van de resultaten vergroot.

Korsten et al. (2013) voegen aan deze benchmarkstappen zes criteria toe die een goed gebruik van een benchmark volgens hen bevorderen:

- eenvoudig te interpreteren benchmarkresultaten, bijvoorbeeld door gebruik van streefcijfers of andere normen;
- gebruik van een aansprekend en goed praktijkvoorbeeld;
- formuleren van aanbevelingen in het eindrapportage;
- aparte rapportage per organisatie die aan de benchmark deelneemt;
- voorafgaand overleg met de te benchmarken organisaties over de invulling van het onderzoek;
- bestuurlijk draagvlak voor een benchmark.

Daarnaast stelt Korsten (2006; Korsten, et al., 2013) nog een aantal inhoudelijke eisen. Zo pleit hij ervoor dat benchmarking rekenschap blijft geven van de complexiteit van de werkelijke situatie. Complexe prestaties zijn bijvoorbeeld vaak niet volledig in eenvoudige indicatoren te vangen. Echter, te veel nadruk op context en het unieke karakter van organisaties heeft als ultieme consequentie dat geen enkele organisatie zich meer kan vergelijken met een andere. Verder stelt Korsten dat benchmarks ruimte dienen te laten voor creativiteit, ambitie en excellentie. Het streven naar een resultaat gelijk aan het beste praktijkvoorbeeld kan volgens hem organisaties remmen in hun creativiteit en ambitie. Deze stelling lijkt echter vooral ingegeven door sturing op het eindresultaat, waarbij het eindresultaat volgens een vaststaande weg bereikt kan worden (via gevonden succesfactoren). Men zou ook kunnen zeggen dat als het leerproces van de organisatie vooral centraal staat, dit de creativiteit en ambitie niet in de weg hoeft te staan (vergelijk de 'responsieve evaluatie' van Abma, 1996).

Omdat organisatie en inhoud van beleid tot stand komen door een complex van factoren, is interesse om te willen leren en verbeteren niet alleen een zaak van bijvoorbeeld individuele ambtenaren, maar moet er liefst ook bestuurlijk commitment zijn (zie OECD, 2000; Korsten et al., 2013). Volgens het OESO (de Organisatie voor Europese Samenwerking en Ontwikkeling) bevordert dit directe implementatie van leerpunten die uit de benchmark voortvloeien. Meer bestuurlijke aandacht voor benchmarking kan volgens de OESO worden verkregen door:

- de aandacht te vestigen op gebrekkige prestaties of verschil in prestaties (prikkelen);
- de haalbaarheid inzichtelijk te maken waarmee goede praktijkvoorbeelden kunnen worden verkregen (stimuleren);
- de consequenties inzichtelijk te maken van het niet invoeren van goede voorbeelden van anderen (risicobewustzijn kweken).

4. Een voorbeeld uit de verkeer- en vervoerwereld: Benchmarking van verkeersveiligheid

Verkeer en vervoer is een van beleidsdomeinen die in Nederland in sterke mate zijn gedecentraliseerd. Het zou in de lijn der verwachting liggen dat vanwege deze sterke decentralisatie, het aantal benchmarks op dit terrein groot is, maar zoals eerder aangegeven is dat niet het geval. In deze paragraaf inventariseren we, met verkeersveiligheid als voorbeeld, wat er op het gebied van benchmarking op dit moment gebeurt. De decentralisatie betekent op het gebied van verkeersveiligheid dat overheden zoals provincies en gemeenten een belangrijke taak hebben om hun wegennet zo veilig mogelijk in te richten en daarnaast ook andere maatregelen treffen om verkeersdoden en ernstig verkeersgewonden in hun beheergebied te voorkomen. Op landelijk niveau zijn de verkeersveiligheidsdoelstellingen formeel vastgesteld (te weten: maximaal 500 verkeersdoden en 10.600 ernstig verkeersgewonden in 2020; (Ministerie van IenM, 2012)), en de decentralisatie van verkeer- en vervoerbeleid betekent ook dat decentrale overheden naar rato geacht worden daaraan hun bijdrage in te leveren. Veel provincies en gemeenten zitten

daardoor in feite met een zelfde vraagstuk, namelijk: hoe voeren we zo effectief en efficiënt mogelijk verkeersveiligheidsbeleid? Zeker omdat het effect van beleid op bespaarde slachtoffers niet direct zichtbaar is, is dit vraagstuk voor veel overheden niet eenvoudig. Ongevallen zijn immers het gevolg van een combinatie van uiteenlopende factoren zoals mobiliteitskeuzen en risico's in het verkeerssysteem die maar deels door (verkeersveiligheids)beleid beïnvloed worden, deels door andere factoren zoals ontwikkelingen in de markt en behoeften en gedrag van burgers. Een methode zoals benchmarking kan dan helpen van elkaars aanpak te leren en succesfactoren te vinden die beleid effectief en efficiënt maken. In een eerste verkenning bij gemeenten, geeft een aantal van hen aan daarbij ook geïnteresseerd te zijn in hoe zij effectieve maatregelen kunnen treffen tegen geringere kosten (zie Aarts en Bax 2015).

4.1 Soorten benchmarks

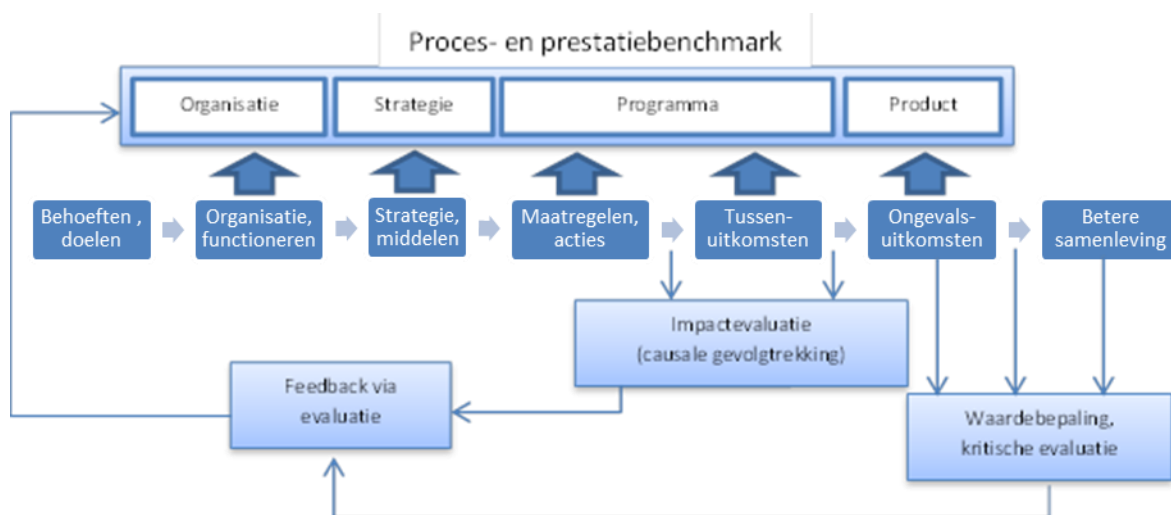
Over benchmarking van verkeersveiligheid is vooral in internationaal en wetenschappelijk verband nagedacht (zie bijvoorbeeld Koornstra et al., 2002; Wegman et al., 2005; Wegman et al., 2008; Gitelman et al., 2010; Hermans et al., 2009b; Bax et al., 2012). Het model dat hierbij veelal centraal staat is de zogeheten 'verkeersveiligheidspiramide' (zie *Figuur 3*).

De prestatie-indicatoren verkeersveiligheid verdienen daarbij een kleine toelichting: dit zijn kenmerken van het verkeerssysteem die een causale relatie hebben met het ontstaan of de ernst van ongevallen. Voorbeelden van deze kenmerken zijn de verkeersveiligheidskwaliteit van het verkeerssysteem (bijvoorbeeld de obstakelvrije afstand in de wegberm in relatie tot de snelheidslimiet of de mate waarin verschillende verkeerssoorten met verschillende snelheden en richtingen van elkaar gescheiden zijn in plaats of tijd), gevaarlijke gedragingen (zoals de mate waarin te hard rijden voorkomt of het aandeel bestuurders dat met alcohol op achter het stuur kruipt) en gebruik van beveiligingsmiddelen (bijvoorbeeld het aandeel bromfietzers dat een helm op de juiste manier draagt).



Figuur 3: Verkeersveiligheidspiramide (Koornstra et al., 2002; LTSA, 2000) met daarin van onder naar boven de elementen van het verkeerssysteem die bijdragen aan of het gevolg zijn van verkeersveiligheid.

Net zoals in de bestuurskundige literatuur (zie §3.1), worden in de verkeersveiligheidsliteratuur-proces-, prestatie en strategische benchmarks onderscheiden (zie Eksler, 2009), waarbij de nadruk ligt op de eerste twee. Eksler stelt vraagtekens bij het nut van strategiebenchmarks van verkeersveiligheid omdat hierbij te weinig in ogenschouw zou worden genomen hoe de implementatie van plannen volgend op een strategie gestalte krijgt. Een strategische benchmark biedt daarmee te weinig zicht op het voor verkeersveiligheidsbeleid essentiële succes in termen van resultaten. In *Figuur 4* zijn de beleidsonderdelen volgens Eksler van proces- en prestatiebenchmarks toegepast op verkeersveiligheid weergegeven. De middelste laag geeft daarbij de opeenvolgende processtappen weer die worden doorlopen, beginnend bij het stellen van doelen of het hebben van bepaalde behoeften (bijvoorbeeld bereiken van meer verkeersveiligheid in een bepaald tijdsbestek), via organisatorische aspecten van het beleid om dit doel te bereiken uiteindelijk leidt tot uitkomsten (veiliger verkeerssysteem en minder ongevallen) die als 'prestaties' gezien kunnen worden. Evaluaties van beleidsimpact op - in dit geval - een veiliger verkeerssysteem en minder ongevallen kunnen zelf ook weer onderdeel zijn van specifieke evaluatie die ertoe kan leiden dat de beleidsinrichting wordt bijgesteld. De mate waarin dergelijke evaluatie en bijstelling plaatsvindt, wordt mede beïnvloed door de maatschappelijke waarde die aan de beleidsuitkomst wordt gehecht.



Figuur 4: Onderdelen van verkeersveiligheidsbeleid als onderdelen van proces- en prestatiebenchmarking volgens Eksler (2009).

Volgens Eksler is er tot nu toe vooral veel aandacht geweest voor benchmarken van resultaten, en ook zijn enkele pogingen ondernomen voor het benchmarken van programma's (zie Wegman et al., 2005) maar niet of nauwelijks voor het benchmarken van organisatie en strategie. Overigens dient hierbij 'benchmarking' veelal opgevat te worden als 'vergelijking'. In veel benchmarkstudies op gebied van verkeersveiligheid worden eerste verklaringen voor verschillen in ongevalsprestaties tussen gebieden (meestal landen) vooral gezocht in verschillen in tussenuitkomsten (prestatie-indicatoren verkeersveiligheid) zoals gevaarlijke gedragingen en kwaliteit van het verkeerssysteem (zie bijvoorbeeld Gitelman et al., 2010; Hermans et al., 2009b; Bax et al., 2012). Presteert een land relatief slecht in termen van verkeersslachtoffers, dan zijn hiervoor meestal verklaringen te vinden in bijvoorbeeld hogere rijnsnelheden, minder gordelgebruik of een lagere kwaliteit van de infrastructuur. In een enkel geval is een poging ondernomen om ook de link te leggen met successen van beleid, bijvoorbeeld door de beleidsinspanningen en de achtergronden van best presterende (of best presterend binnen een vergelijkbare groep) landen te analyseren (zie Bax et al., 2012).

Wat hierbij ook opvalt is dat in de verkeersveiligheidsliteratuur het benchmarken vooral technisch van aard is, waarbij de uitvoering zich concentreert rond de zoektocht van onderzoekers naar indicatoren, criteria en verklaringen.

4.2 Eisen aan benchmarking

In de wetenschappelijke verkeersveiligheidsliteratuur worden voor benchmarking andere – vooral meer technische – stappen onderscheiden dan in de bestuurskundige literatuur. Zo onderscheiden Wegman & Oppe (2010) de volgende stappen:

- identificeren van de belangrijkste componenten van verkeersveiligheidsprestaties;
- identificeren van de vergelijkingsgroep (landen, regio's, inclusief best presterende van de groep);
- construeren van indicatoren om tot een betekenisvolle vergelijking te komen
- vaststellen en verklaren van verschillen in de prestaties;
- vaststellen van na te streven prestaties in de toekomst.

Ook stellen zij dat bij een vergelijking alleen sprake is van een zinnig referentiekader als er *meer dan twee gebieden* vergeleken worden. Een groep van drie wordt dan ook beschouwd als een minimum aantal om goed vast te kunnen stellen wie de best presterende van de groep is en in hoeverre kenmerken van die best-presterende inderdaad afwijken van minder presterende gebieden.

Vervolgens is het ook een vereiste dat gebieden in een benchmarkgroep op een of andere manier goed zijn te vergelijken. Knowles et al. (2010) onderscheiden twee methoden waarop deze vergelijkbaarheid kan worden verkregen:

- samenstellen van vergelijkbare (homogene) gebieden, bijvoorbeeld een groep grote steden of een groep plattelandsgemeenten met maximaal 10.000 inwoners en een sterk agrarisch accent (zie ook Houwing et al., 2012);
- normeren, corrigeren of modelleren van verschillen tussen gebieden om tot een meer vergelijkbaar resultaat te komen (bijvoorbeeld als gebieden van verschillende grootte, inwoneraantal of verschillende relevante kenmerken vergeleken willen worden kan de prestatie per 1.000 inwoners of per ha. grondgebied worden vergeleken).

Homogene gebieden

Wegman & Oppe (2010) zijn expliciet voorstander van vergelijkingen binnen groepen gebieden *die op elkaar lijken*, dit om gemakkelijker beleidsaanbevelingen te kunnen doen. Ook hebben zij in internationaal verband aanwijzingen dat er meer interesse is om te leren van de best presterende als deze soortgelijke eigenschappen heeft als de te benchmarken gebieden.

Houwing et al. (2012) concluderen in onderzoek naar regionale verschillen in verkeersveiligheid dat het vaststellen van vergelijkbare groepen afhankelijk is van de volgende factoren:

- De gehanteerde gebiedsgrenzen. Houwing et al. voerden hun onderzoek uit met gemeenten als eenheid omdat dit Nederland opdeelt in voldoende kleine stukjes homogeen gebied (bijvoorbeeld stedelijk versus landelijk). Zij merkten wel dat gemeentegrenzen voor bepaalde kenmerken te weinig onderscheid boden (bijvoorbeeld kerngemeenten versus gemeenten met verspreid liggende kleine kernen). Uit onderzoek van Eksler (Eksler et al., 2008; Eksler & Lassare, 2008) blijkt dat gebiedskenmerken die van invloed zijn op de verkeersveiligheid, meer verschillen tussen sub-nationale gebieden dan tussen landen (zie ook Eksler, 2009). In deze studies werd gevonden dat de meeste variantie in mortaliteit verklaard kon worden als in het onderzoek groepen gebieden werden vergeleken die op het

laagste disaggregatieniveau (NUTS-3 niveau in dit geval⁴) op elkaar leken. Waarschijnlijk houdt dit verband met het feit dat een gebied over het algemeen homogener kenmerken zal hebben naarmate het kleiner is.

- De gebruikte gegevens. Werd in het onderzoek van Houwing et al. alleen gekeken naar bijvoorbeeld bevolkingsdichtheid, dan kwamen zij tot iets andere homogene groepen gemeenten dan als ook andere kenmerken, zoals grondgebruik, kenmerken van het wegennet en samenstelling van de bevolking in de analyse werden betrokken.
- De gebruikte methode van groeperen. De onderzoekers experimenteerden ook met verschillende methoden om tot een indeling te komen en merkten dat ook dit het resultaat beïnvloedde.

Normering

Er zijn ook benchmarkstudies van verkeersveiligheid die indicatoren normeren of corrigeren voor verschillen tussen de gebieden om te kunnen vergelijken. Zo nemen Eksler et al. (2008) bijvoorbeeld belangrijke lokale invloedsfactoren mee als relevante factor in het onderzoek naar verkeersveiligheidsprestaties tussen gebieden. Knowles et al. (2010) kiezen in hun benchmarkstudie van Noord-Ierland voor modellering nadat zij tot de conclusie zijn gekomen dat geen goede homogene benchmarkgroep is samen te stellen. Door de gemodelleerde relatie tussen factoren te vergelijken met de werkelijke prestatiewaarden van overheden, kan vastgesteld worden of ze onder of boven de gemodelleerde norm zitten.

Een dergelijke modelleringsaanpak lijkt toch vooral geschikt om het effect van kenmerken op de verkeersveiligheid vast te stellen en onderbouwing te leveren van de invloed van succesfactoren. De aanpak is minder geschikt om de rol van organisaties en hun bijdrage aan succesvol verkeersveiligheid in kaart te brengen. Ook levert een modelleringsaanpak over het algemeen relatief complexe en daarmee op het eerste gezicht lastig te interpreteren prestatie-maten op, wat zich daardoor waarschijnlijk minder goed leent voor een analyse- en leerproces met organisaties zoals decentrale overheden.

Kwaliteit en begrijpelijkheid

Niet alleen vanuit de bestuurskundige hoek wordt het belang van aanspreekbare resultaten benadrukt. Ook binnen de verkeersveiligheid zijn hierover aanbevelingen te vinden. Zo benadrukt Eksler (2010) dat, willen eindgebruikers het resultaat van een benchmark accepteren, het belangrijk is dat het gebruikte instrument:

- betrouwbaar en robuust is, maar ook
- eenvoudig en begrijpelijk.

Deze vereisten zijn nogal eens met elkaar in tegenspraak waardoor het nodig is om hierin een evenwicht te vinden.

5. Lessen uit de bestuurskundige en verkeersveiligheidsliteratuur

5.1 Analyse van verschillen

Bovenstaande beschouwing van benchmarking in de bestuurskundige en verkeersveiligheidsliteratuur laat zien dat de bestaande benchmarks verschillende accenten hebben. Zo blijken de 'benchmarks' op het gebied van verkeersveiligheid vooral gericht op een technisch goed uitgevoerde en wetenschappelijk te verantwoorden vergelijking. Dit zou als een

⁴ NUTS staat voor Nomenclature des Unités Territoriales Statistiques. NUTS-1 = landsdelen, NUTS-2 = provincies, NUTS-3 = COROP-gebieden (regio's binnen provincies; Utrecht en Flevoland zijn ieder een COROP-gebied)

verbetering ten opzichte van de meer bestuurskundige invalshoek kunnen worden beschouwd, in die zin dat vanuit theoretische of empirische evidentie naar verklaringen voor successen wordt gezocht. In de bestuurskundige invalshoek blijft de stap van 'verklaren' vaak impliciet of is zelfs afwezig. Er wordt daarbij vooral uitgegaan dat de werkwijze van de best presterende organisatie de verklaring vormt voor de goede prestaties. Het goede voorbeeld volstaat en dient ter inspiratie. De relatie tussen (aspecten van) de werkwijze van het meest succesvolle groepslid en de prestatie wordt niet onderzocht of onderbouwd.

De keerzijde van de medaille is dat de bestaande benchmarks (of beter: vergelijkingsstudies) op het gebied van verkeersveiligheid veel minder gericht zijn op de criteria en inhoudelijke eisen die bijvoorbeeld Korsten et al. (2013) vanuit bestuurskundig oogpunt stellen aan benchmarks: de internationale verkeersveiligheidsbenchmarks zijn door de vooral technische uitwerking bijvoorbeeld voor beleidsmedewerkers en politici niet eenvoudig te interpreteren. Ook wordt er geen gebruik gemaakt van aansprekende praktijkvoorbeelden of aanbevelingen om gebieden (meestal landen) te inspireren tot verbetering. Daarnaast wordt er ook niet actief samengewerkt met de betreffende gebieden om daadwerkelijk tot een leerproces te komen. In bestuurskundige termen zou gesteld kunnen worden dat de benchmarking van verkeersveiligheid tot nu toe beperkt is gebleven tot vergelijken en verklaren van verschillen en ontbreekt het leer- en implementeertraject dat in de bestuurskunde als essentiële onderdelen van benchmarking worden beschouwd.

5.2 Mogelijke verklaringen

Een mogelijke verklaring voor bovengenoemde verschillen ligt in de constatering dat bestuurskundigen vaker praktijkgericht zijn dan wetenschappers in veel andere disciplines, getuige alleen al de titel van het liber amicorum voor emeritus hoogleraren Herweijer en het jubileumboek van de Vereniging voor Bestuurskunde (respectievelijk: "In wetenschap voor de praktijk", Marseille & Winter, 2012 en "Brug tussen theorie en praktijk", Jong & Steur, 2008). Hun referentiekader staat daarbij dichterbij dat van de beleidsmaker dan van de wetenschapper pur sang. Caplan en Merton (Caplan 1979; Merton, 1973) beschrijven de uiteenlopende manieren waarop beleidsmakers en wetenschappers naar de wereld kijken. Zij zien een substantieel verschil in cultuur en institutionele arrangementen tussen de wereld van beleidsmakers en die van wetenschappers. Zowel de taal als de belangen en het beloningssysteem verschillen substantieel tussen deze twee werelden. Deze visie wordt de 'Two Communities' metafoor genoemd. In onderstaande *Tabel 2* heeft Leroy (2007) de verschillen tussen politiek en wetenschap op een rij gezet op het gebied van ambitie, ethische en kwaliteitsstandaarden, kennisbehoefte en methoden voor kwaliteitscontrole.

Tabel 2: Verschillen tussen wetenschap en politiek (Leroy, 2007).

	Wetenschap	Politiek
Is op zoek naar	waarheid	macht
Wordt gedreven door	waardevrijheid	normatieve overtuiging
Is uit op informatie met	diepgang	snellheid, focus op oplossingen
Is uit op informatie	in detail	op hoofdlijnen
Zoekt vooral naar	oorzaken	oplossingen
Kwaliteit gebaseerd op eis voor	validiteit	acceptatie
En op eis voor	betrouwbaarheid	haalbaarheid
Methode voor kwaliteitscontrole	collegiale toetsing	Draagvlak bij bevolking

Door deze verschillende visies op de wereld, kijken wetenschappers en beleidsmakers ook verschillend naar problemen in beleidsterreinen zoals die van de verkeersveiligheid. Hisschemöller en Hoppe (Hisschemöller & Hoppe, 1995; Hoppe, 2002a) hebben het denken over beleidsproblemen onderverdeeld langs twee dimensies: actoren kunnen het al dan niet eens zijn over de normen en waarden van een probleem, en actoren kunnen het al dan niet eens zijn over de vereiste kennis om het probleem op te lossen. Bax (2011) vond in haar onderzoek naar gebruik van verkeersveiligheidskennis door beleidsmakers dat wetenschappers en beleidsmakers in het verkeersveiligheidsveld het over het algemeen eens zijn over de waarden en normen van het probleem: verkeersongevallen, zeker waarbij doden en gewonden vallen zijn zeer onwenselijk. Men verschilt echter regelmatig van mening over de kennis die nodig is om beleidsproblemen in het verkeersveiligheidsveld op te lossen. Wetenschappers vinden over het algemeen verkeersveiligheid geen eenvoudig probleem, omdat het in wetenschappelijke termen niet zo eenvoudig is om bijvoorbeeld de oorzaken van ongevallen te onderzoeken of de effecten van verkeersveiligheidsmaatregelen vast te stellen.

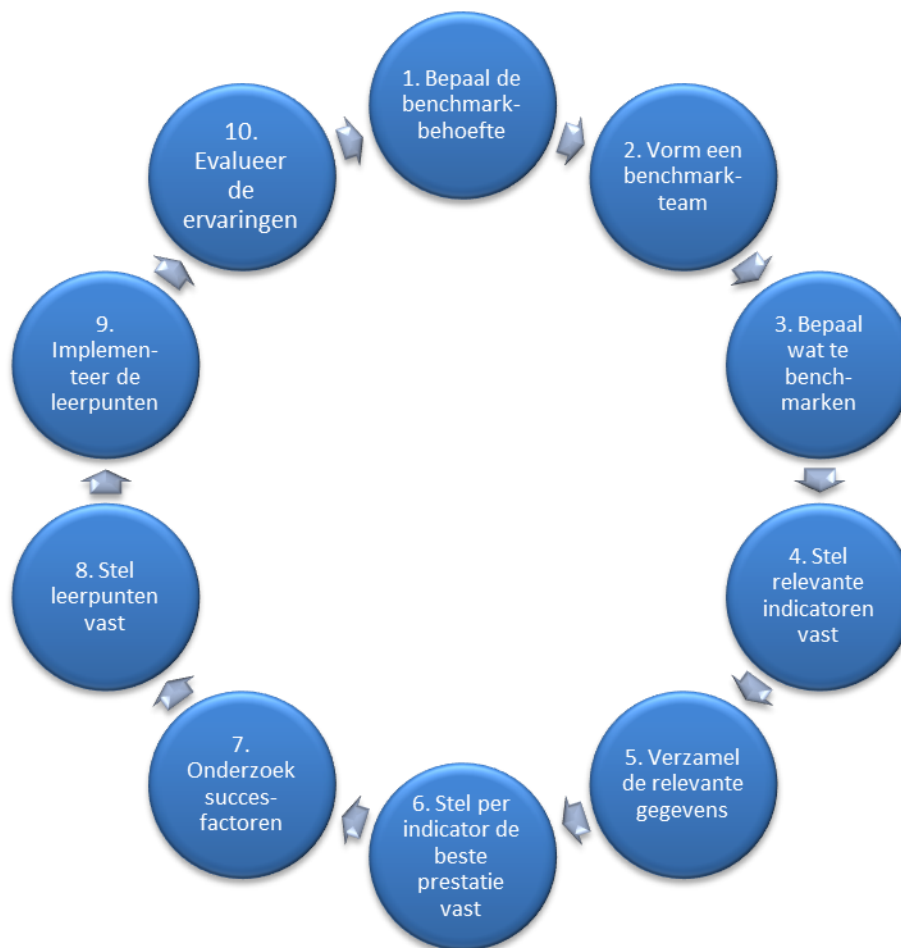
Beleidsmakers beschouwen over het algemeen verkeersveiligheidsproblemen als niet eenvoudige problemen omdat actoren die bij verkeersveiligheidsbeleid zijn betrokken (burgers, overheden, transportorganisaties, belangenverenigingen, hulpdiensten, OV-maatschappijen et cetera) het weliswaar eens zijn over het doel van het beleid (minder ongevallen), maar vaak niet over de maatregelen die genomen moeten worden. Alle actoren hebben daarbij hun eigen belangen, en beleidsmakers hebben de taak deze belangen tegen elkaar af te wegen in hun formele beslismacht over verkeersveiligheidsmaatregelen. Zij hebben daardoor veelal behoefte aan instrumenten die een analyse van het probleem en de oplossingen geven, zodat zij vervolgens de belangen op elkaar af kunnen afstemmen of een keuze kunnen maken. Deze verschillen in de kijk op verkeersveiligheid en op de benodigde kennis kan wellicht een reden zijn waarom de huidige, door wetenschappers ontwikkelde benchmarks op het terrein van verkeersveiligheid meer gericht zijn op technische aspecten dan op betrokkenheid van onderzochte doelgroepen.

5.3 Naar een synthese

Het doel van benchmarken is uiteindelijk dat organisaties (zoals landen, regio's, gemeenten) van elkaar leren en zichzelf verbeteren. Om dit doel te bereiken is het belangrijk dat benchmarking impact heeft op het beleid van de betrokken partijen. Uit de bestuurskundige literatuur leren we het belang van betrokkenheid van deze partijen in het benchmarkproces en het belang van bestuurlijk commitment. Deze werkwijze is bij de bestaande benchmarks of vergelijkingsstudies op het terrein van verkeersveiligheid nog onderbelicht gebleven. Het ligt derhalve voor de hand om samenwerking met overheden ook in dit domein te gaan verkennen.

Naast aandacht voor het proces is het ook belangrijk dat er aandacht blijft voor verklaringen van succesfactoren. De essentie hiervan is dat er verbanden worden gevonden tussen goede prestaties van organisaties en de succesfactoren waaraan deze prestaties kunnen worden toegeschreven. Dit is een sterk punt dat de verkeersveiligheidsstudies te bieden hebben voor de bestaande benchmarks met meer bestuurskundige invalshoek.

Figuur 5 bevat een synthese van de processtappen van benchmarking zoals onderscheiden door Bhutta & Huq (1999), aangevuld met de meer technische stappen van Wegman & Oppe (2010). Daarin wordt gepleit voor voldoende aandacht voor het proces samen met de te benchmarken partijen, als de onderbouwing van verklaringen voor verschillen die in het vergelijkingsproces aan het daglicht komen.



Figuur 5: Stappenplan voor het benchmarken van verkeersveiligheidsbeleid bij decentrale overheden (naar Bhutta & Huq, 1999 en Wegman & Oppe, 2010).

6. Naar een benchmark van verkeersveiligheidsbeleid voor gemeenten in Nederland

In de voorgaande paragrafen is een vergelijking gemaakt tussen de algemene benchmarkliteratuur en de concrete toepassing in de, als voorbeeld gekozen, verkeersveiligheidswereld. In deze paragraaf trekken we dit voorbeeld door en benoemen welke zaken in zo'n verkeersveiligheidsbenchmark nadere uitwerking zouden behoeven. Met de voorgaande analyse als uitgangspunt, zijn er voor een benchmark van verkeersveiligheid voor gemeenten de volgende factoren nader uit te werken:

- het proces rondom de benchmark;
- inhoud van de benchmark;
- indicatoren.

6.1 Het proces rondom de benchmark

Bij benchmarking staat centraal dat gemeenten uit een vergelijking van succesfactoren (stap 6 en 7, zie *Figuur 5*) met elkaar leren (stap 8) en deze leerpunten implementeren (stap 9). Om dit leren en implementeren optimaal te laten plaatsvinden, is commitment van gemeenten belangrijk (zie OECD, 2000; Korsten et al., 2013), ook op bestuurlijk niveau. Als in Nederland een benchmark

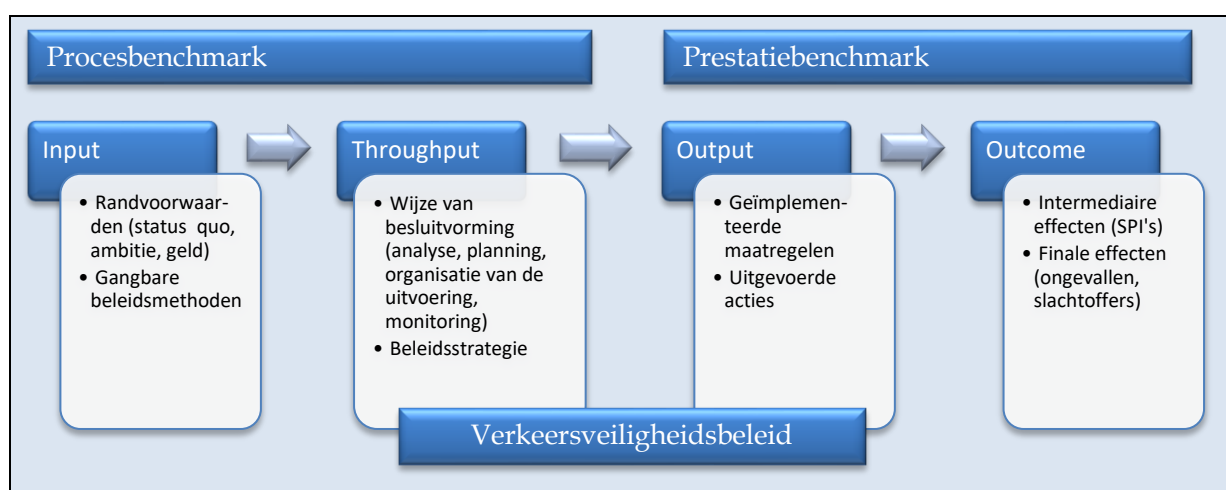
van verkeersveiligheid voor gemeenten van de grond kan komen, dan lijkt het bewerkstelligen van dit commitment een belangrijke eerste stap en randvoorwaarde (stap 1).

Vervolgens is samenwerking met gemeenten die willen benchmarken essentieel (stap 2). Dit vergt dus een proces van afstemming tussen onderzoekers en beleidsmakers: een belangrijke nieuwe dimensie aan de tot nu vooral academisch georiënteerde vergelijkingsstudies op het gebied van verkeersveiligheid. Om tot deze samenwerking te komen is het wel nodig om gemeenten goed te kunnen vergelijken en ook eensgezindheid te hebben over het specifieke onderwerp dat in de benchmark centraal staat (stap 3).

Voor een goede vergelijking is een belangrijke randvoorwaarde dat gemeenten in bepaalde mate op elkaar lijken. Uit diverse studies blijkt daarbij vooral stedelijkheidsgraad een belangrijke factor te zijn (zie voor een overzicht Houwing et al., 2012). Deze kan – eventueel afhankelijk van het specifieke onderwerp dat wordt gekozen – verder worden aangevuld met andere kenmerken (denk bijvoorbeeld aan gemeenten met een accent op fietsbeleid, gemeenten met veel industriële activiteiten of juist een sterk agrarisch karakter; zie bijvoorbeeld Houwing et al., (2012).

6.2 Inhoud van de benchmark

In zowel de bestuurskundige als in de verkeersveiligheidsliteratuur blijkt het onderscheid te worden aangehouden tussen strategische benchmarks, proces- en prestatiebenchmarks. Op grond van de beschouwingen van Eksler (2009), waarin hij constateert dat er weinig oog is voor de doorvertaling van een strategie in implementatieplannen, kunnen echter vraagtekens worden gezet bij het nut van een strategische benchmark ten aanzien van het onderwerp verkeersveiligheid. Het ligt meer in de rede om de proces- en prestatiebenchmark als uitgangspunt te nemen en deze zo mogelijk te combineren. In *Figuur 6* is weergegeven welke onderdelen van het beleidsproces gericht op verkeersveiligheid in deze typen benchmarks kunnen passen.



Figuur 6: Relatie tussen type benchmark en elementen in het beleidsproces, gericht op verkeersveiligheid.

De elementen in de verkeersveiligheids-piramide (zie *Figuur 3*) zijn grotendeels te koppelen aan het model van Input-Throughput-Output-Outcome, dat in de beleidswereld in het algemeen veel gebruikt wordt. Zo kunnen we de verkeersslachtoffers en de intermediaire effecten in termen van de veiligheidskwaliteit van het verkeerssysteem en risicovol gedrag (prestatie-indicatoren verkeersveiligheid, ook welbekend als SPI's) beschouwen als 'outcome'-indicatoren. De laag 'verkeersveiligheidsbeleid' is een diffuse laag en bevat enerzijds output-indicatoren (concrete maatregelen), maar daarnaast ook throughput-indicatoren (proces en strategie).

De verschillende stappen van het beleidsschema kunnen in de twee soorten benchmarks worden gemeten.

1. De *prestatiebenchmark* geeft alleen een prestatiescore ten opzichte van anderen; het gaat daarbij om 'output'- en 'outcome'-indicatoren. Deze scores bieden mogelijkheden om verbeteringen aan te brengen en van anderen te leren. In de verkeersveiligheid zijn verschillende soorten prestatiebenchmarks mogelijk:
 - a. van eindprestaties: aantallen doden en ernstig verkeersgewonden;
 - b. van intermediaire prestaties: kenmerken van gedrag, zoals snelheid, alcohol, helmgebruik en kwaliteit van het verkeerssysteem zoals de weginrichting;
 - c. van concrete maatregelen.
2. De *procesbenchmark* kijkt ook naar de onderliggende werkprocessen, dus naar de organisatie van het beleid: hoe worden de beschikbare verkeersveiligheidsbudgetten verdeeld, hoe zijn de doelstellingen onderbouwd, volgen maatregelen logisch uit een ongevalanalyse, is er een monitor opgesteld en dergelijke. In bovenstaand schema gaat het dan om de relatie tussen 'input'- en 'throughput'-indicatoren, waarbij de wijze van invulling van de 'throughput' centraal staat. Ook hier geeft vergelijking met andere decentrale overheden aanknopingspunten tot verbeteringen.

Om te beoordelen of de te benchmarken organisaties vergelijkbaar zijn of met vergelijkbare omstandigheden te maken hebben (zie ook vorige paragraaf), kan vóórdat met een benchmark begonnen wordt gekeken worden naar de specifieke omstandigheden van organisaties: betreft het bijvoorbeeld plattelandsgemeenten of steden? Is er sprake van een specifiek soort grondgebruik zoals industrie of landbouw? Is er sprake van veel toerisme?

Dit biedt uiteindelijk een raamwerk voor de verdere inhoudelijke uitwerking (stappen 3 t/m 7 in *Figuur 5*) van een benchmark van verkeersveiligheid. Het raamwerk is daarbij zowel van toepassing op een benchmark van de algemene verkeersveiligheid, als van specifiekere invalshoeken daarbinnen.

6.3 Indicatoren voor een verkeersveiligheidsbenchmark

Als randvoorwaarden voor te kiezen indicatoren zijn de volgende zaken in de literatuur als belangrijk aangemerkt (zie bijvoorbeeld Groot et al., 2006; Eksler, 2009; Korsten, 2006):

- de betrouwbaarheid en robuustheid van de indicatoren;
- begrijpelijkheid van de indicatoren voor de gebruikers;
- de behapbaarheid van de indicatoren.

Betrouwbaarheid en robuustheid

Binnen de wetenschappelijke verkeersveiligheidswereld is al veel onderzoek gedaan naar geschikte indicatoren, waaronder ook pogingen om tot een samengestelde indicator voor het presteren van bijvoorbeeld landen te komen (zie voor een overzicht Aarts & Bax, 2014). Al heeft dit onderzoek zich nog maar nauwelijks op gemeentelijk niveau gericht, toch biedt deze literatuur houvast om tot een uiteindelijke set indicatoren voor een verkeersveiligheidsbenchmark te komen. Dit lichten we nog verder toe en vatten we samen in *Tabel 3*.

Begrijpelijkheid

Een punt dat daarbij nog niet altijd evenveel aandacht heeft gehad, maar wel belangrijk is om bij gemeenten draagvlak voor benchmarking van verkeersveiligheid te krijgen, is de begrijpelijkheid van de indicatoren. Het ligt vanuit deze redenering niet voor de hand om in een benchmark voor gemeenten te werken met indicatoren die in één cijfer alle prestaties pogen samen te vatten. Werken met een aantal losse indicatoren die dicht bij de bron blijven sluit naar verwachting beter aan bij de doelgroep. Dit wil niet zeggen dat samengestelde indicatoren in enige mate uitgesloten zouden moeten zijn. Ondanks dat gemeenten in zekere mate vergelijkbaar zijn, blijft het belangrijk om zo veel mogelijk te corrigeren voor expositiefactoren, zoals de grootte van een gemeente of de grootte van een doelgroep, omdat deze factoren invloed hebben op een prestatie maat zoals aantal slachtoffers (zie ook Houwing et al., 2012 en Aarts & Bax, 2013). Hierbij moet eerst worden bepaald wat een geschikte maat is en vervolgens of deze maat daadwerkelijk op gemeentelijk niveau beschikbaar is.

Behapbaarheid

Een aandachtspunt bij het bepalen van een uiteindelijke set aan indicatoren, is vervolgens ook het *aantal* indicatoren dat in de benchmark betrokken gaat worden. Zoals ook Groot et al. (2006) constateren, is er bij gemeenten vaak behoefte om in korte tijd, en met niet al te veel inspanning, iets te kunnen zeggen over prestaties en verbeterpunten. Dat vraagt om een klein aantal indicatoren die met behoorlijk gemak zijn in te vullen. Echter, voor meer begrip van achterliggende oorzaken van prestaties is het soms noodzakelijk om meer de diepte in te gaan en meer informatie te hebben. Dit geldt zeker voor een terrein als verkeersveiligheid, waarvan eindprestaties (doden en gewonden) sterk beïnvloed worden door 'kans' en een complex aan invloedsfactoren kennen.

Voorstel voor indicatoren

Om toch aan zo veel mogelijk verschillende randvoorwaarden tegemoet te komen, is geïnspireerd door de werkwijze op het terrein van de openbare verlichting (Commissie Openbare Verlichting, 2011) een set indicatoren uitgewerkt voor benchmarking van verkeersveiligheid (zie *Tabel 3*). Daarbij is gewerkt met een beperkte set primaire indicatoren, aangevuld met een uitgebreidere set indicatoren. Het gaat hierbij zo veel mogelijk om indicatoren waarvan wetenschappelijk een causale relatie is aangetoond met het ontstaan van ongevallen of letsel of het voorkómen daarvan. Tevens is gebruikgemaakt van beschikbare kennis over gegevens die op gemeentelijk niveau voorhanden zijn.

Tabel 3: Overzicht van primaire en secundaire indicatoren voor een gemeentelijke benchmark van verkeersveiligheid. D = doden, EVG = ernstig verkeersgewonden.

Type indicator	Input	Throughput	Output	Outcome	
	Uitgangs-indicatoren	Procesindicatoren	Beleidsindicatoren	Tussenindicatoren (SPI's)	Eindindicatoren (slachtoffers)
Primaire indicatoren	Beschikbare budget voor verkeersveiligheid	<ul style="list-style-type: none"> • Uitgevoerde probleemanalyse als uitgangspunt voor beleidskeuzen • Of en zo ja hoe relevante beleidsprestaties worden gemonitord 	<ul style="list-style-type: none"> • Uitgevoerde maatregelen op het terrein van: <ul style="list-style-type: none"> - Infrastructuur - Handhaving - Voorlichting - Educatie - Overige maatregelen 	<ul style="list-style-type: none"> • Gemiddelde snelheid per wegtype • Kwaliteit van de infrastructuur 	<ul style="list-style-type: none"> • Doden en ernstig verkeersgewonden (gecorrigeerd met relevante en beschikbare expositiegraad zoals bevolkingsdichtheid, weglengte of verkeersintensiteit)
Secundaire indicatoren	Aanwezigheid van verkeersveiligheidsdoelstellingen en andere ambities	<ul style="list-style-type: none"> • Economische evaluaties • Partijen waarmee wordt samengewerkt 		<ul style="list-style-type: none"> • Spreiding in snelheid, aandeel overtredders • Kwaliteit van de fietsinfrastructuur • Rijden onder invloed van alcohol • Gordeldracht • Lichtvoering fietsers • Helmdracht tweewielers 	<ul style="list-style-type: none"> • Mortaliteit en morbiditeit in de belangrijkste leeftijdsgroepen: adolescenten en ouderen • Dichtheid van D+ EVG in de belangrijkste vervoerswijzen • Dichtheid van D+ EVG in de belangrijkste locaties⁵

In de benchmarkgroep moet uiteindelijk zelf een keuze worden gemaakt welke indicatoren worden meegenomen. Bij het kiezen van indicatoren is het belangrijk om inzichtelijk te maken wat de samenhang is met verkeersveiligheid. Idealiter betreft dit een verband dat uit (verschillende) empirische onderzoeken blijkt; minimaal gaat het om een verband dat theoretisch onderbouwd kan worden. Indicatoren die voor gemeenten interessant zijn, maar waarbij nog onvoldoende bekend is over de relatie met verkeersveiligheid, kunnen de kennisagenda voeden.

Metten van indicatoren

De genoemde indicatoren in Tabel 3 zijn deels beschikbaar in landelijke databestanden, deels beschikbaar bij gemeenten zelf en voor een paar indicatoren geldt dat deze nog verzameld moeten worden (zie Aarts, 2014 voor een overzicht). Voor de primaire indicatoren betekent dit:

Indicatoren die op basis van landelijke bronnen kunnen worden samengesteld betreffen:

- Morbiditeit en mortaliteit: gegevens over verkeersslachtoffers naar gemeente en voor het gemeentelijke wegennet zijn beschikbaar in de politieregistratie (BRON); voor de ernstig

⁵ EVG naar locatie is maar in beperkte mate mogelijk, namelijk alleen voor de EVG die door de politie worden geregistreerd. Dit zijn hoofdzakelijk EVG in ongevallen met gemotoriseerd verkeer.

verkeersgewonden (EVG) is een correctie beschikbaar op basis van medische gegevens uit ziekenhuizen (LMR). De belangrijkste expositiegegevens die worden gebruikt om de omvang van het aantal slachtoffers mee te corrigeren, betreffen bevolkingsomvang of weglengte naar gemeente. Deze gegevens zijn beschikbaar bij het Centraal Bureau voor Statistiek (CBS).

- Budget: informatie over gemeentelijke begrotingen, waarbij ook een post voor verkeer- en vervoer wordt onderscheiden, is beschikbaar bij CBS. Gemeenten kunnen, in aanvulling op de gegevens van CBS, preciezer aangeven welke budgetten direct of indirect voor verkeersveiligheid van belang zijn, zoals budgetten voor weginrichting en -onderhoud, campagnes, educatie.

Indicatoren die gemeenten in principe zelf hebben en niet in landelijke bestanden structureel ontsloten zijn betreffen:

- Snelheidsgegevens: gemeenten verzamelen doorgaans incidenteel snelheidsgegevens van verschillende typen wegen. Deze kunnen gebruikt worden om een indicatie te krijgen van gemiddelde snelheden en eventueel aanvullende maten op de belangrijkste wegtypen in beheer bij gemeenten (30 km/uur, 50 km/uur en 80 km/uur-wegen). Gemeentelijke snelheidsmetingen worden veelal uitgevoerd met radarapparatuur. Om een indruk te krijgen van gemiddelde snelheden kan aanvullend ook naar 'floating car data'-producten worden gekeken (bijvoorbeeld 'speed profiles' of aanverwante producten).
- Uitgevoerde maatregelen: welke maatregelen op het terrein van infrastructuur, verkeerseducatie, verkeersveiligheids campagnes, gemeenten zelf hebben uitgevoerd en op welke speerpunten hoeveel gehandhaafd is door politie, is informatie die alleen bij gemeenten zelf beschikbaar is. Het gaat hierbij om de specifieke maatregelen die zijn gepleegd, de hoeveelheid en locatie (infrastructuur) en de duur (educatie, handhaving en campagnes). Ook kan er sprake zijn van overige maatregelen met als doel om de verkeersveiligheid te verbeteren. Afhankelijk van de aard van deze maatregelen, kan de omvang en de inhoud hiervan verder gespecificeerd worden.
- Probleemanalyse en monitoring: gemeenten weten zelf het beste of er probleemanalyses aan hun beleidsplannen of - acties ten grondslag liggen en of ze hun prestaties monitoren.
- Budget: dit kan gemeten worden door te kijken naar het verkeer- en vervoerbudget (onderdeel wegen) of specifieke verkeersveiligheidsbudget in de begroting van de gemeente, naar de hoogte van de Brede Doeluitkeringen naar andere budgetten, zoals bijvoorbeeld specifieke doeluitkeringen van de provincie voor verkeersveiligheid. Ook kan budget gemeten worden door te kijken naar het aantal FTE dat binnen de gemeente is gealloceerd voor verkeersveiligheid.

In aanvulling op de primaire indicatoren is het voor gemeenten vermoedelijk ook niet heel veel moeite om aan te geven of ze nagaan wat kosten- en baten zijn van hun inspanningen en welke samenwerkingspartners ze op het terrein van verkeersveiligheid hebben en wat de kwaliteit is van die samenwerking.

Tot slot indicatoren die aanvullende gegevensverzameling vereisen:

- Kwaliteit van de infrastructuur (weg en fiets): er bestaan diverse instrumenten om de veiligheidskwaliteit van de infrastructuur te kunnen beoordelen, zoals de Duurzaam-Veilig gehaltemeter of het instrument voor Veilige Snelheden en Geloofwaardige Snelheidslimieten of de Road Protection Score uit EuroRAP (zie bijvoorbeeld Aarts et al., 2014). Deze instrumenten bepalen of - gegeven de snelheidslimiet en eventueel de functie van de weg - de aanwezige kenmerken tot een veilige weginrichting behoren of daar te weinig aan bijdragen. Voor deze instrumenten is het noodzakelijk dat er gegevens worden verzameld over relevante kenmerken van de infrastructuur. Op basis van deze kenmerken kan vervolgens een veiligheidsindicatie worden verkregen.

- Rijden onder invloed (alcohol): het behoort tot de mogelijkheden om door middel van een decentraal alcoholmeetnet een beeld te krijgen van het alcoholgebruik in een bepaald gebied. Voor zover dergelijke decentrale alcoholmeetnetten bestaan, hebben die vooral een regionaal karakter om aan voldoende metingen te kunnen komen die aan de gestelde randvoorwaarden te voldoen (zie bijvoorbeeld Houwing & Aarts, 2013). Het is echter denkbaar dat ook op gemeentelijk niveau metingen worden uitgevoerd die inzicht geven in het alcoholgebruik in de gemeente, eventueel zelfs verschillende gebieden in de gemeente (binnen versus buiten de bebouwde kom bijvoorbeeld). Hiervoor is een samenwerking met de politie noodzakelijk, evenals voldoende metingen in het gebied.

Fietsveiligheid is een belangrijk thema voor de gemeentelijke verkeersveiligheid omdat dit de belangrijkste vervoerwijze is van de ernstig gewonde verkeersslachtoffers. Een van de factoren daarbij is de (gemeentelijke) infrastructuur. Als gemeenten besluiten de veiligheidskwaliteit van de infrastructuur inderdaad in kaart te brengen, dan kan het een relatief kleine moeite zijn om daarbij ook de kwaliteit van de fietsinfrastructuur in kaart te brengen. Hiervoor is onlangs een instrument ontwikkeld: het Safe Cycle Network (Wijlhuizen et al., 2014). Dit instrument gaat na welke en hoeveel knelpunten en aandachtspunten er op de weg aanwezig zijn.

Omdat het aantal slachtoffers in gemeenten jaarlijks sterk kan fluctueren, is het raadzaam om voor een stabiel beeld het gemiddelde van bijvoorbeeld 10 jaar te nemen. Het mag duidelijk zijn dat dit betekent dat van indicatoren zoals gemeentelijke grenzen, expositiefactoren en prestatie-indicatoren ook gegevens moeten worden genomen van dezelfde periode. Voor de overige indicatoren is het van belang om gegevens uit een periode te nemen waarbij aannemelijk is dat ze invloed kunnen hebben gehad op de verkeersveiligheid in de periode die centraal staat in de benchmark.

Voor de gegevens die nog verzameld dienen te worden is het van belang dat deze tussen de te benchmarken gemeenten op een zelfde wijze verzameld worden.

Momenteel zijn er een aantal gemeenten samen met SWOV actief betrokken bij een pilot 'gemeentelijke verkeersveiligheidsbenchmark'. Deze pilot beoogt verdere inzichten te verschaffen in lessen die volgens de beschreven methode te leren zijn en hoe de methode ingezet en/of verder ontwikkeld kan worden voor de kortere en langere termijn.

Referenties

- Aarts, L.T. (2014). *Benchmarking van gemeentelijke verkeersveiligheid in de praktijk. Een verdere uitwerking en toetsing van behoeften bij gemeenten*. R-2014-34. Den Haag: SWOV.
- Aarts, L.T. & Bax, C.A. (2015). *Eerste gemeentelijke benchmarks verkeersveiligheid: uitwerking en verkenning met een aantal gemeenten*. R-2015-19. Den Haag: SWOV.
- Aarts, L., Dijkstra, A. & Bax, C. (2014). *ProMeV: Proactief Meten van Verkeersveiligheid. Inzicht in onveiligheid vóóordat er slachtoffers vallen*. R-2014-10. Den Haag: SWOV.
- Abma, T.A. (1996). *Responsive evaluation*. Proefschrift. Rotterdam: Erasmus Universiteit Rotterdam.
- Annema, J.A., van der Beek, N., Bulthuis, O. en Jansen, J. (2012). Ex-postevaluatie van zes provinciale wegenprojecten. *Tijdschrift Vervoerswetenschap*, 48(3), 3-15.
- Bax, C. A. (2011). *Processes and patterns. The utilisation of knowledge in Dutch road safety policy*. Proefschrift. Leidschendam: SWOV.
- Bax, C., Wesemann, P., Gitelman, V., Shen, Y., et al. (2012). *Developing a Road Safety Index*. Deliverable 4.9 of the EC FP7 project DaCoTA. European Commission, Brussels.
- Bhutta, K.S. & Huq, F. (1999). Benchmarking - best practices: an integrated approach. *Benchmarking: An International Journal*, 6(3), 254-268.
- Boogers, M., Schaap, L., Munckhof, E. D. v. d., & Karsten, N. (2008). *Decentralisatie als opgave. Een evaluatie van het decentralisatiebeleid van de Rijksoverheid, 1993-2008*. Tilburg: Universiteit van Tilburg.
- Breen, J. (2000). Benchmarking road safety in the UK with the rest of Europe. *Targets 2010: taking the road safety strategy forward; Proceedings of a conference organised by the Parliamentary Advisory Council for Transport Safety (PACTS)*, London, UK, 17th October 2000, PACTS.
- Brink, G. v. d. (2002). *Mondiger of moeilijker? Een studie naar de politieke habitus van hedendaagse burgers*. Wetenschappelijke Raad voor het Regeringsbeleid (Ed.). Den Haag: SDU Uitgevers.
- Camp, R.C. (1989). *Benchmarking: the search for industry best practices that lead to superior performance*. Milwaukee, Wisconsin: Quality press for the American Society for Quality Control.
- Caplan, N. (1979). The Two communities theory and knowledge utilization. *American Behavioural Scientist*, 22(3), 459-470.
- Commissie Openbare Verlichting (2011). *Kengetallen openbare verlichting*. Herziening juni 2011. Ede: NSVV.
- Coopers & Lybrand. (1994). *Survey of Benchmarking in the UK*. London: Coopers & Lybrand and CBI National Manufacturing Council.
- Dattakumar, R. & Jagadeesh, R. (2003). A review of literature on benchmarking. *Benchmarking: An International Journal*, 10(3), 176-209.
- Ecorys (2010). *Benchmark Gemeentelijk Ondernemingsklimaat 2010*. Thematische rapportage eindmeting. Rotterdam: Ecorys.
- Eksler, V. (2009). *Road mortality in the EU: a regional approach*. PhD Thesis. Versailles: Versailles St.-Quentin University.
- Eksler, V. (2010). Measuring and understanding road safety performance at local territorial level. *Safety Science*, 48, 1197-1202.

Eksler, V. & Lassarre, S. (2008). Evolution of road risk disparities at small-scale level: example of Belgium. *Journal of Safety Research*, 39(4), 417-427.

Eksler, V., Lassarre, S. & Thomas, I. (2008). Regional analysis of road mortality in Europe. *Public Health*, 122, 826-837.

Felsö, F., Groot, H. de & Heezik, A van (2011). *Benchmark gemeentelijk afoalbeheer. Een empirisch onderzoek naar de productiviteit en kostendoelmatigheid*. Delft: Innovatie en Publieke Sector Efficiëntie Studies, Technische Universiteit Delft.

Groot, M.I., Court, T. de la & Chang, I. (2006). *Randvoorwaarden voor een prikkelende benchmark Duurzaam inkopen*. Delft: CE.

Hermans, E., Brijs, T. & Wets, G. (2009a). *Developing a theoretical framework for road safety performance indicators and a methodology for creating a performance index*. RA-MOW-2008-010 Hasselt: Steunpunt Mobiliteit en Openbare Werken IMOB.

Hermans, E., Brijs, T., Wets, G. & Vanhoof, K. (2009b). Benchmarking road safety: lessons to learn from a data envelopment analysis. *Accident Analysis and Prevention*, 41, 174-182.

Hisschemöller, M. and Hoppe, R. (1995). Coping with intractable controversies: the case for problem structuring in policy design and analysis. *Knowledge and Policy*, 8(4), 40-61.

Hoppe, R. (2002). Cultures of Public Policy Problems. *Journal of Comparative Policy Analysis*, 4(3), 305-326.

Houwing, S. & Aarts, L.T. (2013). *Monitoring rijden onder invloed van alcohol. Handreiking voor een gestructureerd decentraal meetnet*. H-2013-1. Leidschendam: SWOV.

Houwing, S., Aarts, L.T., Reurings, M.C.B. & Bax, C.A. (2012). *Verkennde studie naar regionale verschillen in relatie tot verkeersveiligheid*. R-2012-10. Leidschendam: SWOV.

Jong, P. d., & Steur, B. (2008). *Brug tussen theorie en praktijk*. Vereniging voor Bestuurskunde. 1973-2008. Den Haag: DeltaHage bv.

KING (2011). *Handleiding KING keurmerk benchmarks 2012*. Versie 2.0. Den Haag: Kwaliteitsinstituut Nederlandse Gemeenten KING.

Korsten, A.F.A. (2006). *Ins & outs van benchmarking*. Geraadpleegd 5 december 2013, op URL: <http://www.arnokorsten.nl/PDF/Onderzoek/Benchmarking.pdf>

Korsten, A.F.A., Abma, K., & Meer, A.-D. van der (2013). *Mythen over benchmarking. Feiten en ficties rond het vergelijken van prestaties tussen organisaties, met bijzondere aandacht voor gemeenten*. Geraadpleegd 5 december 2013, op URL: <http://www.arnokorsten.nl/PDF/Onderzoek/Mythen%20over%20benchmarking.pdf>

Kozak, M. (2004). "Chapter 1. Overview of benchmarking theory." In: Kozak, M. (Ed.), *Destination benchmarking: concepts, practices and operations*. Wallingford: CABI Publishing.

Kyrö, P. (2003). Revising the concept and forms of benchmarking. *Benchmarking: An International Journal*, 10(3), 210-225.

Leroy, P. (2007). "Wetenschap, samenleving en politiek: een debat met voorbeelden uit de milieusfeer." In: Funtowicz, L. G. S., Grin J. & Leroy P. (Ed.). *Wetenschap, maatschappij, politiek: wie stuurt wie?* Delft: Eburon.

Merton, R. K. (1973). "The Normative Structure of Science." In: R. K. Merton (Ed.) *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago/London: University of Chicago Press.

Marseille, A. T., & Winter, H. B. (2012). *In wetenschap voor de praktijk*. Liber Amicorum Michiel Herweijer. Nijmegen: Wolf Legal Publishers.

Ministerie van Infrastructuur en Milieu (2012b) *Structuurvisie infrastructuur en ruimte. Nederland concurrerend, bereikbaar, leefbaar en veilig*. Den Haag: Ministerie van Infrastructuur en Milieu.

Moriarty, J.P. & Smallman, C. (2009). En route to a theory of benchmarking. *Benchmarking: An International Journal*, 16(4), 484-503.

OECD/OESO (2000). *Transport benchmarking. Methodologies, applications and data needs*. European Conference of Ministers of Transport, Proceedings of the Paris conference, 1999, Paris: OECD.

Pommer, E., Ooms, I. & Jansen, S. (2013). *Maten voor gemeenten 2013. Prestaties en uitgaven van gemeenten 2006-2011*. Den Haag: Sociaal en Cultureel Planbureau.

Rekenkamercommissie Gemeente Den Haag (2006). *Sturen op verkeersveiligheid. Onderzoek naar de verkeersveiligheid in de stad*. Den Haag: Rekenkamercommissie Gemeente Den Haag.

Rekenkamer Maastricht, (2010). *Onderzoek beheer openbare ruimte*. Maastricht: Rekenkamer Maastricht.

SGBO (2011). *Benchmark wwb2011*. Den Haag: SGBO Benchmarking.

Stern, E. (2003). *The rationale & challenges for evaluation of public policies*. Keynote speech held at the Fifth European Conference on Evaluation of the Structural Funds, Budapest. URL: http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CDEQFjAA&url=http%3A%2F%2Fec.europa.eu%2Fregional_policy%2Fsources%2Fdocconf%2Fbudapev al%2Fstern.pps&ei=2wHZUoSULOqq0QWR6oGQBQ&usg=AFQjCNEONRcogfoF7xxLv12BDFFc3HbiEg&bvm=bv.59568121,d.d2k

Ten Tije, S., Pieterse, W., Dijk, J. van & Jansen, J. (2010). *Balans in benchmarking*. Enschede: Universiteit Twente.

Van Helden, G.J. & Tillema, S. (2005). In search of a benchmarking theory for the public sector. *Financial Accountability & Management*, 21(3), 0267-4424

Wegman, F., Eksler, V., Hayes, S., Lynam, D., Morsink, P., & Oppe, S. (2005). *A comparative study of the development of road safety in the SUNflower+6 studies. Final Report*. Leidschendam: SWOV.

Wegman, F. & Oppe, S. (2010). Benchmarking road safety performances of countries. *Safety Science*, 48, 1203-1211.

Wijlhuizen, G.J. Dijkstra, A. & Petegem, J.W.H. van (2014). *Safe Cycling Network; Ontwikkeling van een systeem ter beoordeling van de veiligheid van fietsinfrastructuur*. R-2014-14. Den Haag: SWOV.

Yasin, M. M. (2002). The theory and practice of benchmarking: then and now. *Benchmarking: An International Journal*, 9(3), 217-243.