

Benchmarking van verkeersveiligheid

Dr. L.T. Aarts & dr. C.A. Bax

R-2014-5

Benchmarking van verkeersveiligheid

Een inventarisatie en aanbevelingen voor de opzet van
verkeersveiligheidsbenchmarks in Nederland

Documentbeschrijving

Rapportnummer:	R-2014-5
Titel:	Benchmarking van verkeersveiligheid
Ondertitel:	Een inventarisatie en aanbevelingen voor de opzet van verkeersveiligheidsbenchmarks in Nederland
Auteur(s):	Dr. L.T. Aarts & dr. C.A. Bax
Projectleider:	Dr. L.T. Aarts
Projectnummer SWOV:	C04.01
Trefwoord(en):	Road, safety, decentralization, policy, region, data acquisition, data processing, analysis (math), evaluation (assessment), decision process, efficiency, Netherlands.
Projectinhoud:	Decentrale overheden zijn in belangrijkere mate verantwoordelijk voor verbetering van de verkeersveiligheid in hun gebied. Een methode die daarbij behulpzaam kan zijn is 'benchmarking', globaal gezegd een vergelijking met andere decentrale overheden met als doel om van de best presterende te leren. In dit rapport is nagegaan wat onder benchmarking verstaan kan worden, en hoe benchmarking van verkeersveiligheidsbeleid kan worden uitgewerkt als methode om beleidsmakers op decentraal niveau te helpen bij effectief en efficiënt beleid.
Aantal pagina's:	54 + 10
Prijs:	€ 12,50
Uitgave:	SWOV, Den Haag, 2014

De informatie in deze publicatie is openbaar.
Overname is echter alleen toegestaan met bronvermelding.

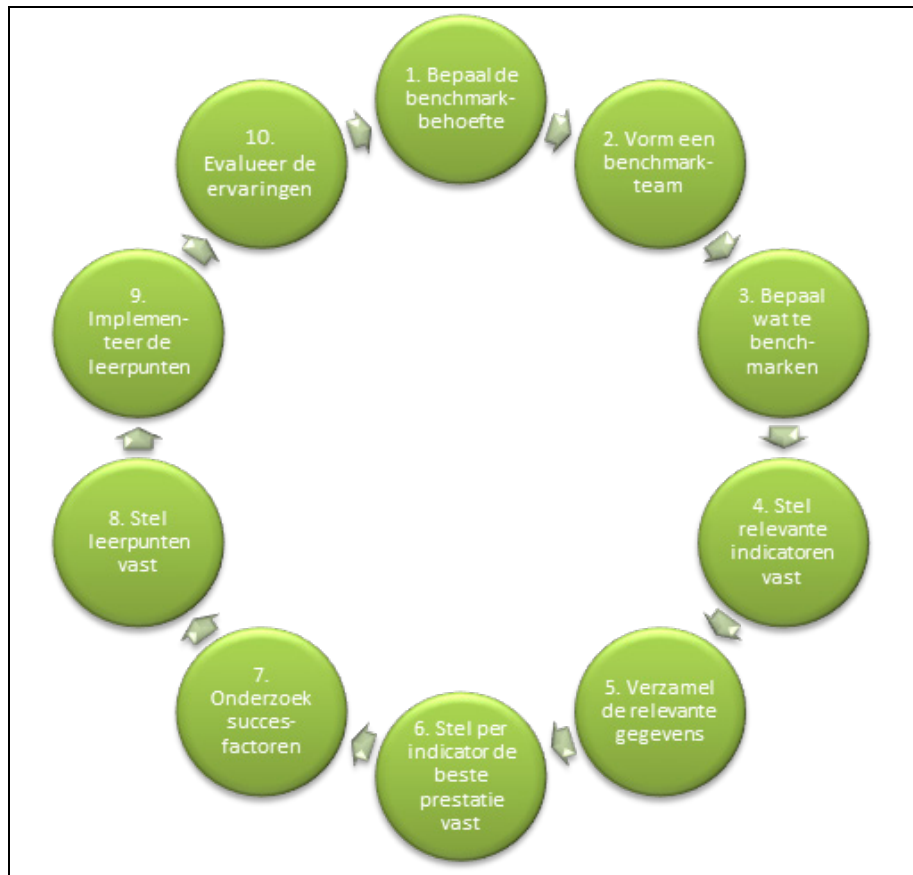
Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
Postbus 93113
2509 AC Den Haag
Telefoon 070 317 33 33
Telefax 070 320 12 61
E-mail info@swov.nl
Internet www.swov.nl

Samenvatting

Sinds het verkeersveiligheidsbeleid gedecentraliseerd is, zijn regionale en lokale overheden in belangrijkere mate verantwoordelijk voor verbetering van de verkeersveiligheid in hun gebied. Methoden om dit proces te ondersteunen, efficiënter en effectiever te maken kunnen daarbij behulpzaam zijn. Een methode die hiervoor interessant kan zijn, is 'benchmarking', globaal gezegd een vergelijking met andere gebieden, waarbij men kijkt wat men van het 'best presterende' gebied kan leren. Doel van dit rapport is dan ook om 'benchmarking' als methode te ontwikkelen om effectief en efficiënt verkeersveiligheidsbeleid bij decentrale overheden te stimuleren. Vragen die daarbij centraal staan zijn: Wat kan onder 'benchmarking' worden verstaan? Wat voor soorten benchmarks kunnen we onderscheiden? Hoe kan een verkeersveiligheidsbenchmark worden uitgewerkt? Welke stappen worden doorlopen? Met welke factoren dient rekening te worden gehouden? Welke voorbeelden van benchmarks zijn er al in Nederland op het gebied van verkeersveiligheid of andere verkeer-en-vervoerthema's? Wat kunnen we daarvan leren en welke kansen bieden deze voorbeelden?

Aan de hand van literatuur uit de bestuurskunde en het verkeersveiligheids-onderzoek is bestudeerd wat de heersende opvattingen zijn over wat benchmarking is, en hoe benchmarking voor de verkeersveiligheid wordt uitgewerkt en ingezet. Het valt daarbij op dat de bestuurskundige literatuur de nadruk legt op de proceskant en het uitvoeren van een benchmarkstudie door, of ten minste mét de organisatie zelf, terwijl de verkeersveiligheids-literatuur veel meer het accent legt op de technische uitwerking van benchmarking en het kiezen van relevante (samengestelde) indicatoren om prestaties van verschillende organisaties (vooral landen) te vergelijken. Dit is veel meer een activiteit van onderzoekers.

Op basis van de literatuur concluderen we dat voor verbetering van verkeersveiligheidsbeleid binnen Nederland, benchmarking kan worden verstaan als: *een proces van vergelijken van prestaties, processen of strategieën van vergelijkbare organisaties (bijvoorbeeld overheden zoals gemeenten) door of samen met deze organisaties, waarbij naar verklaringen voor onderlinge verschillen wordt gezocht, succesfactoren worden aangewezen en leerpunten worden geïmplementeerd.* Gebaseerd op zowel de bestuurskundige als de verkeersveiligheidsliteratuur, zijn er volgens deze definitie tien stappen binnen het benchmarkproces te nemen (zie *Afbeelding 1*).



Afbeelding 1. Tienstappenplan van benchmarking van verkeersveiligheidsbeleid waarbij zowel proces als inhoud aandacht krijgt (op basis van Bhutta & Huq, 1999 en Wegman & Oppe, 2010).

Om daadwerkelijk tot toepassing van benchmarking binnen Nederland te komen, constateren we dat het belangrijk is om de te benchmarken organisaties zelf (voor ons doel vooral decentrale overheden) bij het proces te betrekken en om de technische indicatoren (die meer de focus zijn van bij het onderwerp betrokken wetenschappers) niet te complex te maken. De verkeersveiligheidspiramide die in het verkeersveiligheidsveld vaak als theoretisch raamwerk wordt gebruikt, biedt inhoudelijk een prima uitgangspunt voor te benchmarken indicatoren van prestaties en processen (zie hieronder).

De verschillende typen benchmarks – vergelijking van prestaties, processen of strategieën – kunnen voor toepassing in verkeersveiligheidsbeleid bestaan uit de volgende stappen:

Stappen binnen een *prestatiebenchmark*:

1. vergelijking van genormeerde eindindicatoren (doden en ernstig verkeersgewonden);
2. vergelijking van intermediaire indicatoren (SPI's) en deze in verband brengen met de eindindicatoren;
3. vergelijking van maatregelen en deze in verband brengen met de vastgestelde intermediaire en eindindicatoren.

Stappen binnen een *procesbenchmark*:

1. vergelijking van doelstellingen en ambities;
2. vergelijking van inhoudelijke beleidskeuzen in relatie tot de geformuleerde doelstellingen, inclusief de onderbouwing van die keuzen;
3. vergelijking van budgetten en de allocatie daarvan in relatie tot de beleidskeuzen en -doelen;
4. vergelijking van organisatie van de uitvoering, waarbij speciaal aandacht is voor kwaliteitsbewaking en de relatie tot de geformuleerde beleidsplannen;
5. vergelijking van de uiteindelijk genomen maatregelen in relatie tot de beleidsdoelen.

Bij een *strategische benchmark* is het van belang om strategische keuzen in relatie tot doelen en ambities te vergelijken, te bekijken hoe deze strategische keuzen inhoudelijk en procesmatig worden uitgewerkt en hoe de kwaliteit van die uitwerking wordt gewaarborgd.

Behalve naar literatuur is ook gekeken naar praktijkvoorbeelden in Nederland van benchmarking op het terrein van verkeer en vervoer, en specifiek daarbinnen van verkeersveiligheid. Wat daarbij opvalt is dat er al wel enkele voorbeelden zijn van benchmarks waarin eindprestaties van regio's worden vergeleken. Dit zijn bijvoorbeeld de cijfers op de vergelijkingssite van het Kwaliteitsinstituut Nederlandse Gemeenten (KING), het dashboard Duurzame mobiliteit van het Kennisplatform Verkeer en Vervoer (KpVV), de app verkeersveiligheid van bureau VIA.nl of zelfs vergelijkende studies van de verkeersveiligheid van regio's. Er zijn echter nog geen voorbeelden waarin (samen met die overheden) ook naar achterliggende processen en naar verklaringen van eindprestaties is gekeken. Met benchmarking van verkeersveiligheid zouden de SWOV en decentrale overheden bij organisaties zoals KING en KpVV en hun producten kunnen aansluiten.

Summary

Benchmarking road safety; Stocktaking and recommendations for the development of road safety benchmarks in the Netherlands

Road safety policy has been decentralized in the Netherlands, giving regional and local governments a greater responsibility for road safety improvement in their jurisdiction. Methods to support this process and make it more efficient and effective can be of assistance. One method that may be of interest for this purpose is benchmarking, generally speaking a comparison with other fields in which study is made of what could be learned of the 'best in class'. The present report therefore aims to develop benchmarking as a method to encourage effective and efficient road safety policymaking by regional and local governments. Central questions are: How to define benchmarking? What types of benchmarks can be distinguished? How can a road safety benchmark be developed for regional and local authorities? Which steps must be taken? Which factors must be taken into account? Which examples of benchmarks in the field of road safety or other traffic and transport themes do already exist in the Netherlands? What can these examples teach us and which opportunities do they provide?

Based on administrative sciences literature and road safety research, study was made of the present opinions about benchmarking and of how benchmarking for road safety is being developed. It is noteworthy that administrative sciences literature focuses on the process and performing a benchmark study by or at least with the organization itself, whereas road safety literature focuses more on the technical development of benchmarking and choosing relevant (composite) indicators to compare the performances of different organizations (mainly countries). This is mainly done by researchers.

Based on the literature we conclude that benchmarking for improvement of road safety policy may be defined as: *a process of comparing performances, processes or strategies of similar organizations (e.g. governments like municipalities) by or together with these organizations, in which explanations are sought for differences, success factors are identified and lessons learned are implemented.* Based on both administrative and road safety literature, this definition indicates ten steps that are to be taken within the benchmarking process (see *Figure 1*).

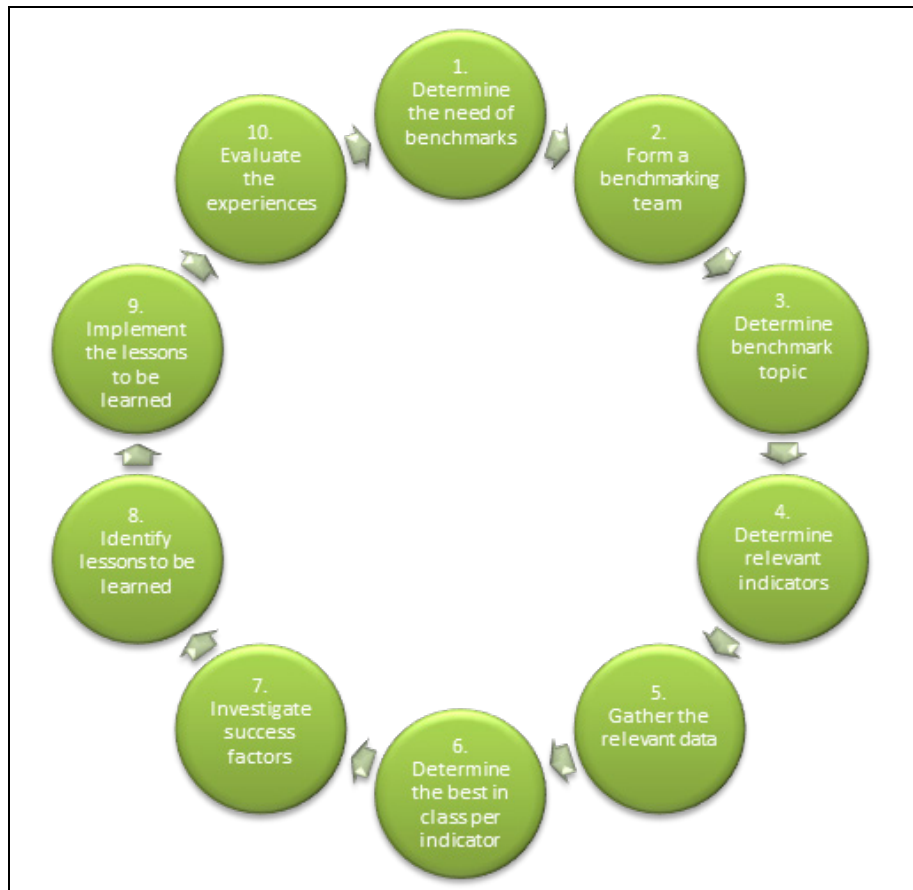


Figure 1. *Ten steps for benchmarking road safety policy, including process as well as content (based on Bhutta & Huq, 1999 and Wegman & Oppe, 2010).*

In order to implement benchmarking in regional and local jurisdictions, we conclude that it is important that the organizations that are to be benchmarked are involved in the process and that the technical indicators that are used in the benchmarking process are not made too complex. The road safety pyramid which is often used as a theoretical framework in road safety, offers an excellent starting point for benchmarking indicators of performances and processes (see below).

The different types of benchmarks – comparison of performances, processes or strategies – for road safety policy may consist of the following steps:

Steps within a *performance benchmark*:

1. comparison of standardized final indicators (fatalities and serious road injuries);
2. comparison of intermediate indicators (SPIs) and relating these to the final indicators;
3. comparison of road safety measures and relating these to the established intermediate and final indicators.

Steps within a *process benchmark*:

1. comparison of objectives and ambitions;
2. comparison of substantive policy choices in relation with the formulated objectives, including the evidence that supports these choices;
3. comparison of budgets and how these are allocated in relation with the policy choices and objectives;
4. comparison of how implementation is organized, with special attention for quality control and the relation with the formulated policy plans;
5. comparison of the final measures taken in relation with the policy objectives.

For a *strategic benchmark* it is important to compare strategic choices in relation with objectives and ambitions, to look at the elaboration and process development of these strategic choices and at how the quality of this elaboration is guaranteed.

Not only literature, but practical examples of benchmarking in the field of traffic and transport in the Netherlands, and road safety in particular, were also considered. Noteworthy is that there are already some examples of benchmarks in which the final performances of regions are compared. Examples are the figures on the comparison website of the Quality Institute of the Dutch Municipalities (KING), the Sustainable Mobility Dashboard of the Knowledge Platform for Traffic and Transport (KpVV), the road safety app of VIA.nl, or even comparative studies into regional road safety. However, so far there are no examples of benchmarks in which the underlying processes and explanations of the final performance outcomes have been compared or have been processed together with the governments involved. These examples of road safety benchmarking products could be used to inspire a starting point and it could be interesting to cooperate with the organisations that are already involved in benchmarking activities focusing on regional and local governments.

Inhoud

1. Inleiding en vraagstelling	11	
1.1. Vraagstelling en doel van in dit rapport	11	
1.2. Wat is benchmarking?	11	
1.3. Definities van benchmarking	12	
1.4. Over dit rapport	13	
2. Achtergronden van benchmarking vanuit bestuurskundige context	15	
2.1. De oorsprong van benchmarking	15	
2.2. Waarom benchmarken?	15	
2.3. Soorten benchmarks	16	
2.4. Eisen aan benchmarking	20	
2.5. Werkt benchmarking?	22	
2.6. Conclusie	22	
3. Onderzoek naar benchmarking in de wetenschappelijke verkeersveiligheidsliteratuur	24	
3.1. Identificeren van de belangrijkste componenten in verkeersveiligheidsprestaties	25	
3.2. Identificeren van de vergelijkingsgroep	25	
3.3. Construeren van indicatoren om te vergelijken	28	
3.4. Vaststellen en verklaren van verschillen in prestaties	30	
3.5. Vaststellen van na te streven prestaties in de toekomst	31	
3.6. Conclusies	32	
4. Uitwerkingen van benchmarking in de praktijk	33	
4.1. Werkwijze KING: stimulering en kwaliteitsbewaking van benchmarking	33	
4.2. Duurzame mobiliteit: KpVV	34	
4.3. Fietsbalans van de Fietsersbond	36	
4.4. ViaStat Dashboard	37	
4.5. SWOV Analyses Verkeersonveiligheid (SAVE)	38	
4.6. Diverse regionale vergelijkingsstudies (SWOV)	39	
5. Conclusies en aanbevelingen	41	
5.1. Definities van benchmarking en consequenties voor de praktijk	41	
5.2. Inzichten uit de bestuurskundige benchmarkliteratuur	42	
5.3. Inzichten uit de wetenschappelijke verkeersveiligheidsliteratuur over benchmarking	42	
5.4. Conclusies voor een vervolg	42	
5.5. Tot slot	48	
Literatuur	49	
Bijlage A	Details ten aanzien van indicatoren van verkeersveiligheid en het gebruik daarvan in benchmarking	55
Bijlage B	Details ten aanzien van producten van KING	59

1. Inleiding en vraagstelling

Sinds het verkeersveiligheidsbeleid sterk gedecentraliseerd is, staan regio's en lokale overheden aan de lat om – samen met de centrale overheid en andere relevante partijen – de verkeersveiligheid op een effectieve en efficiënte manier te verbeteren. De SWOV wil daarin graag ondersteunend zijn, en onderzoekt daarom onder meer methoden die decentrale overheden kunnen benutten bij het vormgeven van effectief en efficiënt verkeersveiligheidsbeleid.

Een methode die interessant kan zijn voor decentrale overheden is 'benchmarking', globaal gezegd een vergelijking met andere decentrale overheden met als doel om van de best presterende te leren. Bij benchmarking kan ook naar oorzaken worden gezocht achter (verschillen in) prestaties en problemen op het gebied van verkeersveiligheid, kan wetenschappelijke kennis over bijvoorbeeld de relatie tussen maatregelen of prestatie-indicatoren en verkeersveiligheid doorwerken in de praktijk en kan inspiratie worden gehaald uit de wijze waarop andere, vergelijkbare decentrale overheden tot goede prestaties komen.

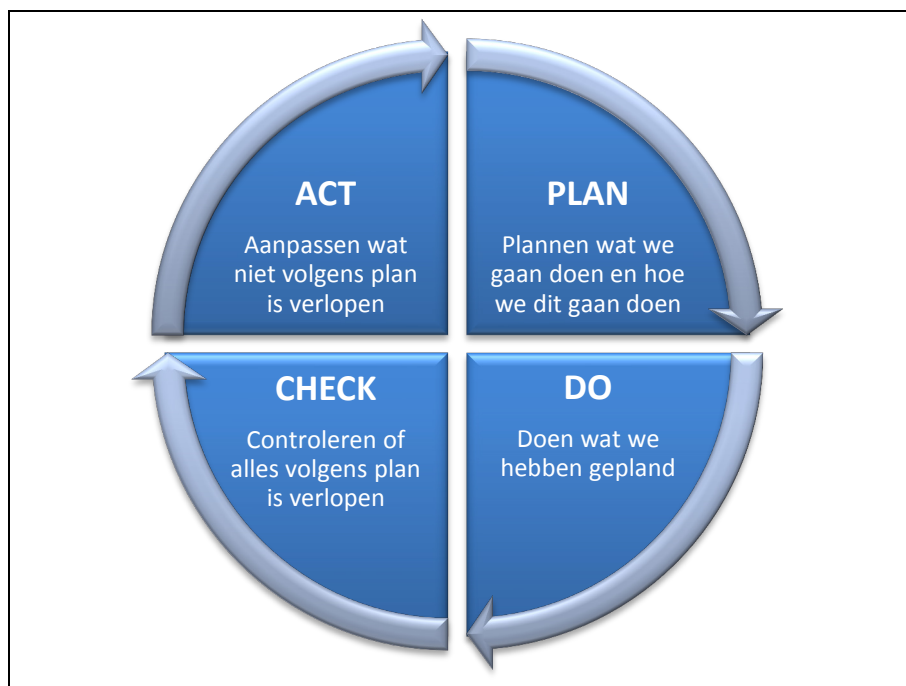
1.1. Vraagstelling en doel van in dit rapport

Doel van dit rapport is om 'benchmarking' als methode te ontwikkelen om effectief en efficiënt verkeersveiligheidsbeleid bij decentrale overheden te stimuleren. Vragen die daarbij centraal staan zijn: Wat kan onder benchmarking worden verstaan? Wat voor soorten benchmarks kunnen we onderscheiden? Hoe kan een verkeersveiligheidsbenchmark worden uitgewerkt? Welke stappen worden doorlopen? Met welke factoren dient rekening te worden gehouden? Welke voorbeelden van benchmarks zijn er al in Nederland op het gebied van verkeersveiligheid of andere verkeer- en vervoersthema's? Wat kunnen we daarvan leren en welke kansen bieden deze voorbeelden?

In dit hoofdstuk gaan we eerst in op wat we onder benchmarking kunnen verstaan en sluiten we af met de leeswijzer voor de rest van dit rapport.

1.2. Wat is benchmarking?

Beleidsmakers staan regelmatig voor de vraag in hoeverre de situatie binnen hun beheergebied voldoet aan gestelde doelen, of daarnaar goed op weg is. Een evaluatie van het beleid is dan op zijn plaats. Evaluatie is een vast onderdeel van 'de beleidscyclus' (zie bijvoorbeeld Hoogerwerf (1993)) Hij onderscheidt binnen het beleidsproces bijvoorbeeld agendasetting, beleidsvoorbereiding, beleidsbepaling, uitvoering, naleving en handhaving, en tot slot de evaluatie van beleid. De uitkomsten van de evaluatie kunnen leiden tot bijstelling van beleid of tot nieuw beleid. Een soortgelijke aanpak volgt de PDCA-cyclus van Shewhart (1939) en Deming (1986) (zie *Afbeelding 1.1*).



Afbeelding 1.1. De Plan-Do-Check-Act- ofwel PDCA-cyclus van Shewhart (1939) en Deming (1986) die als belangrijk actiemodel fungeert voor beleidsmakers.

De PDCA-cyclus gaat uit van een plan dat wordt uitgevoerd, na uitvoering wordt gecheckt of de uitvoering volgens dat plan is verlopen, en de uitvoering zo nodig nog wordt bijgesteld, waarna de cyclus weer opnieuw wordt doorlopen. Dit hele proces is erop gericht om uiteindelijk een doel te verwezenlijken of een doel dichterbij te brengen.

Deze beleidsevaluatie – of ‘check’ in de termen van de PDCA-cyclus – kan gebeuren met behulp van een eenvoudige beleidsevaluatie, door uitkomsten van het beleid met de beleidsdoelen te vergelijken. Een andere manier om beleid te evalueren is door het eigen beleid en de uitvoering te vergelijken met die van andere organisaties en vervolgens te bezien in hoeverre bijstelling van het eigen beleid of van de uitvoering nodig is. Dat is in grote lijnen wat gedaan wordt als we spreken over benchmarking.

1.3. Definities van benchmarking

Definities van benchmarking zijn in allerlei soorten te vinden. Een ‘benchmark’ is het beste te vertalen als: criterium, referentiemaat, referentiepunt of ijkpunt. ‘Benchmarking’ is dan te vertalen als ‘vergelijkend onderzoek’ waarbij er een referentiepunt of ijkpunt wordt gedefinieerd dat richtinggevend wordt voor anderen in het vergelijkend onderzoek.

De meeste definities vatten benchmarking op 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". Moriarty & Smallman (2009) spreken bijvoorbeeld van *"een proces binnen een organisatie om een bestaande werkpraktijk te veranderen in een betere, aan de hand van een voorbeeld"*. Kozak (2004) geeft een overzicht van de definities van tien verschillende auteurs en spreekt zelf van *"een handeling om specifieke praktijken te ontdekken die verantwoordelijk zijn voor goede prestaties, om te begrijpen hoe deze praktijken werken en kunnen worden gerealiseerd, en deze aan te passen aan de eigen organisatie"*.

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"*. Wij hanteren in dit rapport de bovengenoemde definitie van Camp.

Belangrijk in de definitie van benchmarking is dat een organisatie(onderdeel) niet alleen wordt vergeleken met één of meer anderen, maar dat de resultaten daarvan ook worden teruggekoppeld naar de eigen organisatie en dat deze stappen onderneemt om te leren van de analyse en om verbeteringen te implementeren. Dit betekent dat – volgens bovengenoemde definities – benchmarking geen proces is dat buiten de te benchmarken organisatie plaatsvindt, maar samen mét of zelfs dóór de organisatie. De definities gaan er impliciet van uit dat organisaties zelf verantwoordelijk zijn voor de benchmark en deze zelf uitvoeren, eventueel met behulp van een adviseur of onderzoeker. In dit rapport kijken we naar een specifiek soort organisaties, namelijk decentrale overheden en naar prestaties of processen en hun effecten op verkeersveiligheid.

Binnen de verkeersveiligheidsliteratuur blijkt het begrip benchmarking vooral gebruikt te worden om een vergelijkend onderzoek aan te duiden. Zo definiëren Wegman & Oppe (2010) benchmarking als *"[...] a process in which countries or sub-national jurisdictions evaluate various aspects of their performance in relation to other, and so-called 'best-in-class' practices."* (p. 1206). Opvallend aan deze definitie is dat de nadruk wordt gelegd op het vergelijkingsaspect van benchmarken, maar dat niet wordt gesproken over leren en verbeteren. Soms wordt het begrip benchmarken ook gebruikt om de prestaties van een organisatie te vergelijken met een externe norm, een (beleids)doelstelling of taakstelling. In de verkeersveiligheid kan beleid bijvoorbeeld vergeleken worden met de landelijke taakstellingen, of met een ander absoluut criterium als 'nul doden'. In de bestuurskundige literatuur wordt hiervoor echter niet het begrip benchmarking gebruikt maar het begrip *prestatiemeting*, wat een onderdeel is van benchmarking (Korsten, 2006).

1.4. Over dit rapport

Verkeersveiligheidsbeleid kan profiteren van benchmarking wanneer verantwoordelijke organisaties de beleidscyclus (of PDCA-cyclus) toepassen. Daarbinnen kan benchmarking een instrument zijn om tot verbeteringen te komen. Voor Nederland kan dat betekenen dat we leren van andere landen die goede resultaten behalen; de spoeling is daar dun

(zie ook de SUNflower-projecten: Koornstra et al., 2002; Wegman et al., 2005; Wegman et al., 2008). Binnen regio's in Nederland is benchmarking van verkeersveiligheidsbeleid nog niet of nauwelijks uitgewerkt, terwijl het wel kansen biedt voor met name decentrale overheden om hun beleid effectiever en efficiënter te maken. In dit rapport verkennen we vanuit de theorie en de bestaande praktijk hoe een verkeersveiligheidsbenchmark er in Nederland uit kan zien. Doel hiervan is om benchmarking als methode te ontwikkelen om effectief en efficiënt verkeersveiligheidsbeleid bij decentrale overheden te stimuleren.

Wat benchmarking kan betekenen en hoe het kan worden uitgewerkt, benaderen we zowel vanuit de meer algemene bestuurskundige literatuur over dit onderwerp (*Hoofdstuk 2*) als vanuit de verkeersveiligheidsliteratuur (*Hoofdstuk 3*). Ook bekijken we praktijkvoorbeelden van de 'benchmark-gedachte' in de wereld van verkeer en vervoer (*Hoofdstuk 4*). Al dit materiaal biedt bouwstenen voor een uitwerking van 'benchmarking' op het gebied van verkeersveiligheid voor decentrale overheden in Nederland (*Hoofdstuk 5*).

2. Achtergronden van benchmarking vanuit bestuurskundige context

Benchmarking is een begrip dat in eerste instantie vooral in een bestuurskundige context is bestudeerd en uitgewerkt. Daarom benaderen we in dit hoofdstuk dit begrip als eerste vanuit die invalshoek. We gaan kort in op de oorsprong van het begrip, vervolgens gaan we in op doelen van benchmarking, soorten benchmarks die in de literatuur worden onderscheiden en de eisen die er aan benchmarks te stellen zijn. Ook staan we stil bij de vraag of benchmarking werkt: levert benchmarking daadwerkelijk betere resultaten op? We trekken uit deze literatuurbeschouwing de eerste conclusies.

2.1. De oorsprong van benchmarking

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 verder ontwikkeld, eerst in wat in het Engels 'blue collar'-bedrijven wordt genoemd, bedrijven waarin bepaalde processen of producten routinematig in grote hoeveelheden worden uitgevoerd of gemaakt. Later kwam benchmarking ook in zwang voor 'white collar'-organisaties (financiële instellingen, detailhandel en dergelijke) en nog later, onder invloed van de New Public Management-beweging ook bij overheden (Francis & Holloway, 2007; Korsten, 2006; Yasin, 2002).

2.2. Waarom benchmarken?

Grofweg zijn er twee belangrijke redenen om een benchmark op te starten (Korsten, 2006). Allereerst kan een benchmark een organisatie een goed inzicht geven in haar prestaties binnen de branche. Dat kan een vorm van zelfbevestiging zijn, maar kan ook aanleiding geven tot verbeteringen.

Goede praktijkvoorbeelden kunnen inspiratie bieden voor nieuwe werkmethoden. Het gaat hier dan vooral om 'leren'. Een tweede reden voor een benchmark kan zijn om de activiteiten van de organisatie te verantwoorden, bijvoorbeeld tegenover de samenleving of een financier. Een benchmark draagt zo bij aan transparantie. Ook bij een benchmark om te verantwoorden is één van de doelen van benchmarking het verbeteren van de eigen prestaties. Dat kan zorgen voor een spanning tussen verantwoorden en verbeteren: voor een verantwoording geven organisaties meestal een zo *gunstig* mogelijk beeld van hun prestaties; om te kunnen leren en verbeteren is daarentegen een zo *eerlijk* mogelijk beeld nodig. Stern (2003) pleit ervoor deze spanning op te lossen door organisaties 'accountable for learning' te maken, ofwel organisaties af te rekenen op hun leervermogen.

De eerste reden voor benchmarking gaat meestal op voor vrijwillig opgestarte benchmarks, benchmarks opgestart ter verantwoording hebben vaak een verplicht karakter. Bij het benchmarken van verkeersveiligheidsprestaties van decentrale overheden zal het, in ieder geval voorlopig, gaan om een vrijwillige benchmark. Deze kan decentrale overheden inzicht geven in hun positie ten opzichte van anderen en aanknopingspunten bieden voor

nieuwe werkmethode en nieuw beleid om de verkeersveiligheid te verbeteren.

2.3. Soorten benchmarks

Nederlandse en buitenlandse overzichtsliteratuur (Bhutta & Huq, 1999; Korsten, 2006; Korsten et al., 2013; Kyrö, 2003; Poerstamper et al., 2007) laten zien dat er veel manieren zijn om benchmarks in te delen. Zo kunnen de volgende soorten benchmarking onderscheiden worden:

- vrijwillig of onvrijwillig;
- gericht op transparantie, leren of op verantwoording;
- horizontale (tussen organisaties van dezelfde hiërarchie) en verticale (tussen organisaties van verschillende hiërarchische niveaus, bijvoorbeeld tussen Rijk, provincies en gemeenten);
- van één of van meerdere organisaties;
- top-down of bottom-up gestuurd;
- interne of externe.

We lichten deze verschillende vormen niet allemaal uitgebreid toe, omdat de meeste termen voor zich spreken. Wel geven we hieronder een overzicht van de belangrijkste vormen die in de internationale literatuur worden onderscheiden.

2.3.1. Ontwikkeling in de tijd

Wat opvalt in de buitenlandse literatuur (Bhutta & Huq, 1999; Kozak, 2004; Kyrö, 2003; Moriarty & Smallman, 2009) is dat in de loop der jaren de definitie van benchmarking verschuift van intern vergelijken (gelijksoortige afdelingen binnen één organisatie, bijvoorbeeld bij Xerox, zie § 2.1), via extern vergelijken (gelijksoortige organisaties) naar meer nadruk op leren van andere organisaties die soms sterk van de eigen organisatie verschillen. Deze historische ontwikkeling heeft zich in de internationale literatuur vertaald in een reeks van diverse soorten benchmarks (zie bijvoorbeeld Kyrö, 2003). In de literatuur worden deze typen ‘verschillende generaties’ benchmarks genoemd, waarmee niet alleen wordt aangegeven dat ze in opeenvolgende tijdperioden benoemd of ontwikkeld zijn, maar waarbij ook wordt aangegeven dat sommige, oudere typen tegenwoordig minder gebruikt worden.

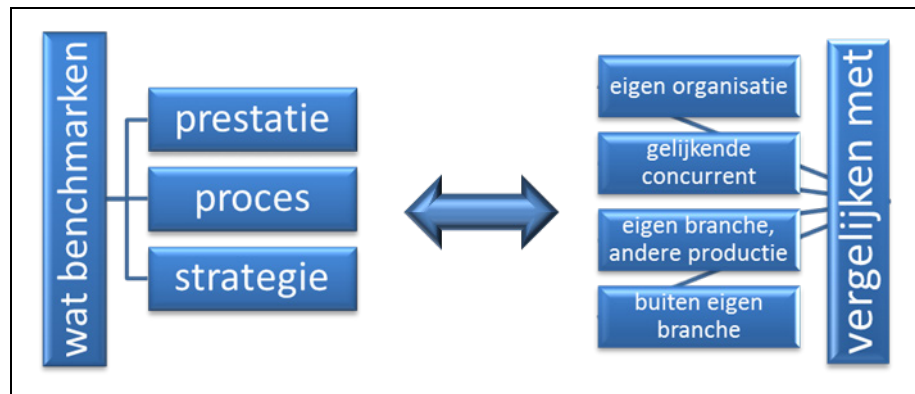
2.3.2. Momenteel veel gebruikte benchmarktypen en vergelijkingsgroepen

Drie typen benchmarks worden tegenwoordig veel gebruikt (Bhutta & Huq, 1999; Kyrö, 2003; Moriarty & Smallman, 2009), elk met een ander doel:

- *Prestatiebenchmarking*: benchmarken van prestatie-indicatoren (outcome- en output-indicatoren; zie *Afbeelding 2.3*) om te bepalen hoe goed een organisatie scoort ten opzichte van anderen en om mogelijkheden voor verbetering op te sporen door van anderen te leren.
- *Procesbenchmarking*: benchmarken van werkmethode en werkprocessen (throughput-indicatoren) om de werkprocessen in de eigen organisatie te verbeteren door van anderen te leren.
- *Strategische benchmarking*: benchmarken van de beleidsuitgangspunten van een organisatie (throughput-indicator) om de strategische koers van de eigen organisatie te veranderen.

Daarnaast onderscheiden de bovengenoemde auteurs vier mogelijke vergelijkingsgroepen waar een benchmarking zich op kan richten (zie ook *Afbeelding 2.1*):

- *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.



Afbeelding 2.1. Soorten benchmarks.

Het spreekt voor zich dat niet elk benchmarkdoel (vergelijken, verbeteren of strategie veranderen) zich even goed verhoudt 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 2.1*).

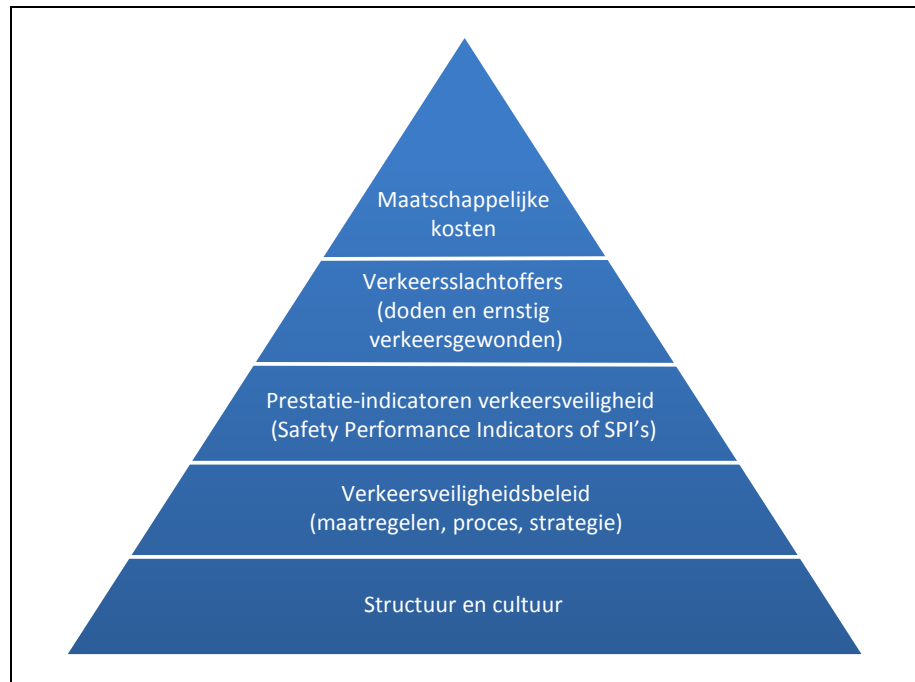
	Eigen organisatie	Gelijkende concurrent	Eigen branche, ander product	Buiten eigen branche
Prestatie	+/-	+	+/-	-
Proces	+/-	-	+	+
Strategie	-	+	-	-

Tabel 2.1. Typen benchmarks waarvoor de verschillende vergelijkingsmogelijkheden meer of minder geschikt zijn volgens Bhutta & Huq (1999).

2.3.3. Vertaling naar benchmarking van verkeersveiligheid

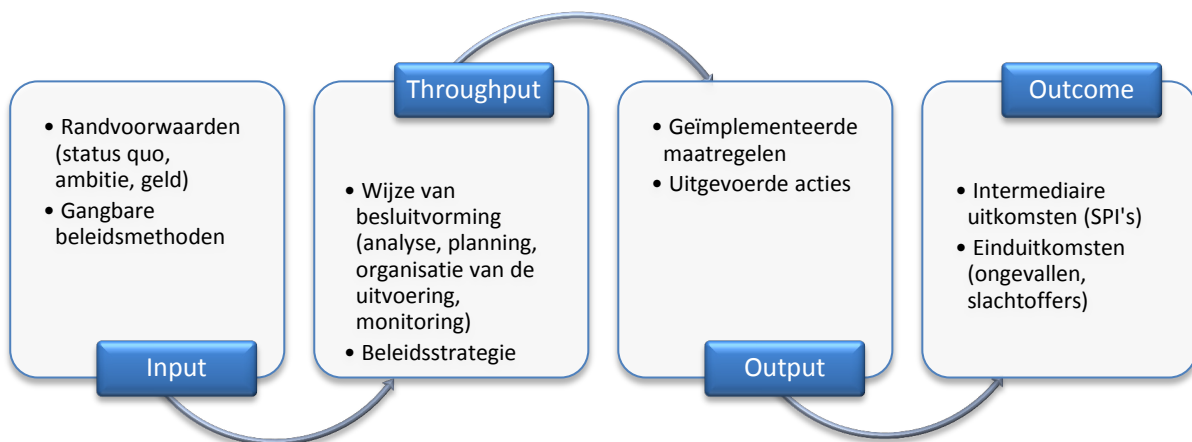
Deze benchmarktypen uit de bestuurskundige context kunnen we vertalen naar verkeersveiligheidsbenchmarks aan de hand van de verkeersveiligheidspiramide (zie *Afbeelding 2.2*) die vaak als theoretisch model wordt

gebruikt om de relatie tussen verkeersveiligheidsbeleid en -prestaties inzichtelijk te maken.



Afbeelding 2.2. Verkeersveiligheidspiramide (Koorstra et al., 2002; LTSA, 2000).

De elementen in deze piramide zijn grotendeels te koppelen aan het model van Input-Throughput-Output-Outcome, dat in de beleidswereld in het algemeen veel gebruikt wordt (zie Afbeelding 2.3). Zo kunnen we de maatschappelijke kosten, de verkeersslachtoffers en de safety performance-indicatoren (SPI's) beschouwen we 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).



Afbeelding 2.3. Beleidschema met daarin de voor verkeersveiligheid relevante onderdelen.

De verschillende stappen van het beleidsschema kunnen in de drie verschillende 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: ook wel bekend als 'Safety Performance Indicators' (SPI's) of kenmerken van gedrag en kwaliteit van het verkeerssysteem, zoals snelheid, alcohol, helmgebruik, kwaliteit van de weginrichting en dergelijke, die een sterke causale relatie hebben met het ontstaan van ongevallen of de ernst van de afloop ervan;
 - 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 ongevallenanalyse, 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.
3. De *strategische benchmark* bekijkt de hele strategie van een organisatie, in dit geval bijvoorbeeld de beleidsuitgangspunten om doelstellingen of ambities te verwezenlijken. In bovenstaand schema betreft dat ook een onderdeel van de 'throughput' van het beleidsschema. Vergelijking van deze strategieën met anderen kan aanleiding zijn de eigen strategie te veranderen.

Om te beoordelen of de te benchmarken organisaties vergelijkbaar zijn of met vergelijkbare omstandigheden te maken hebben, 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?

Als omstandigheden inderdaad vergelijkbaar zijn, kan competitieve benchmarking, bijvoorbeeld tussen plattelandsgemeenten of juist grote steden) worden uitgevoerd. Functionele benchmarking kan bijvoorbeeld worden uitgevoerd tussen verschillende afdelingen van gemeenten en algemene benchmarks zouden kunnen worden uitgevoerd door bijvoorbeeld een verkeersveiligheidsstrategie te vergelijken met een veiligheidsstrategie in de luchtvaart.

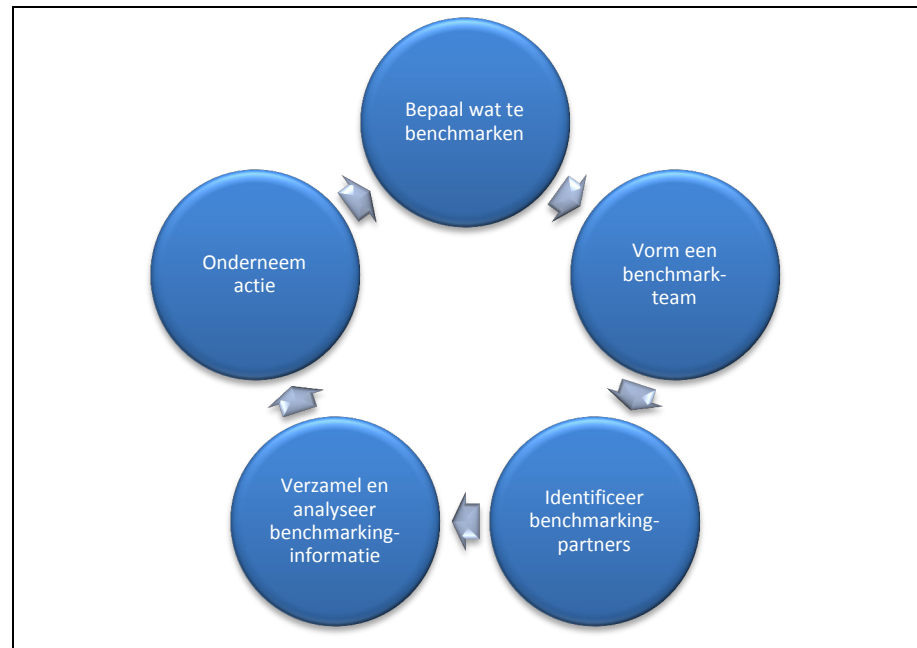
Hoewel de drie soorten benchmarks in oplopende mate verbeterpunten opleveren voor ingrijpen in de eigen organisatie, zijn ze ook in oplopende mate lastiger te operationaliseren: het is eenvoudiger om het aantal doden en gewonden te meten dan concreet te maken wat er gemeten moet worden aan een strategie.

2.4. Eisen aan benchmarking

Verskillende auteurs hebben algemene methoden voor benchmarks beschreven, stappen die daarvoor moeten worden doorlopen, en vereisten waaraan benchmarks zouden moeten voldoen.

2.4.1. Het stappenplan van Bhutta & Huq: techniek versus proces

Bhutta & Huq (1999) beschrijven benchmarking aan de hand van de volgende vijf stappen (zie *Afbeelding 2.4*):



Afbeelding 2.4. *Stappen in benchmarkproces, aangepast op basis van Bhutta & Huq (1999).*

Hoewel deze vijfstapsaanpak erg algemeen is, valt hierin wel een aantal zaken op.

Ten eerste zijn twee van de vijf stappen toebedeeld aan de *technische* aspecten van benchmarking (stap 1: bepaal wat te benchmarken en stap 4: verzamel en analyseer).

Ten tweede zijn drie van de vijf stappen gewijd aan de *interactieve* kanten van benchmarken: teams formeren en benchmarkpartners identificeren impliceert dat benchmarken niet iets is dat een organisatie alleen doet, maar samen met andere organisaties. Ook de laatste stap, actie ondernemen, is iets dat niet vanuit bijvoorbeeld een extern bureau of de wetenschap gedaan kan worden, maar waar de te benchmarken organisatie actief aan moet deelnemen.

Moriarty & Smallman (2009) stellen dan ook enigszins vanzelfsprekend dat benchmarking succesvoller zal zijn als deze wordt uitgevoerd door 'experienced implementers', mensen met beide benen in het werkveld.

2.4.2. *Benchmarkeisen volgens Korsten*

Ook Korsten et al. (2013) benadrukken het belang van het proces en maken dit nog iets concreter door zes criteria te benoemen die een goed gebruik van een benchmark volgens hen bevorderen:

1. eenvoudig te interpreteren benchmarkresultaten, bijvoorbeeld door gebruik van streefcijfers of andere normen;
2. gebruik van een aansprekend goed praktijkvoorbeeld;
3. formuleren van aanbevelingen in het eindrapport;
4. apart rapport per organisatie die aan de benchmark deelneemt;
5. voorafgaand overleg met de te benchmarken organisaties over de invulling van het onderzoek;
6. bestuurlijk draagvlak voor een benchmark.

Veel eisen die aan een goede benchmark gesteld moeten worden, zijn contextafhankelijk en specifiek voor het gekozen beleidsveld. De keuze voor specifieke indicatoren of de beoordeling van vergelijkbaarheid zal bijvoorbeeld sterk verschillen per beleidsterrein.

Korsten (2006; Korsten, et al., 2013) stelt dat alle technische details goed in orde moeten zijn. Dat betekent dat zowel de interne als de externe validiteit moet kloppen. Daarbij gaat het om vragen zoals:

- Zijn de juiste indicatoren gekozen?
- Worden de indicatoren op de juiste manier gemeten?
- Zijn de indicatoren tussen organisaties te vergelijken?
- Zijn de te benchmarken organisaties wel echt vergelijkbaar?

Indien niet aan deze eisen is voldaan, bestaat het gevaar dat de benchmark technisch incorrecte of vertekende antwoorden oplevert. Bovendien bestaat er ook een zogenaamd *perverteringsgevaar*: het kiezen van de verkeerde indicatoren kan aanzetten tot (in strategische zin) pervers gedrag. Dat is met name het geval met 'output-indicatoren' die slechts ten dele verband houden met de 'outcome-doelstellingen'. Een voorbeeld: indien stoelbezetting als een indicator voor het succes van een schouwburg is aangewezen, kan het effect zijn dat schouwburgen louter publiekstrekking programmeren. Indien de outcome-doelstelling breder is (bijvoorbeeld ook aandacht voor kleinschalig theater) meet de outputindicator niet alleen slechts een klein deel van de doelstelling, maar zorgt deze ook voor strategisch gedrag dat de doelstelling niet dichterbij haalt.

Een tweede reeks eisen aan benchmarking die Korsten (2006; Korsten, et al., 2013) formuleert is principiëler van aard. Benchmarking dient zich volgens hem altijd rekenschap te blijven geven van de complexiteit van de werkelijke situatie. Complexe prestaties zijn bijvoorbeeld vaak niet volledig in eenvoudige indicatoren te vangen. Bovendien kunnen complexe organisaties volgens Korsten niet goed vergeleken worden zonder hun context, netwerk en beleidsdoelstellingen mee te nemen: een benchmark op zich zegt immers niets over de achtergronden van de verschillen tussen organisaties. Een gevaar van de nadruk op context en uniciteit van organisaties heeft als ultieme consequentie dat geen enkele organisatie zich meer kan vergelijken met een andere, omdat elke organisatie zekere unieke eigenschappen zal hebben. 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).

2.5. Werkt benchmarking?

In de bestuurskundige literatuur is erg weinig onderzoek te vinden naar de effectiviteit of efficiëntie van benchmarking. Sterker nog, diverse auteurs (Dattakumar & Jagadeesh, 2003; Moriarty & Smallman, 2009; Yasin, 2002) concluderen dat de literatuur 'overwhelmingly pragmatic' is en dat 'iedereen' (in de bewoording van de auteurs) benchmarking eigenlijk wenselijk vindt zonder daarvoor al te veel onderbouwing aan te leveren.

Ook Korsten (2006) concludeert dat (nog) niet bewezen is dat benchmarks echt tot verbetermanagement leiden. Eén van de weinige onderzoeken naar de effectiviteit en efficiëntie van benchmarks is uitgevoerd in Nederland, onder 24 waterschappen. Van Helden & Tillema (2005) stellen dat in hun casussen 18 van de 24 waterschappen effectiever zijn gaan werken na toepassing van een benchmark. Echter, het is niet duidelijk of deze effectiviteitsslag te danken was aan het benchmarkinstrument, of aan de min of meer tegelijkertijd vanuit de overheid opgelegde strengere effectiviteits-eisen. Wat betreft de efficiëntie zijn geen duidelijke conclusies te trekken.

Ouder onderzoek van Coopers & Lybrand (1994, in Francis & Holloway, 2007) laat in een grootschalige enquête zien dat 75% van de geënquêteerde grote bedrijven hun benchmarkproject als positief beoordeelden.

Meer in het algemeen valt op dat in de verschillende soorten benchmarks en ook de definities die we in *Hoofdstuk 1* zagen, niet altijd 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.

2.6. Conclusie

Benchmarking is het vergelijken van activiteiten en/of prestaties van een organisatie met één of meer andere organisaties door het aanwijzen van een of meer goede praktijkvoorbeelden, en is gericht op verbetering door leren.

Benchmarking kan twee doelen hebben: het kan organisaties inzicht bieden in hun prestaties en mogelijke verbeterpunten; een benchmark kan ook gebruikt worden voor het afleggen van verantwoording. Er is een natuurlijke spanning tussen deze twee doelen, die kan worden voorkomen door organisaties aan te spreken op hun 'accountability for learning': hun vermogen om te leren.

Hoewel benchmarks sinds de jaren tachtig veel gebruikt worden in het bedrijfsleven, en in toenemende mate in de publieke sector, is er weinig onderzoek gedaan naar de effectiviteit of efficiëntie van benchmarks. De schaarse onderzoeken suggereren een positief effect. Een verbeterpunt is de onderbouwing van verklaringen voor succes.

2.6.1. *Drie mogelijke verkeersveiligheidsbenchmarks*

Voor een verkeersveiligheidsbenchmark zijn verschillende soorten benchmarks denkbaar.

1. De *prestatiebenchmark* geeft alleen een prestatiescore ten opzichte van anderen als basis voor leren en verbeteren. In de verkeersveiligheid kunnen prestaties gemeten worden in:
 - a. eindprestaties: aantallen doden en ernstig verkeersgewonden,
 - b. intermediaire prestaties: Safety Performance Indicators zoals snelheid, alcohol, helmgebruik, staat van de weg en dergelijke,
 - c. concrete maatregelen.

Uiteraard geeft zo'n benchmark niet enkel een lijst van scores, maar ook mogelijkheden om verbeteringen aan te brengen en van anderen te leren.

2. De *procesbenchmark* vergelijkt vooral de onderliggende werkprocessen, dus de organisatie van het beleid: zijn er budgetten? Zijn er doelstellingen opgesteld? Volgen maatregelen logisch uit een ongevallanalyse? Is er een monitor opgesteld? Dat soort vragen. Ook hier kan vergelijking met andere organisaties aanknopingspunten bieden voor verbeteringen.
3. Een *strategische benchmark* is een derde mogelijkheid. Hierbij kunnen strategische routes om de verkeersveiligheid te verbeteren worden vergeleken met strategieën in andere verkeersveiligheidsorganisaties of zelfs in andere (vervoers)domeinen. Dit type benchmark lijkt vooralsnog het lastigst te operationaliseren.

2.6.2. *Belang van het proces en van valide methoden bij benchmarken*

Een belangrijk punt bij benchmarking is dat het beoogde 'leren en verbeteren' alleen te bereiken is door organisaties actief te betrekken bij een benchmark, of hen de benchmark zelf te laten uitvoeren. De in de bestuurskundige literatuur beschreven stappenplannen voor het uitvoeren van een benchmark besteden dan ook meer aandacht aan interactieve processen dan aan de technische uitwerking. (Bestuurlijk) draagvlak om de benchmark uit te voeren wordt cruciaal genoemd.

Andere belangrijke voorwaarden voor een goede benutting van een benchmark zijn bijvoorbeeld eenvoudig te interpreteren en voldoende organisatiespecifieke resultaten en aanbevelingen. Daarnaast moet worden voldaan aan methodologische voorwaarden rondom interne en externe validiteit, maar moet ook de complexiteit van de werkelijke situatie in het oog worden gehouden bij het uitvoeren van een benchmark. Deze kan niet altijd volledig worden gevat in te meten indicatoren.

3. Onderzoek naar benchmarking in de wetenschappelijke verkeersveiligheidsliteratuur

Op het terrein van verkeersveiligheid, blijkt het merendeel van de studies over benchmarking zich te richten op het vergelijken van verkeersveiligheidsprestaties van landen of gebieden (zie bijvoorbeeld Bax et al., 2012; Breen, 2000; Hermans et al., 2009a; 2009b; Eksler, 2010; Wegman & Oppe, 2010), of op onderdelen van het beleid van landen of gebieden zoals de veiligheidscore van wegen (zie bijvoorbeeld Grove et al., 2002; Kimber, 2003; Klassen & Wahlström, 2005).

In de wetenschappelijke verkeersveiligheidsliteratuur worden voor benchmarking andere stappen onderscheiden dan in de bestuurskundige literatuur uit het vorige hoofdstuk (door bijvoorbeeld Bhutta & Huq, 1999). Het stappenplan voor benchmarking uit de wetenschappelijke verkeersveiligheidsliteratuur wordt in dit hoofdstuk als uitgangspunt genomen. De stappen worden het duidelijkst omschreven door Wegman & Oppe (2010):

1. identificeren van de belangrijkste componenten van verkeersveiligheidsprestaties (§ 3.1);
2. identificeren van de vergelijkingsgroep (landen, regio's, inclusief best presterende van de groep; § 3.2);
3. construeren van indicatoren om tot een betekenisvolle vergelijking te komen (§ 3.3);
4. vaststellen en verklaren van verschillen in de prestaties (§ 3.4);
5. vaststellen van na te streven prestaties in de toekomst (§ 3.5).

Aan de hand van dit stappenplan bespreken we in de achtereenvolgende paragrafen de literatuur over benchmarking van verkeersveiligheid. Deze literatuur is geselecteerd door te zoeken op 'benchmarking' en 'road safety' en gevoegd bij de literatuur over dit onderwerp die al bekend was bij de onderzoekers. Alleen die literatuur is vervolgens in beschouwing genomen die ingaat op het vergelijken van verkeersveiligheidsprestaties of -beleid van landen, gebieden of onderdelen daarvan. Literatuur waarin het begrip 'benchmarking' voorkwam maar bijvoorbeeld de vergelijking van twee verschillende soorten risicogedragingen behandelde, is buiten beschouwing gelaten.

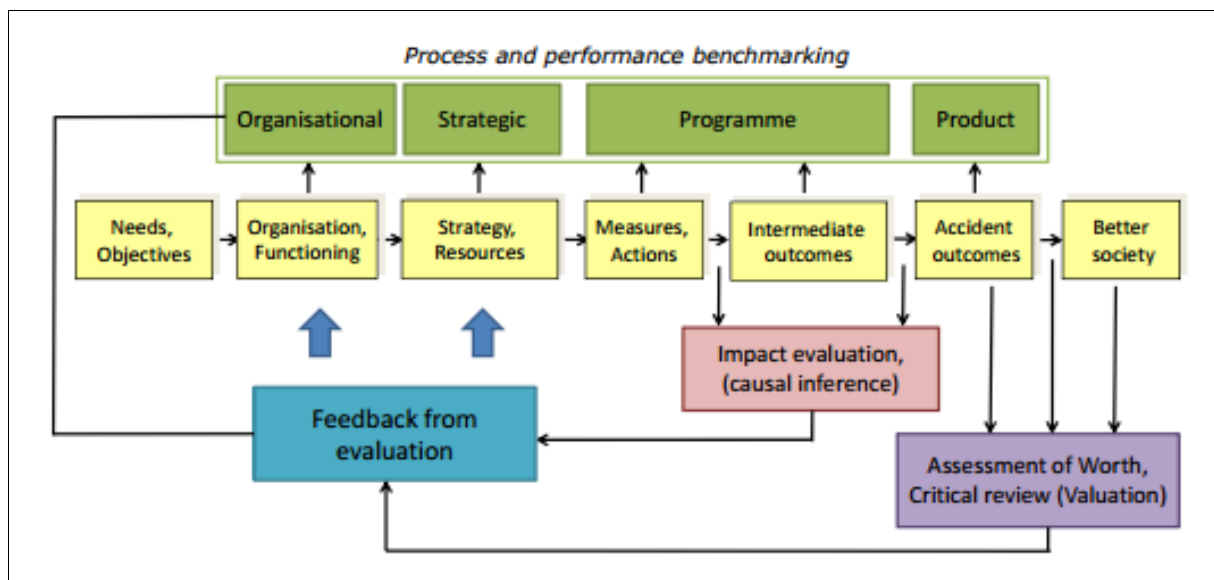
Aan het einde van het hoofdstuk brengen we de inzichten uit de bestuurskundige literatuur (*Hoofdstuk 2*) en die uit de wetenschappelijke verkeersveiligheidsliteratuur (dit hoofdstuk) bij elkaar om uiteindelijk mee te nemen in de uitwerking van een stappenplan voor verkeersveiligheidsbenchmarking binnen Nederland.

3.1. Identificeren van de belangrijkste componenten in verkeersveiligheidsprestaties

Een van de onderzoekers die binnen de wetenschappelijke verkeersveiligheidsliteratuur veel geschreven heeft over benchmarking is Vojtech Eksler. Als het gaat om het benoemen van belangrijke componenten van verkeersveiligheidsprestaties, onderscheidt Eksler (2009) de volgende:

- organisatie;
- strategie;
- programma;
- resultaten.

Eksler spreekt alleen over prestatie- en procesbenchmarks (zie ook het vorige hoofdstuk), en stelt vraagtekens bij het nut van strategiebenchmarks op het gebied van verkeersveiligheid. Volgens hem is benchmarken van strategie wellicht te beperkt, omdat het niet de implementatie van plannen volgend op de strategie in ogenschouw neemt, en daarmee te weinig zicht biedt op het succes in termen van resultaten. *Afbeelding 3.1* geeft de verschillende benchmarkonderdelen weer en hoe deze aansluiten bij de stappen in het verkeersveiligheidsbeleid volgens Eksler.



Afbeelding 3.1. Proces- en prestatiebenchmarking toegepast op verkeersveiligheidsbeleid volgens Eksler (2009).

Volgens Eksler is er tot nu toe vooral veel aandacht geweest voor het 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.

3.2. Identificeren van de vergelijkingsgroep

Wegman & Oppe (2010) stellen dat bij een vergelijking alleen sprake is van een zinnig referentiekader als er *meer dan twee gebieden vergeleken worden*. Er is dan een minimum aantal gebieden aanwezig 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 een groep (≥ 2) minder presterende gebieden.

Vervolgens is de vraag hoe die gebieden te definiëren. In eerder onderzoek van Wegman et al. (2008) wordt gepleit voor het definiëren van gebieden aan de hand van drie mogelijke indelingen:

- overheidsgrenzen (zoals landsgrenzen, provinciegrenzen of gemeentegrenzen);
- politiedistricten;
- gebieden met gelijke culturele normen en waarden.

Deze indelingen sluiten volgens Wegman en collega's zo goed mogelijk aan bij reeds bestaande indelingen, samenwerkingsverbanden en eenheden, en helpen volgens hen daarmee het potentiële draagvlak voor de uitkomsten van het onderzoek zo groot mogelijk te maken.

Er worden vervolgens twee manieren onderscheiden om gebieden te kiezen en te vergelijken (zie ook Knowles et al., 2010):

1. 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);
2. 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).

Om tot een goede benchmark te komen zijn gebieden idealiter vergelijkbaar, vooral omdat het anders lastig is vast te stellen of verschillen in prestaties niet ook aan onderliggende verschillen geweten kunnen worden. Deze vergelijkbaarheid is belangrijker naarmate de benchmarking meer op het proces achter de prestaties focust. Daarnaast dient echter ook opgemerkt te worden dat er zelfs bij vergelijkbare gebieden kleine verschillen zijn, waardoor het verstandig kan zijn prestaties te normeren.

In de volgende twee subparagrafen gaan we nog wat specifiek in op onderzoek dat illustratief is voor bovengenoemde twee sporen.

3.2.1. *Homogene gebieden vaststellen*

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 ervaren dat er meer interesse is om te leren van de best presterende als deze soortgelijke eigenschappen heeft als de te benchmarken gebieden.

De onderzoekers beschrijven diverse mogelijkheden om gebieden (met name landen) te groeperen, en wel op basis van:

- expertbeoordeling (kwalitatief);
- niveau van en ontwikkeling in verkeersveiligheid;
- verschillende kenmerken van het gebied die samenhangen met verkeersveiligheid.

In onderzoek van Houwing et al. (2012) naar het vaststellen van voor de verkeersveiligheid relevante homogene gebieden in Nederland, wordt geconcludeerd dat het *resultaat* van groepering (dat wil zeggen: welke gebieden tot één groep behoren) afhangt van:

- De gehanteerde gebiedsgrenzen. In het genoemde onderzoek werden gemeenten als eenheid genomen omdat dit Nederland opdeelde in voldoende kleine stukjes homogeen gebied (bijvoorbeeld stedelijk versus landelijk), maar aan de andere kant gaven gemeentegrenzen ook weer net te weinig onderscheid (kerngemeenten versus gemeenten met verspreid liggende kleine kernen; eventueel zou hier wijkniveau een uitkomst voor bieden, maar nadeel hiervan is het grote aantal gebieden dat dan weer ontstaat en ook het probleem dat niet alle gegevens op wijkniveau te koppelen zijn).
- De gebruikte gegevens. Werd alleen gekeken naar bijvoorbeeld inwonerdichtheid, dan kwamen er iets andere homogene groepen tevoorschijn 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.

Op basis hiervan concluderen Aarts & Houwing (te verschijnen) en Aarts & Bax (2013) dat het gekozen referentiegebied bij vergelijkend onderzoek het beste kan afhangen van de gestelde vraag. Dat wil zeggen dat als bijvoorbeeld het vergelijken van verkeersveiligheidsprestaties in het algemeen centraal staat, referentiegebieden gekozen worden op basis van algemene kenmerken die met verkeersveiligheidsprestaties samenhangen (bijvoorbeeld omvang van de bevolking). Als bijvoorbeeld een vergelijking ten aanzien van fietsveiligheid centraal zou staan, dan zouden referentiegebieden gekozen kunnen worden op basis van indicatoren die te maken hebben met fietsveiligheid (bijvoorbeeld de mate waarin aandacht aan fietsbeleid wordt besteed). Aarts & Bax constateren daarbij dat het ook belangrijk is dat de te benchmarken organisaties zichzelf ook in de vergelijkingsgebieden herkennen.

Eksler (zie bijvoorbeeld 2009; 2010) breekt een lans voor het benchmarken van sub-nationale gebieden. Dit gebeurt in zijn ogen nog niet of nauwelijks, dit in tegenstelling tot vergelijking van landen (in internationaal verband) of het vergelijken van meer gedetailleerde situaties aan een norm (denk aan de sterrensystemen EuroNCAP voor auto's en EuroRAP voor wegen). Het blijkt dat gebiedskenmerken die van invloed zijn op de verkeersveiligheid (zie de paragraaf over contextindicatoren, *Bijlage A*), meer verschillen tussen sub-nationale gebieden dan tussen landen (zie ook Eksler, 2009). In studies van Eksler en collegae (Eksler et al., 2008; Eksler & Lassare, 2008) is gevonden dat op het laagste disaggregatieniveau van gebieden (NUTS¹⁻³ niveau in dit geval), de meeste variantie in mortaliteit verklaard kon worden. Waarschijnlijk houdt dit verband met het feit dat een gebied over het algemeen homogener kenmerken zal hebben naarmate het kleiner is (zie ook Houwing et al., 2012).

¹ 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)

3.2.2. *Corrigeren voor verschillen*

Er zijn ook benchmarkstudies op het gebied van verkeersveiligheid die verkeersveiligheidsprestaties normeren of corrigeren voor verschillen tussen de te benchmarken gebieden en zo prestaties tussen gebieden vergelijken. Zo nemen Eksler et al. (2008) bijvoorbeeld belangrijke lokale invloedsfactoren mee als covariaat in het onderzoek naar verkeersveiligheidsprestaties tussen gebieden. Knowles et al. (2010) kiezen in hun benchmarkstudie van Noord-Ierland ook 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. Uiteraard hangt die norm af van de gebieden en data die in de modellering betrokken worden (zie ook vorige paragraaf). Knowles et al. namen bijvoorbeeld diverse expositiematen, zoals het aantal voertuigen per aantal inwoners en de verkeersprestatie voor vijf typen wegen. Echter, van sommige geconstateerde verschillen tussen de vergeleken landen, bleken onvoldoende data beschikbaar om ook hiervoor te corrigeren.

Een dergelijke modelleringsaanpak is vooral geschikt om het effect van kenmerken op de verkeersveiligheid vast te stellen en onderbouwing te leveren van de invloed van succesfactoren. Een dergelijke aanpak lijkt 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 prestatiematen op, wat zich daardoor waarschijnlijk minder goed leent voor een analyse- en leerproces met organisaties zoals decentrale overheden.

3.3. **Construeren van indicatoren om te vergelijken**

Benchmarking is onlosmakelijk verbonden met het definiëren en vergelijken van waardevolle indicatoren (zie onder andere Eksler, 2009; Eksler et al., 2009; Wegman et al., 2009; Wegman & Oppe, 2010). Een indicator kan worden gedefinieerd als een kwantitatieve of kwalitatieve maat die is afgeleid van geobserveerde feiten om de relatieve toestand van een onderwerp vast te kunnen stellen (zie Nardo et al., 2005). Wat opvalt als we de wetenschappelijke verkeersveiligheidsliteratuur over benchmarking bekijken, is dat de nadruk sterk ligt op het construeren van ingenieuze samengestelde indicatoren. In deze derde stap die Wegman en Oppe onderscheiden (Construeren van indicatoren om te vergelijken), gaat het dus vooral om het bewerken van gegevens tot (samengestelde) indicatoren waarop de gekozen gebieden uiteindelijk vergeleken gaan worden. De volgende subparagrafen beschrijven onderzoek op dat punt.

3.3.1. *Verkeersveiligheidsindicatoren*

Zagen we in de bestuurskundige literatuur (*Hoofdstuk 2*) een indeling van indicatoren of type benchmark (prestatie, proces of strategie), in de wetenschappelijke verkeersveiligheidsliteratuur worden weer net iets andere indicatoren onderscheiden. Zo definiëren Wegman en collegae (2009) drie typen indicatoren op basis waarvan landen of andere jurisdicties vergeleken kunnen worden:

1. verkeersveiligheidsprestatie-indicatoren;
2. implementatieprestatie-indicatoren;
3. beleidsprestatie-indicatoren.

Verkeersveiligheidsprestatie-indicatoren kunnen bestaan uit eindindicatoren, zoals ongevallen en slachtoffers, maar ook uit intermediaire indicatoren of prestatie-indicatoren voor verkeersveiligheid zoals onveilig gedrag en de systeemtoestand (ook wel bekend als Safety Performance Indicators of SPI's; zie ook *Afbeelding 2.2*). Deze tweede groep van indicatoren kan organisaties belangrijke inzichten bieden in onveilige aspecten van hun wegverkeer, mits goed naar de causale relaties van die indicatoren met ongevallen wordt gekeken en ook goede data beschikbaar zijn (zie ook Bax et al., 2012; Gitelman et al., 2010).

Implementatieprestatie-indicatoren zijn voor de verkeersveiligheid niet of nauwelijks ontwikkeld. Het kan daarbij bijvoorbeeld gaan om indicatoren die aangeven welke en hoeveel effectieve maatregelen er door beleid in hoeveel tijd geïmplementeerd zijn. Dit zou echter ook opgevat kunnen worden als een onderdeel van beleidsprestatie-indicatoren, immers, beleid is pas af als plannen en strategieën daadwerkelijk tot maatregelen in de praktijk hebben geleid.

Beleidsprestatie-indicatoren kunnen bestaan uit maatregelen (zie ook implementatieprestatie-indicatoren), het proces om tot een selectie van maatregelen te komen, het hebben en formuleren van doelstellingen, budget, evaluatie en samenwerking.

In *Hoofdstuk 2* en ook § 3.1 hebben we gezien dat verkeersveiligheidsprestatie-indicatoren gebruikt kunnen worden in prestatiebenchmarks en implementatie- en beleidsprestatie-indicatoren zouden relevant zijn in een procesbenchmark en de laatste eventueel ook in een strategische benchmark.

Los van deze indicatoren worden ook achtergrondkenmerken als belangrijke variabelen genoemd (de structuur- en cultuurlaag van *Afbeelding 2.2*). Wegman en collega's (Wegman et al., 2009; Wegman & Oppe, 2010) benadrukken het belang van deze indicatoren om prestaties mee te *normeren*. Ze kunnen ook gebruikt worden om *homogene groepen* gebieden samen te stellen (zie Houwing et al., 2012; Aarts & Bax, 2013; zie § 3.2).

Meer achtergrondbeschouwingen over indicatoren op onderdelen van de verkeersveiligheidspiramide (zie *Afbeelding 2.2*) zijn te vinden in *Bijlage A*. Dit is echter een algemeen overzicht en zeker niet uitputtend. Welke indicatoren in een benchmark uiteindelijk worden betrokken, hangt sterk af van het onderwerp dat in de benchmark centraal staat en hangt ook samen met beschikbaarheid van gegevens en de kwaliteit van die gegevens. Als bijvoorbeeld een paar gemeenten een benchmark willen uitvoeren naar hun verkeersveiligheidsprestaties in het algemeen, dan vergt dat andere – algemenere – indicatoren, dan als een specifiek onderwerp zoals bijvoorbeeld fietsveiligheid of de veiligheid van gebiedsontsluitingswegen in die gemeenten centraal staat. Als hiervoor idealiter bepaalde indicatoren nodig zijn waarvan geen gegevens beschikbaar zijn, dan zal bijvoorbeeld gewerkt moeten worden met surrogaatmaten of zullen bepaalde indicatoren buiten beschouwing moeten worden gelaten.

3.3.2. Methode-ontwikkeling

Op het gebied van verkeersveiligheid is er al behoorlijk wat benchmark-literatuur met aandacht voor het construeren van waardevolle indicatoren. Op dit onderwerp is bijvoorbeeld de Universiteit van Hasselt erg actief. Zo hebben Hermans en collega's vooral onderzoek gedaan naar de Data Envelopment Analysis- ofwel DEA-methode, die ze geschikt hebben gemaakt voor het vergelijken van verkeersveiligheidsprestaties van verschillende landen (zie onder andere Hermans et al., 2008; 2009a; 2009b). De DEA-techniek werd gekozen uit diverse andere datawegings-technieken omdat deze de beste resultaten gaf bij het indelen van landen op basis van mortaliteit.

Een voordeel van de DEA-methode is dat het de combinatie van diverse prestatie-elementen op het gebied van de verkeersveiligheid tot één indicator mogelijk maakt. Binnen die gecombineerde indicator kunnen per organisatie (in dit geval landen) toch ook nog de losse indicatorelementen worden aangewezen waarop progressie mogelijk is en kan per indicatorelement bepaald worden welke landen als benchmark kunnen fungeren.

Los van een aantal subjectieve elementen waarvan de methode gebruikmaakt, kan in het licht van *Hoofdstuk 2* als belangrijkste nadeel worden genoemd dat de methode tamelijk technisch van aard is, en daarmee minder goed toegankelijk voor beleidsmakers. Niet alleen vanuit de bestuurskundige hoek wordt het belang van aanspreekbare resultaten benadrukt. Ook binnen de verkeersveiligheid zijn hiervoor indicaties 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; het is echter belangrijk om hierin een evenwicht te vinden.

3.4. Vaststellen en verklaren van verschillen in prestaties

In de verkeersveiligheidsbenchmarks die zijn uitgevoerd, overheersen internationale vergelijkingen. Met name de SUNflower-studies (Koorstra et al., 2002; Wegman et al., 2005; Wegman et al., 2008) zijn hierin bekende en belangrijke voorbeelden. In deze studies is de verkeersveiligheidspiramide uit *Afbeelding 2.2* als basis gebruikt om verklaringen te vinden voor afwijkingen in verkeersveiligheidsprestaties. In eerste instantie werd echter gekeken wat de beste praktijkvoorbeelden (De SUN-landen: Zweden, het Verenigd Koninkrijk en Nederland) aan eigenschappen gemeenschappelijk hadden die als verklaringen voor hun succes konden dienen. In de eerste SUN-flowerstudie werd vastgesteld dat de drie landen overeenkwamen in het jarenlang vasthoudend en planmatig werken aan verbeteringen, op overeenkomstige beleidsthema's. Op meer detailniveau bleken ze toch ieder hun eigen accenten te leggen, niet verwonderlijk gezien de verschillen die er op detailniveau tussen de landen blijken te bestaan. Ook hierbij speelt weer het in § 3.2.1 genoemde punt van homogeniteit en vergelijkbaarheid versus verschillen.

In veel benchmarkstudies op gebied van verkeersveiligheid worden eerste verklaringen voor verschillen in prestaties tussen gebieden (meestal landen) vooral gezocht in verschillen in SPI's (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 slechte prestaties op het gebied van SPI's (bijvoorbeeld: er wordt harder gereden dan in andere gebieden, of de gordel wordt minder (goed) gedragen, er wordt minder gehandhaafd of de infrastructuur is minder veilig ingericht). In deze studies worden in een enkel geval wel pogingen ondernomen om ook de link te leggen met successen van beleid (secundaire verklaring voor zowel SPI-prestaties als prestaties op het gebied van verkeersslachtoffers), maar met name door gebrek aan evidentie voor de mogelijke succesfactoren, blijven echte verklaringen uit (zie bijvoorbeeld Bax et al., 2012).

De European Transport Safety Council (ETSC) publiceert regelmatig vooral praktisch opgestelde benchmarks in de vorm van vergelijkingen om de verkeersveiligheid in Europese landen te stimuleren (de zogenoemde PIN-reports; zie bijvoorbeeld Jost et al., 2013). In deze voorbeelden wordt niet naar achterliggende redenen voor verschillen in prestaties gezocht, maar verklaringen worden expliciet aan de landen zelf overgelaten. De landen die jaarlijks in het zonnetje worden gezet omdat ze volgens de ETSC-benchmark een uitzonderlijke progressie hebben geboekt, worden vooral op politiek niveau geïnterviewd en naar succesfactoren gevraagd. In hoeverre deze interviews onderbouwde inzichten geven in daadwerkelijke succesfactoren is echter onbekend.

De conclusie lijkt ook hier dus te zijn dat, ondanks de aandacht die er expliciet is voor evidentie van succes, het een lastig punt blijft om daadwerkelijk te onderzoeken en aannemelijk te maken welke factoren aan succes hebben bijgedragen. Wat hierbij ook opvalt is dat in de verkeersveiligheidsliteratuur het benchmarken vooral technisch gericht is, waarbij de uitvoering zich concentreert rond de zoektocht van onderzoekers naar indicatoren, criteria en verklaringen. Er zijn geen voorbeelden aangetroffen waarbij in dit proces ook beleidsmakers of andere verantwoordelijke actoren betrokken zijn. We concluderen hieruit dat de verkeersveiligheidsliteratuur over het thema benchmarking een ander accent legt dan de bestuurskundige literatuur. Verklaringen hiervoor liggen in verschillende vraagstellingen die aan het onderzoek ten grondslag liggen.

3.5. Vaststellen van na te streven prestaties in de toekomst

In het verlengde van vorige paragraaf constateren we dat ook de vijfde stap in verkeersveiligheidsbenchmarks (het vaststellen van na te streven prestaties in de toekomst) vooral aandacht heeft gekregen in technische zin en in termen van SPI's (risicofactoren) en methodeontwikkeling (zie bijvoorbeeld Hermans et al., 2009b; Wegman et al., 2008). Op basis van technische methoden zoals DEA zijn onderzoekers zelfs in staat gebleken heel precies te definiëren – weliswaar gegeven de beschikbare data – wat op SPI-niveau benchmarkwaarden voor verschillende landen zijn om hun prestaties op slachtofferniveau te verbeteren (Hermans et al., 2009b). Het is onbekend in hoeverre dergelijke bevindingen hebben geleid tot praktische toepassingen door beleidsmakers.

3.6. Conclusies

In de wetenschappelijke verkeersveiligheidsliteratuur over benchmarken wordt een aantal stappen onderscheiden. Deze zijn samengevat (als in Wegman & Oppe, 2010):

1. identificeren van de belangrijkste componenten in verkeersveiligheidsprestaties (indicator(en));
2. identificeren van de vergelijkingsgroep;
3. construeren van indicatoren om tot een betekenisvolle vergelijking te komen;
4. vaststellen en verklaren van verschillen in de prestaties;
5. vaststellen van na te streven prestaties in de toekomst.

Wat hierbij opvalt is het verschil met de benchmarkstappen die zijn onderscheiden in de bestuurskundige literatuur uit het vorige hoofdstuk (bijvoorbeeld in *Afbeelding 2.4* door Bhutta & Hug, 1999). Bij deze laatste ligt het accent veel meer op processtappen, bij Wegman & Oppe ligt het accent juist zeer sterk op de technisch te nemen stappen. Dit accent zien we ook terug in de wetenschappelijke verkeersveiligheidsliteratuur over benchmarken. Met name de technische uitwerking van (samengestelde) indicatoren, het onderzoek naar verklaringen voor verschillen in eindprestaties door te kijken naar de relatie met tussenprestaties (SPI's), en de methode-ontwikkeling hebben daarbij veel aandacht gehad.

Als het doel van benchmarks inderdaad is om te leren van andere goede praktijkvoorbeelden, zullen benchmarkingonderzoek en –uitkomsten impact op het beleid moeten kunnen hebben. Hoe die impact te bewerkstelligen is dan een vraag van betekenis. In de bestuurskundige literatuur is daarover al het een en ander gezegd; zie § 2.4 over eisen aan benchmarking. Deze eisen of aanbevelingen kunnen we als waardevolle aanvullingen zien bij het benchmarken van verkeersveiligheid bij decentrale overheden.

4. Uitwerkingen van benchmarking in de praktijk

Naast wetenschappelijke literatuur, hebben we ook praktijkvoorbeelden op het gebied van benchmarking beschouwd. We hebben voorbeelden van benchmarking beschreven die bij SWOV reeds bekend waren en daarnaast op internet verder gezocht naar benchmarkingsvoorbeelden binnen de sector verkeer en vervoer en specifiek binnen verkeersveiligheid. Dit hoofdstuk geeft daarvan een overzicht. Dit is niet per se uitputtend, maar geeft een aantal voorbeelden van uitwerkingen in de praktijk bij partijen die voor verkeersveiligheid en decentrale overheden relevant zijn.

4.1. Werkwijze KING: stimulering en kwaliteitsbewaking van benchmarking

Het Kwaliteitsinstituut Nederlandse Gemeenten (KING) is gelieerd aan de Vereniging van Nederlandse Gemeenten (VNG) en adviseert bij verbetering van de kwaliteit van gemeenten. Benchmarking is een van de instrumenten die het KING promoot om de prestaties en ambities van gemeenten te bevorderen (www.kinggemeenten.nl). Het KING verstaat onder benchmarking meer dan alleen vergelijken van prestatiecijfers, namelijk ook een onderzoek in de diepte naar de redenen waarom een gemeente minder presteert dan vergelijkbare andere (zie ook *Afbeelding 4.1*). Benchmarking kan daarmee een hulpmiddel zijn om van anderen te leren en niet zelf het wiel uit te hoeven vinden.

WAARSTAATJEGEMEENTE.NL onderdeel van **KING** KWALITEITS INSTITUUT NEDERLANDSE GEMEENTEN

HOME | **BURGERPERSPECTIEF** | VERKLARINGSMODELLEN | CIJFERS | BENCHMARKS | TOEZICHTINFORMATIE | MONITORS

VERGELIJKEN, VERBAZEN EN VERBETEREN!

BLIK VAN DE SAMENLEVING Wat vindt de burger van uw gemeente? Lees meer >>	VERKLARINGSMODELLEN Welke kenmerken beïnvloeden de prestaties van uw gemeente? Lees meer >>
CIJFERS Hoe presteert uw gemeente ten opzichte van collega-gemeenten? Lees meer >>	BENCHMARKS Is uw gemeente op zoek naar een verdiepende benchmark? Lees meer >>

Over Waarstaatjegemeente.nl
Gemeenten kunnen op Waarstaatjegemeente.nl hun prestaties vergelijken met die van collega-gemeenten. Zo ontdekken gemeenten op welke maatschappelijke thema's ze zich nog kunnen verbeteren. En van welke gemeenten ze kunnen leren. [Lees meer >](#)

STEL HIER UW VRAAG

Afbeelding 4.1. Instrumentenpaneel van *waarstaatjegemeente.nl* als onderdeel van het Kwaliteitsinstituut Nederlandse Gemeenten.

Benchmarking zelf wordt geïnitieerd, betaald en bepaald door gemeenten. Het KING vervult drie rollen als het gaat om benchmarking.

1. Promoten van benchmarkstudies: op een openbare website worden diverse prestatiegegevens van gemeenten getoond in relatie tot verschillende referenties (dat wil zeggen gemiddelde score van: de provincie, van vergelijkbare gemeenten en van Nederland; zie *Bijlage A* voor meer informatie). Deze prestatiegegevens zijn bedoeld om

- gemeenten te prikkelen om uit te zoeken wat de achtergronden zijn van hun prestaties en wat ze kunnen leren van andere (vergelijkbare) gemeenten die beter presteren.
2. KING-keurmerk benchmarks: dit keurmerk wordt uitgereikt aan benchmarks die aan door het KING gestelde kwaliteitseisen voldoen (zie *Bijlage B*). Het keurmerk is vooral bedoeld om gemeenten te helpen bij het bepalen van een keuze als ze een benchmark willen uitvoeren en om hen zekerheid te bieden dat de gecertificeerde instantie aan bepaalde basiseisen voldoet. De vraagformulering en opdrachtverstrekking liggen verder geheel bij de gemeenten zelf.
 3. Makelaarsfunctie: benchmarks die het KING-keurmerk hebben, worden thematisch geordend openbaar gemaakt. Deze openbaarmaking houdt in dat bekend wordt gemaakt wat de vraagstelling is van de benchmark, wie hem aanbiedt, wat de kosten zijn en hoeveel gemeenten deelnemen. De inhoud blijft tussen de gemeenten die eraan deelnemen om zo een veilige leeromgeving te bieden (zie ook *Hoofdstuk 2*). Tevens bundelt het KING verzoeken van gemeenten voor benchmarks en aanbiedingen voor benchmarks en brengt het vraag en aanbod zo mogelijk met elkaar in contact.

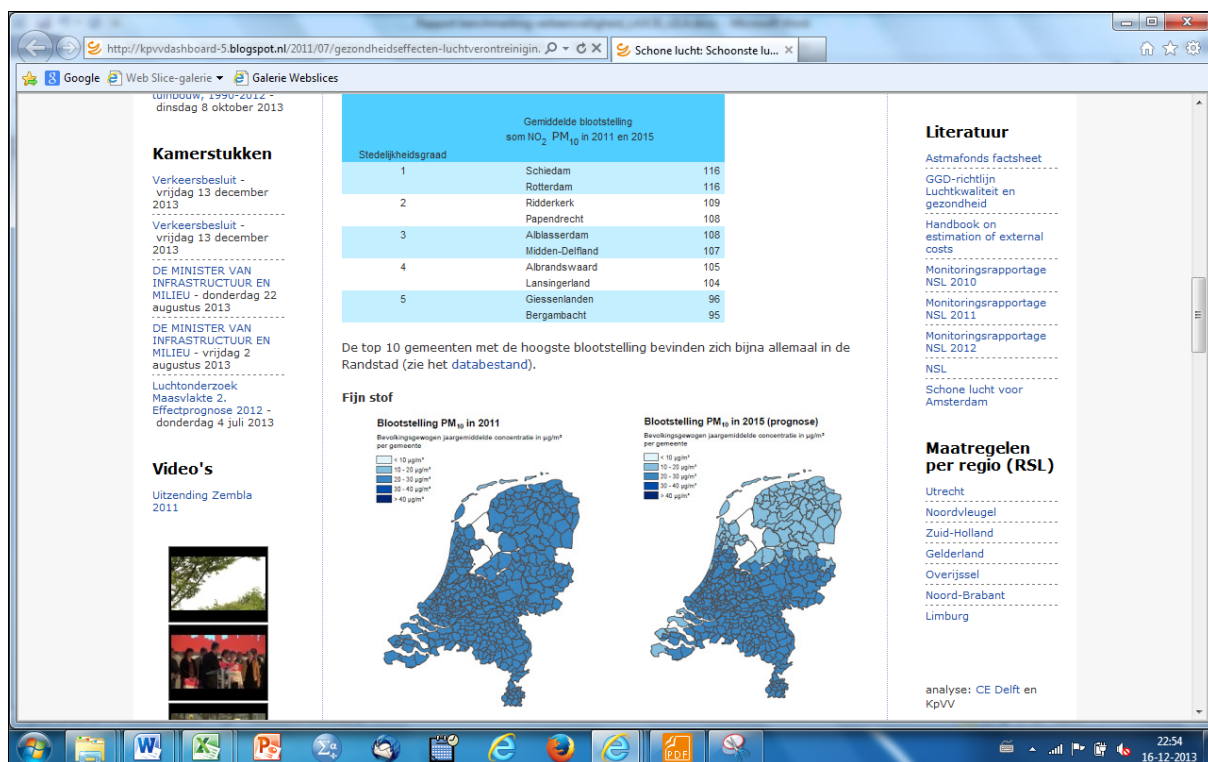
Voor verkeersveiligheid (en andere verkeer-en-vervoersonderwerpen) bestaan momenteel geen benchmarks met het KING-keurmerk. Algemeen bestuurlijke onderwerpen, maatschappelijke ondersteuning, publieke gezondheid en milieu zijn thema's waarbij de meeste benchmarks (met keurmerk) te vinden zijn. Wel presenteert KING op waarstaatjegemeente.nl gegevens over verkeersveiligheid. Dit zijn meningen van burgers over de verkeersveiligheid in de gemeente en de eigen buurt en eenvoudige (geregistreerde) verkeersveiligheidscijfers (zie voor meer details *Bijlage B*). Sinds kort is het KING ook samen met Atlas.nl bezig algemene verklarende modellen te ontwikkelen om gemeenten naast de cijfers en burgerpeilingen ook meer algemene achtergrondinformatie te verschaffen over de samenhang van verschillende indicatoren. Voor verkeer-en-vervoerthema's en dus ook voor verkeersveiligheid, hebben de genoemde partijen een dergelijk model (nog) niet ontwikkeld.

4.2. Duurzame mobiliteit: KpVV

Het Kennisplatform Verkeer en Vervoer (KpVV) heeft tot doel als kennis-makelaar voor decentrale overheden te fungeren. Een van de middelen die KpVV daarvoor gebruikt, is een vergelijkingsinstrument voor 'duurzame en slimme mobiliteit'. KpVV heeft daartoe samen met bureau Goudappel-Coffeng een internetapplicatie ('dashboard') ontwikkeld waarin gebruik wordt gemaakt van vergelijkingen of benchmarking (Korver & De Haan, 2012; zie bijvoorbeeld *Afbeelding 4.2*, <http://kpvvdashboard.blogspot.nl/> of <http://www.kpvv.nl/Dashboard-duurzame-en-slimme-mobiliteit-Schone-lucht>). De doelgroep bestaat uit gemeenten.

De aanleiding voor dit vergelijkingsinstrument was de ervaring dat veel gemeenten niet (goed) op de hoogte zijn van duurzaamheidsgegevens over hun eigen beheergebied. Naast allerlei achtergrondinformatie over duurzaamheid, toont de applicatie ook de prestatie op een aantal duurzaamheidsindicatoren van alle Nederlandse gemeenten in één kaart. Ook de brondata hiervan worden vrij beschikbaar gesteld. Expliciet worden de best en slechtst presterende gemeenten genoemd (top 10-lijsten). Door deze

gegevens in één overzichtelijke internetomgeving beschikbaar te stellen, kunnen gemeenten hun eigen situatie vergelijken met die van andere gemeenten.



Afbeelding 4.2. Voorbeeld van het dashboard duurzame en slimme mobiliteit van KpVV.

Gemeenten worden in deze vergelijkingsoverzichten niet op één hoop gegooid, maar worden gegroepeerd op basis van stedelijkheidsgraad. Hiervoor worden vijf klassen gebruikt: van sterk stedelijk tot sterk ruraal gebied. Voor elk van deze klassen zijn gemiddelde waarden beschikbaar. De uiteindelijke benchmarkgegevens worden weergegeven per hoofd van de bevolking. Ook worden gegevens met elkaar gecombineerd en als verklarende factoren gepresenteerd (het aandeel bewegingsarmoede per provincie wordt bijvoorbeeld in verband gebracht met de gemiddelde duur dat mensen in de betreffende provincie fietsen). Ook wordt achtergronddocumentatie verstrekt.

De ambitie van het KpVV is om iedere maand een nieuw item in de applicatie gereed te hebben. Gemeenten kunnen hiervan op de hoogte worden gehouden via de gebruikelijke KpVV-berichten en via een RSS-feed.

Voorbeelden van onderwerpen die binnen de applicatie aan bod komen zijn:

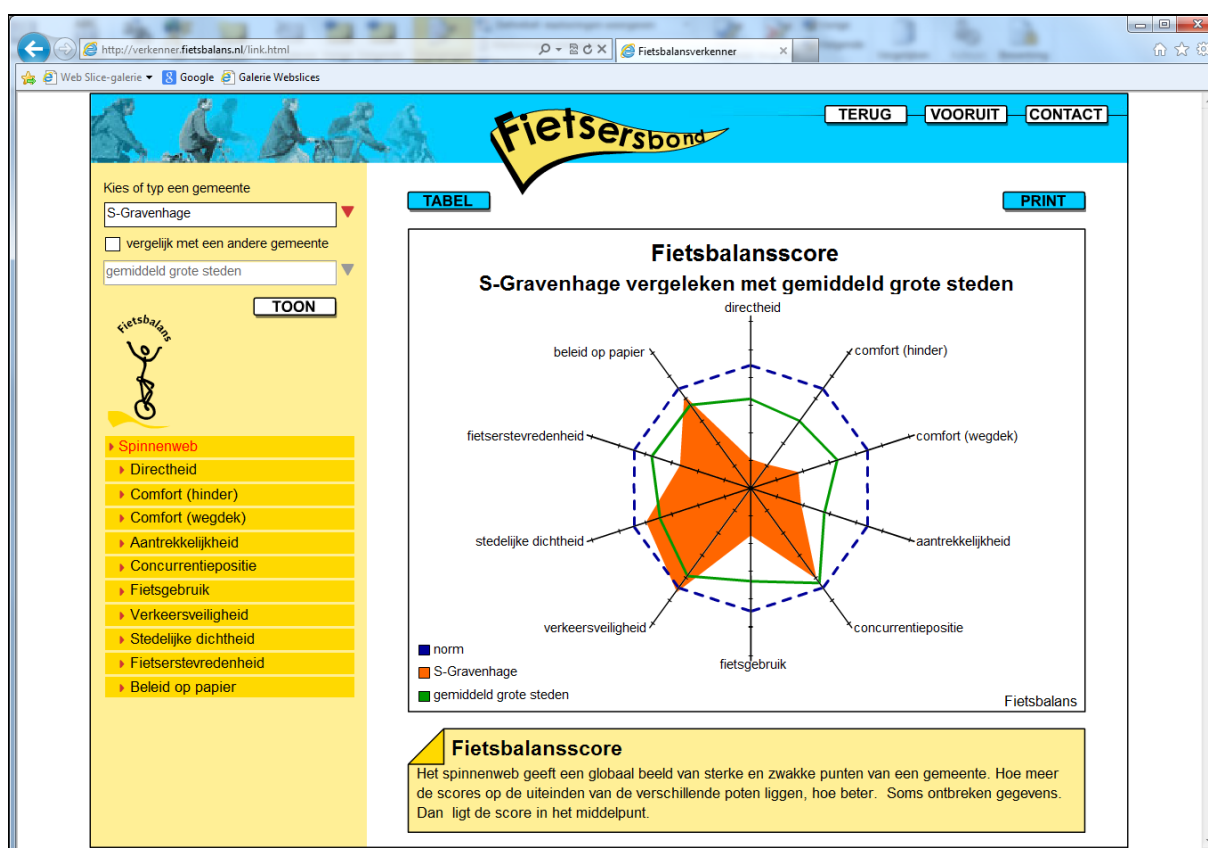
- schone vervoerswijzen, zoals ov, lopen en fietsen;
- luchtkwaliteit;
- slimme mobiliteit, zoals autodelen, slim werken en reizen;
- gezondheid;
- energie.

Verkeersveiligheid is (nog) geen onderwerp in de applicatie.

KpVV heeft een evaluatie gehouden naar de bekendheid en het gebruik van het dashboard. Uit een steekproef van 98 gemeenten bleek ruim 80% het dashboard te kennen (verkeer en vervoer, milieu en ruimtelijke ordening). Van de geïnterviewden geeft ruim 40% aan dankzij het dashboard meer aan duurzame mobiliteit te werken.

4.3. Fietsbalans van de Fietsersbond

De Fietsbalans van de Fietsersbond is ook een vorm van een benchmark omdat hierin gegevens van organisaties (in dit geval gemeenten) worden beschouwd binnen een of meer referentiekaders. De eerste versie van deze Fietsbalans is in de periode 2000-2004 toegepast op 125 gemeenten, en de resultaten daarvan zijn in te zien op een website waarbij gebruikers zelf de metingen van een gemeente kunnen vergelijken met die van andere gemeenten (zie www.fietsbalans.nl/verkenner; zie Afbeelding 4.3).



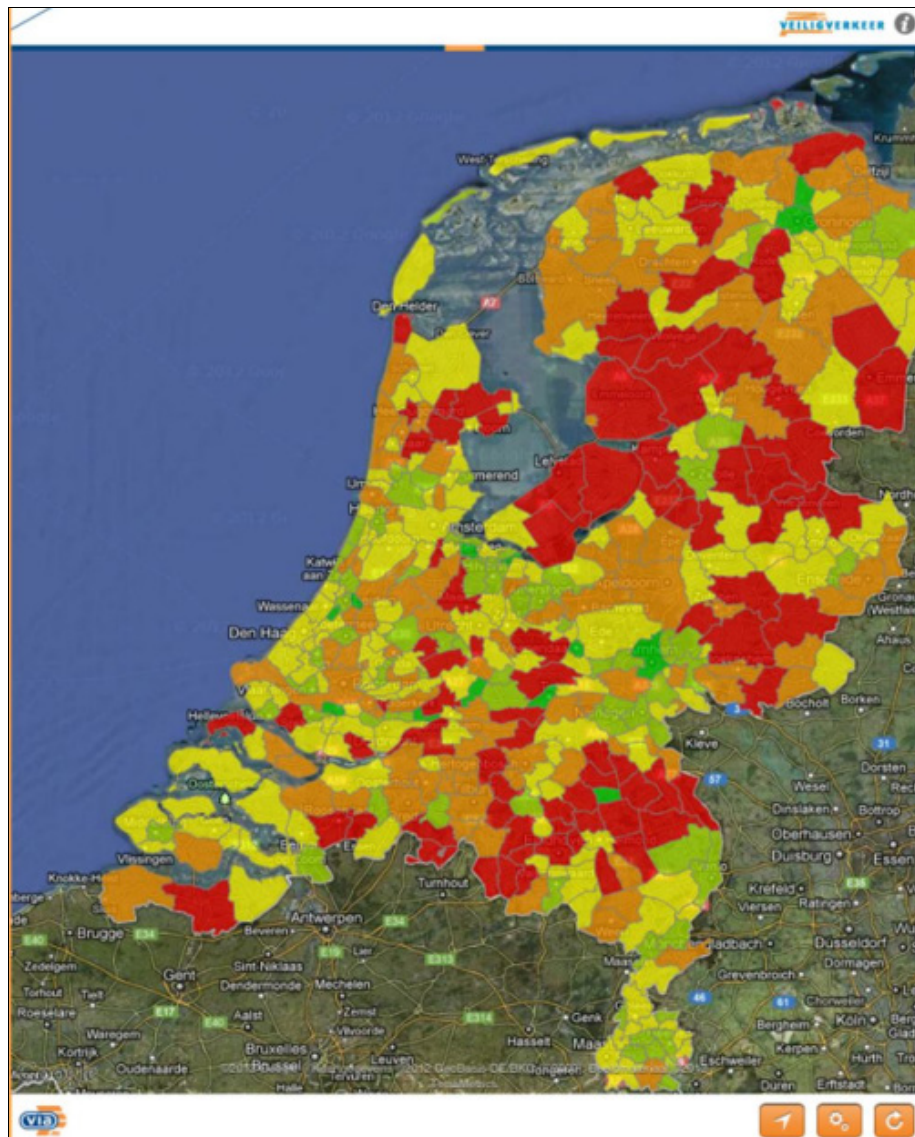
Afbeelding 4.3. Voorbeeld van een resultaat uit de eerste Fietsbalans, waarin gemeenten hun metingen kunnen vergelijken met die van een andere gemeente of een groep andere gemeenten.

Door voor verschillende gemeenten telkens dezelfde indicatoren te meten, konden de prestaties van die gemeenten op die indicatoren met elkaar vergeleken worden. Na 2004 is een tweede versie van de Fietsbalans ontwikkeld, die uiteindelijk in ruim 50 gemeenten is toegepast. De uitwerking van deze metingen zijn terug te vinden in specifiek voor deze gemeenten opgestelde rapporten, waarin ook verschillende vergelijkingen zijn gemaakt met andere gemeenten uit de tweede Fietsbalans. Om praktische redenen zijn deze metingen echter gestaakt. De uitwerking (zie Afbeelding 4.3) kan interessant zijn voor gebruik in een benchmark verkeersveiligheid, vooral

ook op onderdelen zoals fietsveiligheid. Ook zijn de ervaringen met de gegevens en het inwinnen van deze gegevens belangrijk om mee te nemen in een benchmark die uiteindelijk iets kan zeggen over de prestatie van gemeenten op het gebied van fietsbeleid.

4.4. ViaStat Dashboard

VIA.nl is een bureau dat op het gebied van verkeer en vervoer applicaties ontwikkelt. De afnemers daarvan zijn vooral decentrale overheden. Een van de producten die VIA.nl heeft ontwikkeld en waarbij expliciet gebruik wordt gemaakt van vergelijking of benchmarking tussen gemeenten, is het ViaStat Dashboard, een app over gemeentelijke verkeersveiligheid en bereikbaarheid (zie Donkers & Coppen, 2012). Per gemeente wordt inzichtelijk gemaakt hoe men scoort op verkeersveiligheid (doden en ernstig verkeersgewonden 2007-2009 gerelateerd aan het aantal inwoners; zie *Afbeelding 4.4*).



Afbeelding 4.4. Voorbeeld van de presentatie van gemeentelijke verkeersveiligheid in het ViaStat Dashboard.

Daarnaast geeft de app voor aangesloten gemeenten de mogelijkheid om meer detailinformatie op te vragen over onderliggende indicatoren, zoals vervoerswijze, leeftijd en locatie, en tendensen door de tijd heen, waarbij de score van de gemeente vergeleken wordt met die van Nederland gemiddeld. Ook is er informatie over gereden snelheden op te vragen (op basis van speed profiles van TomTom) en informatie uit het Meldpunt Veilig Verkeer (van Veilig Verkeer Nederland).

Deze applicatie gebruikt – behalve voor de ongevalslocaties – per gemeente *opgehoogde geregistreerde* aantallen op basis van de regionale registratiegraad. Het is immers bekend dat de onderregistratie van met name fietsongevallen zeer hoog is. Het is echter nog onvoldoende bekend hoe per gemeente aan *betrouwbare* werkelijke aantallen te komen is. Zeker bij gemeenten met kleine aantallen ongevallen, kan het corrigeren van geregistreerde aantallen op basis van regionale registratiegraden tot een behoorlijke vertekening leiden. Dat kan de geloofwaardigheid en daarmee het gebruik en de impact van een dergelijke benchmark sterk ondermijnen. Wel is het ViaStat Dashboard interessant als het gaat om de vormgeving en de mogelijkheid ervan om het gemeenten inzichtelijk te maken hoe goed het eigenlijk met hun verkeersveiligheid gaat.

4.5. SWOV Analyses Verkeersonveiligheid (SAVE)

In de jaren negentig heeft de SWOV een aantal gemeentelijke benchmarks uitgevoerd onder de noemer SAVE (SWOV Analyse Verkeersonveiligheid). Het doel daarvan was om de gemeentelijke verkeersveiligheid te verbeteren. De methode werd alleen ingezet als daarvoor interesse was vanuit de gemeenten zelf. De benchmarks werden uitgevoerd door de doelgemeente te vergelijken met andere gemeenten op alle circa twintig ongevalskenners van verkeersslachtoffers die de politie registreerde (zie bijvoorbeeld SWOV, 1992). Voor de betreffende gemeente werd gekeken naar de verdeling van het aantal slachtoffers over de ongevalskenners in de laatste drie jaar en werd deze vergeleken met de verdeling drie jaar eerder. Deze patronen werden daarnaast ook vergeleken met die in een aantal 'referentiegemeenten'. Referentiegemeenten werden vastgesteld aan de hand van een gecombineerde score van het aantal inwoners, de oppervlakte van het grondgebied, de lengte van het openbare wegennet, het aantal voertuigen, en drie indicatoren voor de samenstelling van de bevolking (aandeel kinderen, 13-15-jarigen en 18-24-jarigen). Voor iedere geïnteresseerde gemeente werd een rapport met uitleg over de methoden en de gevonden resultaten in tabellen geleverd, evenals een oplegrapport waarin de belangrijkste afwijkingen van de referentiegroep benoemd en in beeld gebracht waren.

De SAVE-opzet is een van de weinige benchmarkuitwerkingen waarbij daadwerkelijk statistisch getoetst is of afwijkingen tussen de onderzoeksgemeente en de referentiegemeenten aan toeval kon worden toegeschreven. Wel dient opgemerkt te worden dat een aantal van de door de politie geregistreerde ongevalskenners, zeker op het detailniveau van gemeenten, tegenwoordig zodanig onderhevig is aan registratieproblemen, dat veel van deze gegevens niet zo veel over de werkelijkheid zouden zeggen bij eenzelfde toepassing als destijds bij SAVE. Bij nieuwe gemeentelijke benchmarks zal hier goed naar gekeken moeten worden en dienen conclusies op basis van de cijfers met voorzichtigheid betracht te worden.

Deze wijze van benchmarking van verkeersveiligheidsprestaties is nog maar het begin van een benchmarkproces zoals het eruit hoort te zien volgens de literatuur uit de vorige hoofdstukken. Verklaringen voor afwijkingen van de referentie en aandacht voor het proces met de betrokken organisatie ontbreken hier nog.

4.6. Diverse regionale vergelijkingsstudies (SWOV)

Naast de methodisch vastomlijnde SAVE-exercities voor gemeenten, heeft de SWOV in de loop van de tijd ook diverse vergelijkingsstudies uitgevoerd op regionaal niveau. In een enkel geval betreft het een vergelijking tussen gemeenten. Vrijwel altijd was een concrete vraag van de betreffende organisaties aanleiding voor de vergelijking en werd de benchmark ingevuld aan de hand van de gestelde vraag en de op dat moment voorhanden inzichten en gegevens.

4.6.1. Verklaringen voor onveiligheid in Noord-Brabant

In de jaren zeventig leefde bij de Provincie Noord-Brabant de vraag of deze provincie onveiliger was dan andere (SWOV, 1976). Die vraag leidde tot een grondig onderzoek naar verkeersveiligheidscijfers en -kenmerken, expositiekenmerken (zoals inwoneraantal, oppervlakte en weglengte) en SPI's zoals alcoholgebruik, rijsnelheden en gordelgebruik. Al deze kenmerken van Noord-Brabant werden vergeleken met dezelfde kenmerken van de andere provincies en met het gemiddelde van Nederland (SWOV, 1984; Brouwer & Wegman, 1999). De vergelijkingen hadden betrekking op zowel de absolute als de relatieve aantallen ongevallen, SPI's en (voor zover beschikbaar) detailkenmerken daarvan, en op zowel de situatie in een bepaald jaar (of een gemiddelde van een aantal jaren) als op (geïndexeerde) ontwikkelingen. Dit leidde uiteindelijk tot conclusies over daadwerkelijke afwijkingen van de verkeersonveiligheid in de provincie Noord-Brabant ten opzichte van de rest van Nederland. Verklaringen voor de hoge slachtofferaantallen in deze provincie werden gevonden in zowel de expositiegegevens als SPI's.

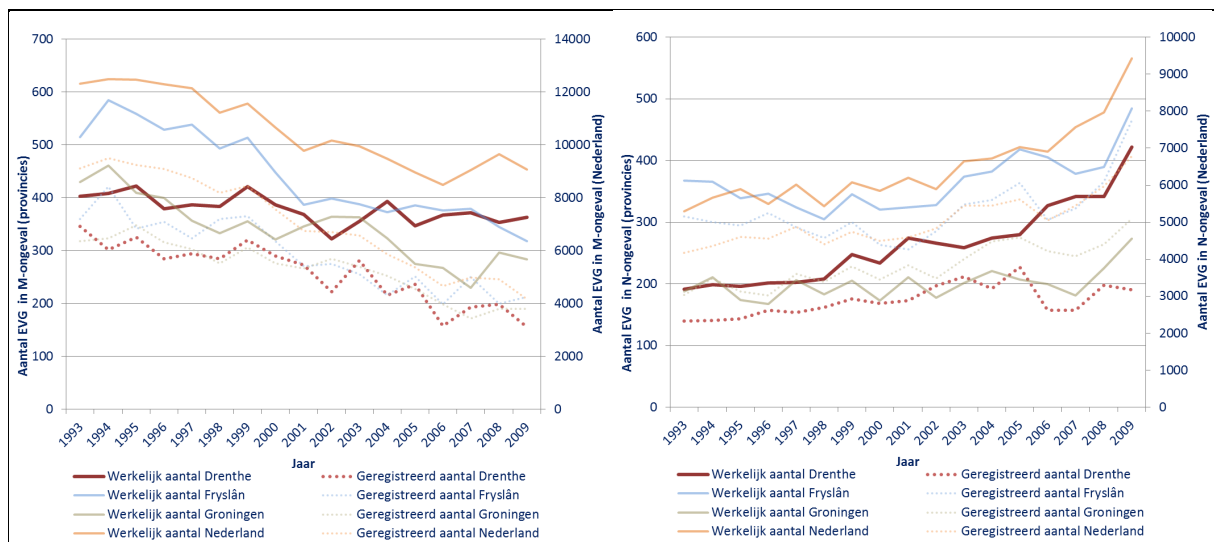
4.6.2. Vergelijkend onderzoek van verkeersonveiligheid in Zeeland

Ook voor de provincie Zeeland zijn soortgelijke analyses uitgevoerd als voor Noord-Brabant, al dan niet met SPI's (Hagenzieker & Wittink, 1995; Aarts et al., 2010). Hagenzieker & Wittink voerden hierbij ook analyses uit van de verkeersonveiligheid in vier Zeeuwse regio's. Deze regio's waren gedefinieerd op basis van geografische grenzen en betroffen de gebieden Noord-Zeeland, Walcheren, de Bevelanden en Zeeuws-Vlaanderen. Zo konden de onderzoekers een fijnere profilering geven van de verkeersonveiligheid binnen de provincie. In de analyse van Aarts et al. (2010) was vooral het doel om de typerende kenmerken van de Zeeuwse verkeersonveiligheid ten opzichte van die gemiddeld in Nederland duidelijk te maken en naar aanleiding daarvan aanbevelingen te doen voor beleidsaccenten.

Ook bij deze onderzoeken vormde een externe vraag de basis voor het onderzoek. Het onderzoek zelf was vooral een exercitie van de onderzoekers. Wel is zoveel mogelijk tussentijds afstemmingsoverleg geweest en zijn daar waar mogelijk lokale informatiebronnen benut.

4.6.3. Benchmarking van verkeersveiligheid als onderbouwing van beleidskeuzen

Meer recent voerde de SWOV vergelijkingen uit voor de gemeente Amsterdam (Reurings et al., 2012) en de provincie Drenthe (Aarts et al., 2013). Aanleiding was in beide gevallen een toekomstig beleidsplan en de wens dit te baseren op een analyse van ongevallen en oorzaken daarvan. In de studie voor Amsterdam werden ongevallencijfers van de stad vergeleken met die in de drie andere grootste steden in Nederland; in het geval van Drenthe was het verzoek om de situatie in die provincie niet alleen te vergelijken met het Nederlandse gemiddelde, maar expliciet ook te kijken naar de provincies Fryslân en Groningen, waar de provincie Drenthe zich mee verbonden voelt (zie *Afbeelding 4.5*).



Afbeelding 4.5. Voorbeeld van vergelijkingen van verkeersveiligheidsindicatoren van Drenthe met Nederland en de twee andere noordelijke provincies (bron: Aarts et al., 2013).

Hoewel er bij deze projecten geen intensieve interactie was tussen beleidsmakers en onderzoekers, was er in beide projecten wel duidelijk sprake van een wens vanuit de beleidsmakers en een daarop afgestemd benchmarkingsonderzoek. Dit is te beschouwen als een waardevolle stap op weg naar meer aandacht voor het proces. Een volgende stap zou zijn dat de organisaties zelf een benchmarkgroep formeren die samen met de onderzoekers de benchmark uitwerkt

5. Conclusies en aanbevelingen

In dit rapport zijn we nagegaan wat onder benchmarking verstaan kan worden, en hoe benchmarking van verkeersveiligheidsbeleid kan worden uitgewerkt als methode om beleidsmakers op decentraal niveau te helpen bij effectief en efficiënt beleid. Om deze vragen te beantwoorden hebben we zowel inzichten beschouwd uit de bestuurskundige literatuur – waarin veel geschreven is over benchmarking – en uit de wetenschappelijke verkeersveiligheidsliteratuur, waarin er ook aandacht is geweest voor dit onderwerp. Daarnaast is ook gekeken naar voorbeelden van uitwerkingen die in de praktijk aan benchmarking zijn gegeven.

In dit hoofdstuk vatten we de inzichten samen en komen we met een voorstel voor de wijze waarop benchmarking van verkeersveiligheid voor Nederlandse decentrale overheden, en dan met name gemeenten, vorm kan worden gegeven.

5.1. Definities van benchmarking en consequenties voor de praktijk

Op basis van de literatuur concluderen we dat voor verbetering van verkeersveiligheidsbeleid binnen Nederland, benchmarking het beste kan worden verstaan als: *een proces van vergelijken van prestaties, processen of strategieën van vergelijkbare organisaties (bijvoorbeeld overheden zoals gemeenten) door of samen met deze organisaties, waarbij naar verklaringen voor onderlinge verschillen wordt gezocht, succesfactoren worden aangewezen en leerpunten worden geïmplementeerd.* Andere vormen van benchmarking die alleen prestaties vergelijken of verklaringen zoeken zonder de organisaties die onderwerp zijn van onderzoek erbij te betrekken, noemen we verder geen benchmarking maar ‘vergelijkingen’ of ‘vergelijkingen en verklaringen’. Volgens bovenstaande definitie is benchmarking dus een interactief proces waarbij wetenschappers een ondersteunende, analyserende en kwaliteitsbewakende rol hebben, maar niet eenzijdig uitkomsten en leerrichtingen dicteren. Betrokken organisaties nemen zelf belangrijke stappen in acceptatie en implementatie van leerpunten, waarbij consensus binnen de benchmarkende groep organisaties van groot belang is. Naast feiten, cijfers en technisch valide conclusies, is het proces van benchmarking daarbij minstens zo relevant.

Binnen benchmarks voor verkeersveiligheid onderscheiden we:

1. prestatiebenchmarks:
 - a. benchmark van eindprestaties (ongevallen en slachtoffers);
 - b. benchmark van intermediaire prestaties (de inmiddels bekende SPI's zoals snelheid, alcohol en de kwaliteit van de inrichting van de infrastructuur);
 - c. benchmark van maatregelen (output van beleid);
2. procesbenchmarks;
3. strategiebenchmarks.

Benchmarking (en ook vergelijken en verklaren) is vooral een methode om minder goed presterende organisaties op een hoger plan te brengen en organisaties die willen verbeteren leerpunten te laten ontdekken.

5.2. Inzichten uit de bestuurskundige benchmarkliteratuur

In de bestuurskundige literatuur vinden we vooral een nadruk op het proces en het belang van samenwerking met de te benchmarken organisaties. Wel komen we tot de conclusie dat de stap van 'verklaren' vaak impliciet blijft of zelfs afwezig is. Er wordt van uitgegaan dat de werkwijze van de best presterende organisatie de verklaring vormt voor de goede prestaties, maar de relatie tussen (aspecten van) deze werkwijze en de prestatie wordt niet onderzocht of onderbouwd.

5.3. Inzichten uit de wetenschappelijke verkeersveiligheidsliteratuur over benchmarking

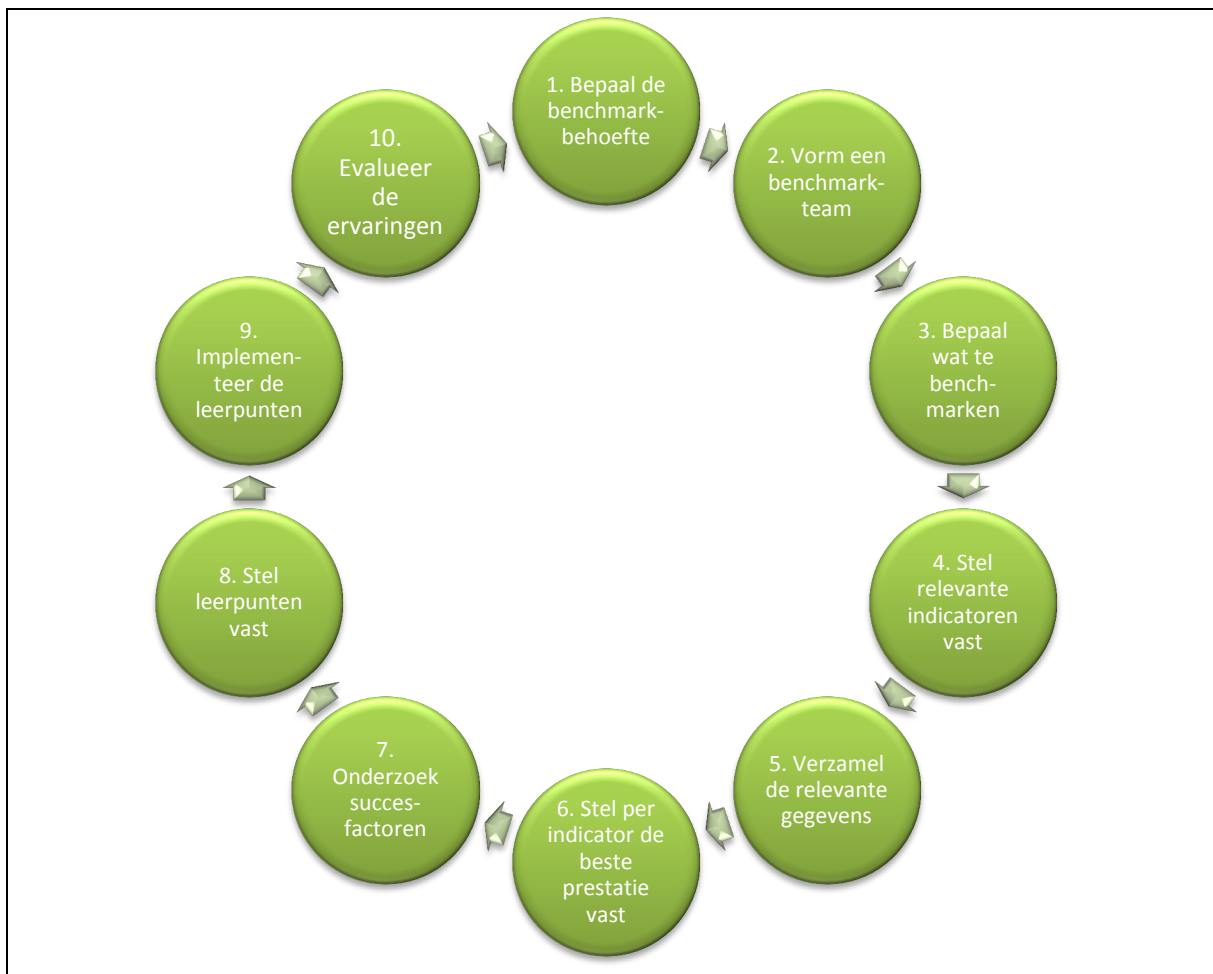
De wetenschappelijke verkeersveiligheidsliteratuur over benchmarken belicht eerder het technische deel van de benchmarking. Hier ligt vooral de nadruk op vergelijken en verklaren: op een technisch-wetenschappelijke uitwerking van indicatoren om prestaties te kunnen vergelijken en op methoden om indicatoren te kunnen combineren en ten opzichte van elkaar te kunnen wegen. Dit werk speelt zich echter buiten de interactie met beleidsmakers af, die ervoor kunnen zorgen dat de bevindingen daadwerkelijk tot verbeteringen in de verkeersveiligheid leiden.

5.4. Conclusies voor een vervolg

Het doel van benchmarken is uiteindelijk dat organisaties (zoals landen, regio's, gemeenten) leren en zichzelf verbeteren aan de hand van een vergelijking met andere, vergelijkbare organisaties die betere prestaties behalen. Om dit doel te bereiken is het belangrijk dat benchmarking impact heeft op het beleid van de te benchmarken partijen en dat er een accent ligt op betrokkenheid van deze partijen in het benchmarkproces. Zoals we in de literatuur hebben vastgesteld, is deze werkwijze bij verkeersveiligheids-benchmarks nog onderbelicht gebleven, dus het ligt voor de hand om deze weg ook hier te gaan verkennen.

Naast aandacht voor het proces is het ook belangrijk dat er aandacht blijft voor het zoeken en vinden van verbanden tussen goede prestaties van organisaties en de succesfactoren waaraan deze prestaties kunnen worden toegeschreven. Hier ligt een logische rol voor onderzoekers weggelegd.

Afbeelding 5.1 bevat een voorstel voor het proces van benchmarking van verkeersveiligheidsbeleid bij decentrale overheden in tien stappen. In dit schema zijn de inzichten van Bhutta & Huq (1999) en van Wegman & Oppe (2010) samengevoegd



Afbeelding 5.1. *Stappenplan voor het benchmarken van verkeersveiligheidsbeleid bij decentrale overheden (naar Bhutta & Huq, 1999 en Wegman & Oppe, 2010).*

5.4.1. *Bepaal de benchmarkbehoefte*

Nemen we als uitgangspunt dat organisaties – bijvoorbeeld gemeenten – zelf betrokken moeten zijn bij een benchmark, dan is het belangrijk om bij hen die behoefte te inventariseren of hen bewust te maken van de mogelijkheden die benchmarking kan bieden voor het verbeteren van hun prestaties. Dat kan bijvoorbeeld door in eerste instantie op basis van een aantal algemene cijfers inzichtelijk te maken hoe gemeenten er nu voorstaan (zie voorbeelden in *Hoofdstuk 4* en ook Houwing et al., 2012) en duidelijk te maken hoe organisaties achter mogelijke verklaringen voor die uitkomsten en leerpunten voor verbetering kunnen komen. Extra bekendheid voor benchmarking van verkeersveiligheid binnen Nederland kan mogelijk gegenereerd worden door aansluiting te zoeken bij het KpVV-dashboard Duurzame en slimme mobiliteit en bij het KING-keurmerk. Verdere keuzen in het benchmarkproces dienen samen met de te benchmarken organisaties in kaart te worden gebracht. Hierover gaan de volgende stappen.

5.4.2. Hoe een benchmarkteam te vormen?

Bij de vraag hoe tot een benchmarkteam te komen, is een aantal zaken van belang. Ten eerste moet er bij betrokken partijen interesse zijn om deel te nemen aan een benchmark; organisaties moeten immers bereid zijn het geleerde in de praktijk te brengen (zie Korsten, 2006). Daarnaast is het belangrijk dat organisaties vergelijkbaar zijn op een aantal algemene kenmerken omdat is gebleken dat vergelijkbare organisaties sneller geneigd zijn om van elkaar te leren (zie Wegman & Oppe, 2010). Bovendien vergemakkelijkt vergelijkbaarheid het vinden van verklaringen voor verschillen in prestaties en het aanwijzen van succesfactoren.

Verder zal de te vormen benchmarkgroep afhangen van de keuze van 'de benchmark'. Is het referentiepunt bijvoorbeeld een ideale norm, het gemiddelde van de te benchmarken organisaties of de best presterende? Om met de te benchmarken organisaties daadwerkelijk succesfactoren te kunnen vaststellen, is eigenlijk alleen de laatste vorm een optie. Wel kan het zo zijn dat op onderdelen de ene organisatie betere prestaties toont dan de rest, en op een ander onderdeel een andere. Bij het formeren van een benchmarkgroep is het van belang om te kijken of er tussen de te benchmarken partijen inderdaad prestaties zijn te vinden die boven die van de anderen uitsteken en dus als voorbeeld kunnen dienen.

Deze uitgangspunten in ogenschouw nemend, lijkt het dus verstandig om zowel breed (bottom-up) te inventariseren welke organisaties (van gelijk bestuursniveau) geïnteresseerd zijn om te benchmarken als vanuit contextuele kenmerken (top-down) te bekijken welke organisaties goed bij elkaar zouden kunnen passen en welke organisaties op welke verkeersveiligheidsonderdelen goede prestaties laten zien. Bij het vaststellen van vergelijkbare groepen is aan te bevelen om in ieder geval uit te gaan van bevolkingsdichtheid (stedelijk versus landelijk gebied) en dit zo nodig aan te vullen met andere structuur- en cultuurfactoren (zie ook Eksler, 2009; Aarts & Bax, 2013).

5.4.3. Wat te benchmarken?

In dit rapport ligt de focus op verkeersveiligheid en daarmee kan de vraag wat te benchmarken deels al worden beantwoord. Echter, we hebben in dit rapport geïllustreerd dat er in het proces van verkeersveiligheidsbeleid tot eindresultaat verschillende onderwerpen zijn aan te wijzen waarop een benchmark zich kan richten. Zo hebben we gezien dat in de literatuur onderscheid wordt gemaakt tussen het benchmarken van:

- prestaties (zowel eindindicatoren als intermediaire indicatoren);
- processen of de inhoud van beleid;
- strategieën, waarbij het nut van deze laatste voor verkeersveiligheid wel in twijfel wordt getrokken (zie Eksler, 2009).

In een *prestatiebenchmark* is achtereenvolgens inhoudelijk naar de volgende elementen te kijken:

1. vergelijking van genormeerde eindindicatoren (doden en ernstig verkeersgewonden);
2. vergelijking van intermediaire indicatoren (SPI's) en deze in verband brengen met de eindindicatoren;

3. vergelijking van maatregelen en deze in verband brengen met de vastgestelde intermediaire en eindindicatoren.

Bij een *procesbenchmark* in relatie tot verkeersveiligheid worden de volgende elementen onderzocht:

1. vergelijking van doelstellingen en ambities;
2. vergelijking van inhoudelijke beleidskeuzen in relatie tot de geformuleerde doelstellingen, inclusief de onderbouwing van die keuzen;
3. vergelijking van budgetten en de allocatie daarvan in de relatie tot beleidskeuzen en –doelen;
4. vergelijking van organisatie van de uitvoering, waarbij speciaal aandacht is voor kwaliteitsbewaking en de relatie tot de geformuleerde beleidsplannen;
5. vergelijking van de uiteindelijk genomen maatregelen in relatie tot de beleidsdoelen.

Bij een *strategische benchmark* is het van belang om strategische keuzen in relatie tot doelen en ambities te beschouwen en te vergelijken, te bekijken hoe deze strategische keuzen inhoudelijk en procesmatig worden uitgewerkt en hoe de kwaliteit van die uitwerking wordt gewaarborgd.

Binnen deze typen benchmarks en het onderwerp verkeersveiligheid kunnen nog verder inhoudelijke keuzen worden gemaakt. Zo kunnen de algemene verkeersveiligheidsprestaties van gemeenten vergeleken worden, of het algemene verkeersveiligheidsbeleid, maar ook kan ervoor worden gekozen om op een onderdeel daarvan te concentreren, bijvoorbeeld fietsveiligheid. De voorkeuren van de te benchmarken partijen zijn hierin belangrijk. In het benchmarkproces zelf kunnen vervolgens enkele standaardstappen gevolgd worden (zie verderop). Wetenschappers kunnen in algemene zin, maar ook specifiek bij dit onderdeel waardevol zijn om het proces te bewaken en de organisaties begeleiden in de te nemen stappen. Ze kunnen ook daar waar nodig bijdragen aan bewustwording van de keuzemogelijkheden wat betreft de typen benchmarks, en toezien op valide (gebruik van) gegevens en het trekken van verantwoorde conclusies.

5.4.4. *Hoe betekenisvolle indicatoren vast te stellen?*

Over betekenisvolle indicatoren is veel geschreven en onderzocht. In eerste instantie zijn de keuzen in de vorige stap daarbij primaire randvoorwaarden, en daarnaast ook specifieke wensen en kenmerken van de benchmark-groep. Een prestatiebenchmark vraagt andere indicatoren dan een procesbenchmark, en indicatoren voor een algemene verkeersveiligheidsbenchmark zullen algemener van aard zijn dan bijvoorbeeld een benchmark die zich richt op uitsluitend de fietsveiligheid in gemeenten.

Bij het kiezen van indicatoren die door te benchmarken organisaties geaccepteerd en gebruikt zullen worden is het in ieder geval belangrijk (zie Eksler, 2010; Korsten, 2006) dat de gebruikte indicatoren :

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

Omdat deze twee eisen met elkaar op gespannen voet staan, is het zaak om hier een goede balans in te vinden.

5.4.5. *Hoe relevante gegevens te verzamelen*

Voor verkeersveiligheid zijn heel wat gegevens beschikbaar, ook tot een vrij ver gedisaggregeerd niveau (bijvoorbeeld bij Centraal Bureau voor de Statistiek, Dutch Hospital Data, Ministerie van Infrastructuur en Milieu, SWOV). Onvermijdelijk heeft elke databron haar beperkingen; deze dienen voor gebruik goed in ogenschouw genomen te worden. Grootste voordeel van deze data is dat ze op uniforme wijze zijn verzameld en daarmee geschikt zijn voor onderlinge vergelijkingen op decentraal niveau.

Er zijn echter ook diverse typen data nodig die niet te verkrijgen zijn uit landelijk goed ontsloten databestanden. Voor deze data is daarom de uniformiteit tussen decentrale eenheden niet gegarandeerd. Zeker als het beleid en het beleidsproces zelf onderwerp zijn van benchmarking, is het zaak hierover van de organisaties zelf goede gegevens aangeleverd te krijgen, volgens een zo uniform mogelijk protocol.

Een andere mogelijkheid is om data alsnog te gaan verzamelen. Dit biedt weliswaar de meest kans op precies die data die men zoekt, anderzijds is het tijdrovend, is de medewerking van derden nodig (en zeker niet altijd gegarandeerd) en zijn er vaak nogal wat kosten aan verbonden.

5.4.6. *Hoe de beste prestatie vast te stellen?*

Het vaststellen van de beste prestatie komt er in wezen op neer dat prestaties gerangordend worden en dat de hoogste in de rangorde hiermee wordt vastgesteld. De rangordening zelf kan op eenvoudige wijze gebeuren (aan de hand van de hoogte van de gekozen indicator; zie bijvoorbeeld Eksler et al., 2009) of op complexere wijze (bijvoorbeeld door combinatie van meerdere indicatoren; zie bijvoorbeeld Hermans et al., 2008).

Een voordeel van de complexere werkwijze is dat meerdere indicatoren in één vervat kunnen worden en zo de 'best presterende' op basis van een totaalplaatje gekozen kan worden. Een nadeel hiervan is dat vaak de relatie met onderliggende, en voor beleidsmakers meer betekenisvolle indicatoren niet inzichtelijk is.

Eenvoudigere vergelijkingen hebben niet het voordeel van gecombineerde indicatoren, maar bieden wel weer kansen voor de te benchmarken organisaties om zich op deelterreinen te profileren en als best presterende uit de bus te komen. Zeker met het oog op het proces, lijkt dit daarom de voorkeur te verdienen.

Op het onderwerp verkeersveiligheid dient er rekening mee gehouden te worden dat prestaties fluctueren over de tijd en dat het verstandig kan zijn om zowel de huidige situatie als de verandering over de tijd mee te nemen. (zie bijvoorbeeld Bax et al., 2012; Eksler et al., 2009).

5.4.7. *Hoe succesfactoren te onderzoeken?*

In dit rapport hebben we geconstateerd dat het onderzoeken van succesfactoren in de bestuurskundige literatuur sterk onderbelicht is. Men lijkt aan te nemen dat het gedrag en de processen van de best presterende organisatie zonder meer bijdragen aan haar succes, en dat dat ook het

geval zal zijn als anderen ze overnemen. De werkelijkheid is echter complexer en de mogelijkheid bestaat dat ook niet-essentiële factoren worden overgenomen, of zelfs factoren die helemaal niet aan de goede prestaties bijdragen. Ook zal de context waarin het beleid plaatsvindt van invloed zijn op de bijdrage aan succes.

Bij het zoeken naar succesfactoren, dient men zich te realiseren dat oorzaken niet altijd goed of eenduidig te vinden zijn (zie bijvoorbeeld Stipdonk et al., 2006). In de verkeersveiligheid geldt dit zeker, omdat ook veel stochastische processen een rol spelen, die voor meer 'ruis op de lijn' zorgen naarmate kleinere gebieden worden onderzocht (zie bijvoorbeeld Davidse et al., 2007, Reurings et al., 2012).

Om voort te bouwen op de kennis die reeds bestaat ten aanzien van succes- (en faal)factoren, bevelen we aan om factoren die in de praktijk samen lijken te hangen met succesvolle prestaties en afwezig blijken in minder goed presterende organisaties, ook te verifiëren in de beschikbare onderzoeks-literatuur. Mocht die niet voorhanden blijken te zijn, dan brengt dit een leemte in kennis aan het licht. Benchmarking kan zo ook bijdragen aan een actuele kennisagenda en samenwerking met de te benchmarken organisaties kan er mogelijk toe bijdragen dat die kennisleemte opgevuld gaat worden, bijvoorbeeld doordat kennis en data worden verzameld. Aanvullend op de benchmark kan vervolgens onderzoek worden uitgevoerd dat voldoet aan de wetenschappelijke kwaliteitscriteria.

Met het oog op het benchmarkproces, is aan te bevelen in deze stap nadrukkelijk ook te horen van de deelnemende organisaties wat zij zelf als succesfactoren beschouwen. Als deze factoren ook met (literatuur)onderzoek gestaafd kunnen worden, biedt dit een meer solide fundering voor vervolgstappen.

5.4.8. *Hoe leerpunten vast te stellen?*

Als de succesfactoren in de vorige stap zijn vastgesteld, dan is het zaak om in de eigen organisatie te bezien in hoeverre men daarin van de beste prestaties afwijkt. Deze afwijkingen kunnen belangrijk of minder belangrijk zijn, groot of klein en op basis hiervan kan de organisatie een stappenplan opstellen welke verbeteringen, wanneer en hoe te implementeren. Uiteraard hangt dit stappenplan sterk samen met de wensen, mogelijkheden en omstandigheden van de organisatie.

5.4.9. *Hoe leerpunten te implementeren*

Het in de vorige stap vastgestelde 'leerplan' dient verder geoperationaliseerd te worden. Hoe het verdere proces van invoering eruit zal zien, hangt sterk af van de leerpunten, de afstand van de huidige tot de gewenste situatie, de mogelijkheden en de betrokken eenheden, afdelingen en mensen binnen de organisatie. Het implementeren zelf is vooral een interne aangelegenheid en kan een proces op zich zijn, zeker als het relatief grote hervormingen vergt.

5.4.10. *Hoe de ervaringen te evalueren*

Met het uitvoeren van een evaluatie is de beleidscyclus of PDCA-cyclus met daarin benchmarking als methode rond. Deze stap wordt helaas maar

weinig uitgevoerd, terwijl het voor verdere kennisontwikkeling en verbetering essentieel is dat ervaringen opgetekend en gedeeld worden. Van deze ervaringen kan ook weer in stap 7 (door anderen) geprofiteerd worden en kunnen anderen geholpen worden om op een hoger plan te komen. Bij een evaluatie is het zaak goed te kijken naar de succes- maar ook naar faalfactoren en de context waarin deze zich hebben afgespeeld.

Net als bij stap 7 speelt hierbij de moeilijkheid dat – als effecten op ongevallen en slachtoffers worden beschouwd – de kansfactor een grote rol speelt. Alvorens een evaluatie te starten is het dan ook verstandig om eerst voldoende tijd te laten verstrijken. Hoeveel tijd voldoende is, hangt mede af van het onderwerp van benchmarken (zie stap 3). Was het doel om prestaties op het niveau van ongevallen en slachtoffers te verbeteren, dan zijn meerdere jaren van verdere dataverzameling nodig om successen goed te kunnen vaststellen. Voor het benchmarken van SPI's kan met een kortere periode worden volstaan; dat kan ook gelden voor het benchmarken van beleidsprocessen.

5.5. Tot slot

Op basis van de inzichten die zijn opgedaan in dit rapport, is de SWOV voornemens in de praktijk te onderzoeken hoe benchmarking van verkeersveiligheidsbeleid bij gemeenten gestalte kan krijgen en bij kan dragen aan verbeteringen in het decentrale verkeersveiligheidsbeleid.

Literatuur

Aarts, L.T., Schermers, G., Hoekstra, A.T.G. & Goldenbeld, Ch. (2010). *Op weg naar nul vermijdbare verkeersslachtoffers in Zeeland. Beschouwing van effectieve maatregelen*. H-2010-2. SWOV, Leidschendam.

Aarts, L.T., Wesemann, P., Goldenbeld, Ch., Petegem, J.W.H. van, & Wijnen, W. (2013). *Verkeersveiligheidsanalyse Drenthe; Samenhang tussen Drentse kenmerken en verkeersveiligheid*. H-2013-3. SWOV, Den Haag.

Aarts, L.T. & Bax, C.A. (2013). *Regionale verschillen in relatie tot verkeersveiligheid; Nadere verkenning in de praktijk op basis van drie Zeeuwse gemeenten*. R-2013-14. SWOV, Leidschendam

Aarts, L.T. & Houwing, S. (te verschijnen). *Benchmarking road safety performance by grouping local territories: a study in the Netherlands*. Paper ingediend bij Journal of Transportation Research, Part A.

Aarts, L.T., Weijermars, W.A.M., Schoon, C.C. & Wesemann, P. (2008). *Maximaal 500 verkeersdoden, waarom eigenlijk niet? Maatregelpakketten en effectschattingen om te komen tot een aangescherpte verkeersveiligheidsdoelstelling*. R-2008-5. SWOV, Leidschendam

Abma, T.A. (1996). *Responsive evaluation*. Proefschrift Erasmus Universiteit Rotterdam.

Allsop, R.E., Sze, N.N. & Wong, S.C. (2011). *An update on the association between setting quantified road safety targets and road fatality reduction; short communication*. In: Accident Analysis and Prevention, vol. 43, p. 1279-1283.

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*. In: Benchmarking: An International Journal, vol. 6, nr. 3, p. 254-268.

Bliss, T. & Breen, J. (2009). *Country guidelines for the conduct of road safety management capacity reviews and the specification of lead agency reforms, investment strategies and safe system projects*. The World Bank, Washington, D.C..

Breen, J. (2000). *Benchmarking road safety in the UK with the rest of Europe*. In: 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, p. 3-15.

Brouwer, M. & Wegman, F.C.M. (1999). *Verkeersveiligheid in Noord-Brabant; Terug in het peloton*. R-99-8. SWOV, Leidschendam.

Camp, R.C. (1989). *Benchmarking: the search for industry best practices that lead to superior performance*. Quality press for the American Society for Quality Control. Milwaukee, Wisconsin.

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*. In: *Benchmarking: An International Journal*, vol. 10, nr. 3, p. 176-209.

Davidse, R.J., Aarts, L.T. & Stipdonk, H.L. (2007). *Analyse van ernstige verkeersongevallen in Zeeland in 2006 en beleidsaanbevelingen*. R-2007-7. SWOV, Leidschendam.

Deming, W.E. (1986). *Out of the Crisis*. MIT Center for Advanced Engineering Study

Donkers, E. & Coppen, H (2012). *ViaStat Dashboard. Dé app (iOS) voor verkeersveiligheid en bereikbaarheid. Een app voor de iPhone en iPad, inclusief de nieuwe iPad*. DNL9100_202-R02. VIA.nl, Vught.

Eksler, V. (2009). *Road mortality in the EU: a regional approach*. PhD Thesis, Versailles St.-Quentin University, Versailles.

Eksler, V. (2010). *Measuring and understanding road safety performance at local territorial level*. In: *Safety Science*, vol. 48, p. 1197-1202.

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

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

Eksler, V., Popolizio, M. & Allsop, R. (2009). *How far from Zero? Benchmarking of road safety performance in the Nordic countries*. European Transport Safety Council ETSC, Brussels.

ETSC (2001). *Transport safety performance indicators*. European Transport Safety Council ETSC, Brussels.

Francis, G., & Holloway, J. (2007). *What have we learned? Themes from the literature on best-practice benchmarking*. In: *International Journal of Management Reviews*, vol. 9, nr. 3, p. 171-189.

Gitelman, V., Doveh, E., & Hakkert, S. (2010). *Designing a composite indicator for road safety*. In: *Safety Science*, vol. 48, p. 1212–1224.

Grove, R., Mihai, F., Bishop, M., Eckeroth, K. & Shepherd, P. (2002). *Road network asset management: international benchmarking study*. AP-R204/02, AUSTRROADS. Haymarket, NSW.

- Hagenzieker, M.P. & Wittink, R.D. (1995). *Verkeersonveiligheid in de provincie Zeeland; Resultaten van een analyse van de verschillen met de rest van Nederland en van de verschillen tussen vier regio's binnen de provincie*. R-95-38. SWOV, Leidschendam.
- Hakkert, A.S. & Gitelman, V. (red.) (2007). *Road Safety Performance Indicators: Manual*. Deliverable D3.8 of the EU FP6 project SafetyNet. European Commission, Brussels.
- Helden, G.J. van, & Tillema, S. (2005). *In search of a benchmarking theory for the public sector*. In: *Financial Accountability & Management*, vol. 21, nr. 3, p. 337-361.
- Hermans, E., Bossche, F. van den & Wets, G. (2008). *Combining road safety information in a performance index*. In: *Accident Analysis and Prevention*, vol. 40, p. 1337–1344.
- 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 Steunpunt Mobiliteit en Openbare Werken IMOB, Hasselt.
- Hermans, E., Brijs, T., Wets, G. & Vanhoof, K. (2009b). *Benchmarking road safety: lessons to learn from a data envelopment analysis*. In: *Accident Analysis and Prevention*, vol. 41, p. 174-182.
- Hoogerwerf, A. (red.). (1993). *Overheidsbeleid*. Samsom H.D. Tjeenk Willink, Alphen aan den Rijn.
- 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. SWOV, Leidschendam.
- Huggins, R. (2010). *Regional competitive intelligence: Benchmarking and policy-making*. In: *Regional Studies*, vol. 44, nr. 5, p. 639-658.
- Jost, G., Allsop, R. & Steriu, M. (2013). *Back on track to reach the EU 2020 Road Safety Target? 7th road safety Performance Index PIN report*. ETSC PIN report 7. European Transport Safety Council ETSC, Brussels.
- Kimber, R. (2003). *Benchmarking the safety of roads: learning from Euro NCAP'S star rating of safety*. TRL, Crowthorne, Berkshire.
- KING (2010). *Standaard Waarstaatjegemeente.nl. Uitvoering burgerpeiling en klanttevredenheidsonderzoek*. Oktober 2010, versie 2.0. Kwaliteitsinstituut Nederlandse Gemeenten KING, Den Haag.
- KING (2011). *Handleiding KING keurmerk benchmarks 2012*. Versie 2.0. Kwaliteitsinstituut Nederlandse Gemeenten KING, Den Haag.
- Klassen, N. & Wahlstroem, H. (2005) *Closing the gap - benchmarking European roads within EuroRAP*. In: *Proceedings of the European Transport Conference ETC, Strasbourg, France, 18-20 September 2005*.

- Knowles, J., Sexton, B., Lawton, B. & Charman, S. (2010). *Benchmarking road safety in Northern Ireland*. PPR478. TRL, Crowthorne.
- Koornstra, M., Lynam, D., Nilsson, G., Noordzij, P., et al. (2002). *SUNflower. A comparative study of the development of road safety in Sweden, the United Kingdom, and the Netherlands*. SWOV Institute for Road Safety Research, Leidschendam.
- Korsten, A.F.A. (2006). *Ins & outs van benchmarking*. Geraadpleegd 5 december 2013, op <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 <http://www.arnokorsten.nl/PDF/Onderzoek/Mythen%20over%20benchmarking.pdf>
- Korver, W. & Haan, J. de (2012). *Creating a sense of urgency for sustainable mobility by benchmarking cities in the Netherlands*. Presentatie voor het internationale POLIS congres, op 4 december 2012, Perugia, Italië.
- Kozak, M. (2004). *Chapter 1. Overview of benchmarking theory*. In: Kozak, M. (Ed.), *Destination benchmarking: concepts, practices and operations*. CABI Publishing, Wallingford.
- Kyrö, P. (2003). *Revising the concept and forms of benchmarking*. In: *Benchmarking: An International Journal*, vol. 10, nr. 3, p. 210-225.
- LTSA (2000). *Road safety strategy 2010: A consultation document*. National Road Safety Committee, Land Transport Safety Authority LTSA, Wellington.
- Moriarty, J.P. & Smallman, C. (2009). *En route to a theory of benchmarking*. In: *Benchmarking: An International Journal*, vol. 16, nr. 4, p. 484-503.
- Muhlrad, N., Gitelman, V. & Buttler, I. (Ed.) (2011). *Road safety management investigation model and questionnaire*. Deliverable 1.2 of the EC FP7 project DaCoTa.
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., et al. (2005). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Statistics Working Papers, 2005/3, Organisation for Economic Co-operation and Development OECD, Paris.
- Noordhoek, M. (2013). *Municipal benchmarking; Organisational learning and network performance in the public sector*. PhD Thesis Aston University. Eburon, Delft.
- Poerstamper, R.J.C., Mourik-van Herk, A. van & Veltman, A.C. (2007). *Benchmarking in de zorg : op weg naar een excellente organisatie*. Elsevier Gezondheidszorg, Maarssen.

Reurings, M.C.B., Aarts, L.T. & Schermers, G. (2012). *Analyse van het aantal verkeersdoden in 2011 in Amsterdam; Is er reden tot zorg?* R-2012-6. SWOV, Leidschendam.

Shewhart, W.A. (1939). *Statistical method from the viewpoint of quality control*. Dover, New York.

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.

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%2Fbudapeval%2Fstern.pps&ei=2wHZUoSULOqq0QWR6oGQBQ&usg=AFQjCNEONRcogfoF7xxLv12BDFFc3HbiEg&bvm=bv.59568121,d.d2k

Stipdonk, H.L., Aarts, L.T., Schoon, C.C. & Wesemann, P. (2006). *De essentie van de daling in het aantal verkeersdoden; Ontwikkelingen in 2004 en 2005, en nieuwe prognoses voor 2010 en 2020*. R-2006-4. SWOV, Leidschendam.

SWOV (1976). *De verkeersonveiligheid in de provincie Noord-Brabant I en II*. R-76-5. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, Voorburg.

SWOV (1984). *De verkeersonveiligheid in de provincie Noord-Brabant X. Eindrapport*. R-84-16. Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, Leidschendam.

SWOV (1992). *Analyse van de verkeersonveiligheid in Weert. SWOV analyse verkeersonveiligheid; Tabellen en toetsingsresultaten*. A-92-23. SWOV, Leidschendam.

Wegman, F., Commandeur, J., Doveh, E., Eksler, V., et al. (2008). *SUNflowerNext: Towards a composite road safety performance index*. Deliverable D6.16 of the EU FP6 project SafetyNet.

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*. SWOV, Leidschendam.

Wegman, F. & Oppe, S. (2010). *Benchmarking road safety performances of countries*. In: Safety Science, vol. 48, p. 1203–1211.

Wegman, F., Vis, M. & Gitelman, V. (2009). *A composite Road Safety Performance Index using the SUN-flower approach*. In: ITE Journal, vol. 79, nr. 8, p. 26-30.

Weijermars, W.A.M., Aarts, L.T. & Schoon, C.C. (2009). *Hoe veilig is Gelderland op (de) weg? Analyse van de huidige verkeersveiligheid in Gelderland en toekomstverkenning en beleidsaanbevelingen voor 2010 en 2020*. R-2009-13. SWOV, Leidschendam

Weijermars, W.A.M. & Wijnen, W. (2012). *Verkeersveiligheidsverkenning 2020: effecten van extra maatregelen. Effectschattingen voor bijstelling van het Strategisch Plan Verkeersveiligheid*. R-2012-14. SWOV, Leidschendam.

Wong, S.C., Sze, N.N., Yip, H.F., Loo, B.P.Y., et al. (2006). *Association between setting quantified road safety targets and road fatality reduction*. In: *Accident Analysis and Prevention*, vol. 38, p. 997-1005.

Wong, S.C. & Sze, N.N. (2010). *Is the effect of quantified road safety targets sustainable?* In: *Safety Science*, vol. 48, p. 1182-1188.

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

Bijlage A

Details ten aanzien van indicatoren van verkeersveiligheid en het gebruik daarvan in benchmarking

Aan de hand van de verschillende lagen in de verkeersveiligheidspiramide (zie *Hoofdstuk 2*) gaan we concreter in op indicatoren op elk van de piramidelagen.

Slachtoffer- en ongevallenindicatoren

De mate van verkeersveiligheid wordt primair afgemeten aan het aantal (ernstige) ongevallen en slachtoffers in het wegverkeer. Omdat we het hier over indicatoren hebben die sterk onderhevig zijn aan stochastische processen, is het belangrijk om rekening te houden met statistische validiteit wanneer dergelijke cijfers worden gebruikt bij het vergelijken van gebieden.

Zo doet Eksler (2010) uitspraken over de validiteit van het gebruik van ongevallen- en slachtoffercijfers voor verschillende detailniveaus van gebieden: vergelijking van landen laat andere – grovere – indicatoren toe dan vergelijking van kleinere gebieden zoals regio's of gemeenten, waarbij kleine aantallen al snel een probleem vormen. Eksler beveelt daarom aan om bij vergelijking van kleinere gebieden:

1. ook de ernstige gewonden erbij te betrekken.
2. te normeren door middel van bayesiaanse statistiek.

Bij dit laatste gaat het om de kans op een bepaalde prestatie gegeven de mate van voorkomen van relevante expositiematen (bijvoorbeeld de hoeveelheid daadwerkelijke ernstige ongevallen met jongeren, gegeven de kans daarop gezien hun risico en hun aandeel in de populatie).

Prestatie-indicatoren verkeersveiligheid

Indicatoren uit de laag 'prestatie-indicatoren' geven de output van een gebied weer en kunnen worden benut voor verklaring van verschillen in 'outcome'. Prestatie-indicatoren zijn gedefinieerd als toestanden in het verkeerssysteem die een sterke causale relatie hebben met het ontstaan van (ernstige) verkeersongevallen en verkeersslachtoffers (ETSC, 2001). Het betreft factoren zoals:

- snelheid
- alcohol
- gebruik van beveiligingsmiddelen
- gebruik van motorvoertuigverlichting
- voertuigveiligheid
- kwaliteit van de infrastructuur
- traumamanagement

(Hakkert & Gitelman, 2007; Bax et al. 2012)

Voor zover prestatie-indicatoren gebruikt worden om naar verschillen in verkeersveiligheidsprestaties tussen landen te kijken, blijken dataproblemen (beschikbaarheid en homogeniteit) veelal een probleem (zie onder andere Gitelman, 2010; Hermans et al., 2008). Dit leidt ertoe dat een iets ander beeld van de verkeersveiligheidsprestaties naar voren komt als naar deze

indicatoren wordt gekeken dan wanneer naar aantallen slachtoffers of ongevallen wordt gekeken.

Maar ook worden prestatie-indicatoren individueel gebruikt om gebieden of producten onderling te kunnen vergelijken. De bekendste op het gebied van verkeersveiligheid zijn de Road Protection Score van EuroRAP voor de kwaliteit van infrastructuur en de EuroNCAP-scores voor de botsveiligheid van personenauto's (zie bijvoorbeeld Kimber, 2003).

Beleidsindicatoren

De laag daar weer onder, zou in belangrijke mate weer moeten geven hoe het beleid in een gebied is vormgegeven en in welke maatregelen energie wordt gestoken. Bevatte deze laag oorspronkelijk vooral de concrete beleidsplannen en –acties (LTSA, 2000), later is deze laag geoperationaliseerd in vijf kwaliteitsindicatoren voor beleid (zie Wegman et al., 2008):

- verkeersveiligheidsdoelstellingen,
- gekozen interventies,
- economische evaluatie,
- monitoring van beleidsprestaties en
- betrokken partijen in het beleidsprogramma.

Parallel aan deze ontwikkeling stellen Bliss & Breen (2009) in een invloedrijk wereldbankrapport dat de beleidslaag van de piramide bestaat uit a) institutionele factoren en b) interventies.

In *Afbeelding 2.2* zijn bovengenoemde factoren uitgeplitst in 'ouput' van beleid (maatregelen) en troughput (proces en strategie).

Vervolgens is de vraag in hoeverre bovengenoemde factoren inderdaad aantoonbaar samenhangen met verkeersveiligheid in termen van ongevallen en slachtoffers. In de literatuur is vooralsnog weinig evidentie gevonden voor een relatie tussen vooral institutionele beleidsfactoren en succes in het verbeteren van verkeersveiligheid (zie Muhrad et al., 2011), Bax et al., 2012) niet in de laatste plaats omdat veel factoren alleen nog maar globaal beschreven zijn en verdere operationalisatie behoeven. Een enkele uitzondering is de positieve relatie tussen verkeersveiligheidsdoelstellingen en reductie in verkeersdoden in een land (Wong et al., 2006; Wong & Sze, 2010; Allsop et al., 2011). Deze beleidsfactor is echter niet onderscheidend als het gaat om vergelijkingen binnen Europa of vergelijkingen op sub-nationaal niveau, aangezien op al deze niveaus verkeersveiligheidsdoelstellingen bestaan.

De kennis over effecten van interventies op verkeersveiligheid is veel beter ontwikkeld en betreft kennis die afkomstig is van evaluaties van maatregelen (zie *Tabel A.1* voor een voorbeeld). Om dergelijke kennis in vergelijkingsstudies tussen gebieden te kunnen gebruiken, zijn wel de juiste gegevens nodig over hoeveelheid en bereik van de implementatie. Dergelijke informatie is veelal niet (uniform) voorhanden (zie Bax et al., 2012)

Maatregel	Besparing in letselongevallen
Rotonde (buiten bebouwde kom)	Ca. -70%
Vrijliggend fietspad	-25 tot -60%
Ongelijkvloerse kruisingen	Ca. -50%
Drempels/plateaus	Ca. -30%
60 km/uur zone	Ca. -25%
30 km/uur zone	-20 tot -30%
Parallelweg	Ca. -15%
Intelligente maatregelen	b.v. ISA: -60% doden, -50 % EVG
Handhaving	-10% tot -40% doden; -5% tot -30% EVG
Educatie en voorlichting	?

Tabel A.1. *Enkele voorbeelden van effecten van beleidsmaatregelen op de verbetering van verkeersveiligheid, in dit geval maatregelen gericht op de reductie van rijnsnelheden.*

De kennis over effecten van interventies kan ook worden gebruikt in vergelijkingsstudies waarbij een absoluut criterium centraal staat. Dit is het geval als beleidsplannen (in termen van concrete maatregelen) in verband worden gebracht met een bepaald beleidsdoel. In Nederland wordt dit regelmatig gedaan op landelijk niveau (de zogenoemde ‘verkenningen verkeersveiligheid’; zie bijvoorbeeld Aarts et al., 2008; Weijermars & Wijnen, 2012); op subnationaal niveau is dit echter veel minder gebruikelijk en zijn er slechts incidentele voorbeelden (zie bijvoorbeeld Weijermars et al., 2009).

Contextindicatoren

De onderste laag van de verkeersveiligheidspiramide – de structuur en cultuurlaag - wordt vooral gebruikt als contextuele laag die als basis dient om gebieden te groeperen (zie ook volgende paragraaf). Uiteraard wordt hiervoor alleen die context in ogenschouw genomen waarvan een relatie bekend is met verkeersveiligheidsprestaties.

De meest prominente contextuele factor die in verband is gebracht met verkeersveiligheidsprestaties van een gebied, betreft urbanisatiegraad of nog beter: populatiedichtheid (zie Aarts & Bax, 2013; Aarts & Houwing, te verschijnen; Eksler, 2009; Ekseler & Lasarre, 2008). Zo vond Eksler (2009) bijvoorbeeld dat 60% van de variantie in mortaliteit verklaard werd door bevolkingsdichtheid van een gebied. In België bleek daarnaast ook dat de buitenwijken van steden een relatief hoge mortaliteit hebben en zich ook minder gunstig ontwikkelen dan andere gebieden.

Andere contextuele factoren die in relatie zijn gebracht met verkeersveiligheid maar veel minder verklarend zijn voor verschillen tussen gebieden zijn (zie Aarts & Houwing, te verschijnen; Eksler, 2009, 2010; Wegman et al., 2008):

1. Klimaat (direct: bijvoorbeeld door veroorzaken van gladheid of langere remweg; indirect: via invloed op gekozen vervoerswijzen)

2. Reliëf (bergachtig gebied geeft beperkingen aan het wegennet, het verloop en inrichting van wegen; heeft ook invloed op vervoerswijzen)
3. Wegennetwerk (via wegenstructuur, snelheidslimieten, verkeersintensiteiten, V/C-ratio, kruispuntendichtheid, weglengte)
4. Demografische kenmerken (aandeel adolescenten (vooral mannen) en ouderen)
5. Economische factoren (werkloosheid en commerciële activiteiten werken vooral door in expositie; daarnaast hangen investeringen van overheden af van beschikbare budgetten)
6. Vervoerswijzen (sommige vervoerswijzen zijn gevaarlijker dan andere)

Volgens Eksler (2009) wordt een belangrijk deel van de variantie in verkeersmortaliteit verklaard door bovengenoemde min of meer fysieke kenmerken, meer dan door verschillen in risico's op de weg.

Naast deze fysieke kenmerken (die zowel statisch als dynamisch kunnen zijn), onderscheidt Eksler (2009) ook nog:

1. Operationele structuren (organisatie van actoren) en
2. Cultuur (attituden, normen en waarden).

De relatie die deze factoren hebben met verkeersveiligheid lopen via risico (factoren kunnen risicoverhogend of –verlagend zijn) en blootstelling (de mate waarin verkeersdeelnemers aan een bepaald risico worden blootgesteld).

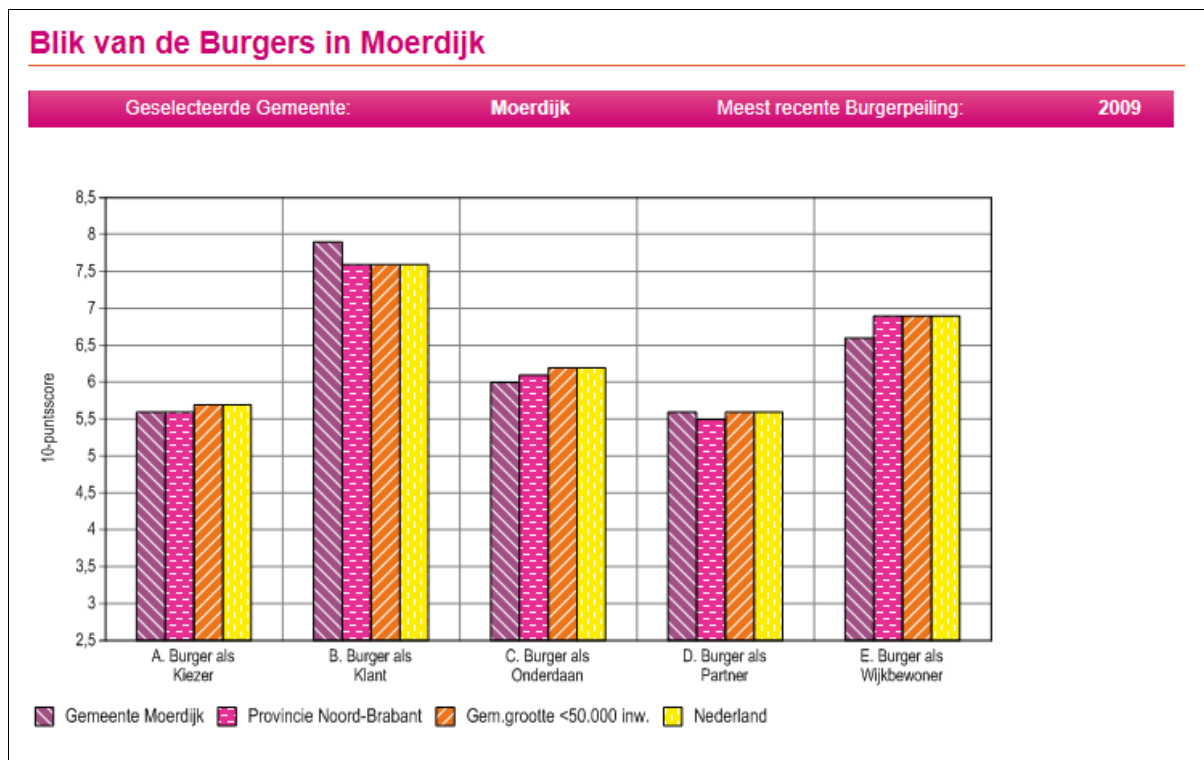
Bijlage B

Details ten aanzien van producten van KING

In de prestatieoverzichten is verkeersveiligheid op de volgende wijzen bij KING gepresenteerd:

Burgerparticipatie

In regelmatig terugkerende enquêtes wordt burgers gevraagd naar onder andere hun mening over de verkeersveiligheid in hun gemeente en in hun buurt. Naar verdere details wordt niet gevraagd. Informatie wordt per gemeente vergeleken met de gemiddelde score per provincie, de gemiddelde score van gemeenten van gelijke grootte (3 klassen²) en de gemiddelde score in Nederland (zie *Afbeelding B.1* en *Afbeelding B.2*). Vergelijkingscores zijn altijd afkomstig uit hetzelfde jaar als waarin de peiling van de betreffende gemeente is gehouden.



Afbeelding B.1. Weergave van burgerpeiling op waarstaatjegemeente.nl voor een willekeurig geselecteerde gemeente.

² 1) < 50.000 inwoners, 2) 50.000 - 100.000 inwoners en 3) > 100.000 inwoners.

E5. De wegen en fietspaden binnen de gemeente zijn goed onderhouden

	Moerdijk	Provincie Noord-Brabant	Gem.grootte <50.000 inw.	Nederland	Max	Min
Oordeel burger onderhoud wegen en fietspaden	6,3	6,7	6,7	6,7	8,0	5,1

Bron: Burgerpeiling algemene dienstverlening Waarstaatjegemeente.nl

[Meer informatie](#)

E6. De verkeersveiligheid in de gemeente is goed

	Moerdijk	Provincie Noord-Brabant	Gem.grootte <50.000 inw.	Nederland	Max	Min
Oordeel burger verkeersveiligheid gemeente	6,5	6,3	6,5	6,5	7,4	5,1
Oordeel burger verkeersveiligheid buurt	6,4	6,3	6,4	6,4	7,0	5,5

Bron: Burgerpeiling algemene dienstverlening Waarstaatjegemeente.nl

[Meer informatie](#)

Afbeelding B.2. Weergave van de burgerpeilingsgegevens over verkeersveiligheid op waarstaatjegemeente.nl voor een willekeurig geselecteerde gemeente.

De burgerpeiling kan jaarlijks worden uitgevoerd op basis van een vaste vragenlijst. Gemeenten kunnen deze vragenlijst zelf afnemen of hem laten afnemen door enquêtebureaus die door het KING hiervoor gecertificeerd zijn. De eerstvolgende peiling zou in 2013 plaatsvinden, maar tijdens de meting is men gestuit op een aantal problemen waardoor de meting tot nader order is uitgesteld.

De opzet van de burgerpeiling en ook het klanttevredenheidsonderzoek is strak vastgelegd door KING (KING, 2010). Zo wordt er een ondergrens gesteld van 400 ingevulde enquêtes per gemeente (bij minder dan 375 ingevulde formulieren worden de resultaten niet opgenomen in waarstaatjegemeente.nl) en een minimale respons van 30%. Personen moeten random getrokken worden uit de gemeentelijke basisadministratie, minimaal 18 jaar zijn en niet wonen in een verzorgings- of verpleegtehuis.

Cijfers

De laatst bekende feiten over geregistreerde ongevallen en slachtoffers (doden, ziekenhuisgewond en lichtgewond) worden hier per gemeente gepresenteerd, alsmede de tendens over de tijd. Er wordt geen melding gemaakt van onderregistratie (zie *Afbeelding B.3, B.4, en B.5*).

Verkeersveiligheid in Moerdijk (Min. van Infrastructuur en Milieu)

Geregistreerde ongevallen

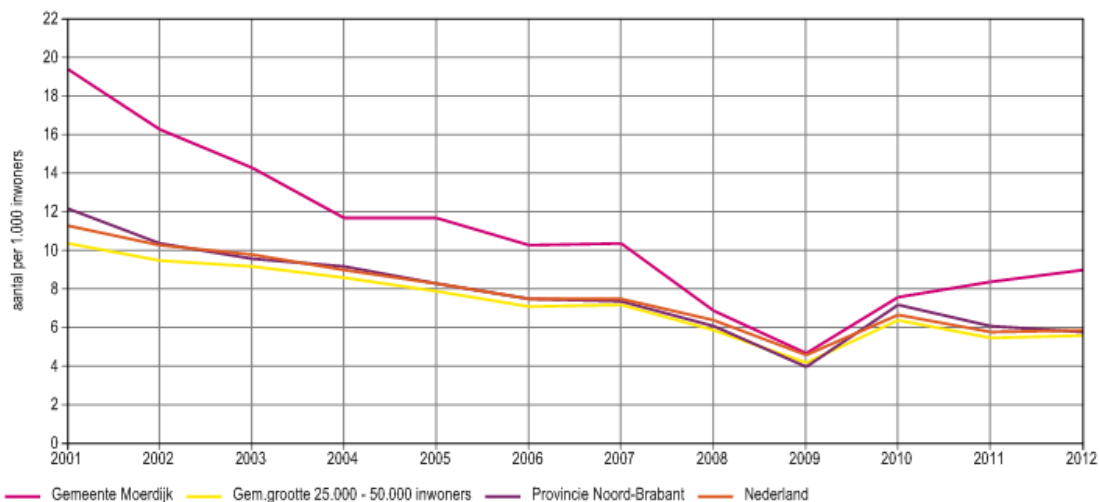
In de gemeente Moerdijk vonden in 2012 327 (geregistreerde) ongevallen plaats, waarvan 322 met uitsluitend materiële schade en 1 met dodelijke afloop. Ongevallen met lichamelijk letsel komen in de gemeente Moerdijk relatief minder vaak voor dan gemiddeld in gemeenten uit dezelfde gemeentegrootteklasse (Gem.grootte 25.000 - 50.000 inwoners). Ongevallen met dodelijke afloop komen relatief minder vaak voor.

Geregistreerde ongevallen in 2012 (abs., rel.)

		Moerdijk		Gem.grootte 25.000 - 50.000 inwoners	Nederland
		abs.	%	%	%
Dodelijke afloop	2012	1	0,3	0,7	0,5
Letsel ziekenhuis	2012	2	0,6	2,1	2,1
Spoedeisende hulp	2012	2	0,6	1,6	1,8
Licht gewonde(n)	2012	0	0,0	0,7	0,6
Uitsluitend materiële schade	2012	322	98,5	94,9	95,0

Het aantal ongevallen in de gemeente Moerdijk is sinds 2001 met 53,8% afgenomen. De afname is substantieel. Sinds 2008 bedraagt de mutatie 29,5%.

Geregistreerde ongevallen in de periode 2001 - 2012 (per 1.000 inwoners)



Afbeelding B.3. Geregistreerde ongevallencijfers voor een willekeurige gemeente zoals gepresenteerd in waarstaatjegemeente.nl.

Slachtoffers

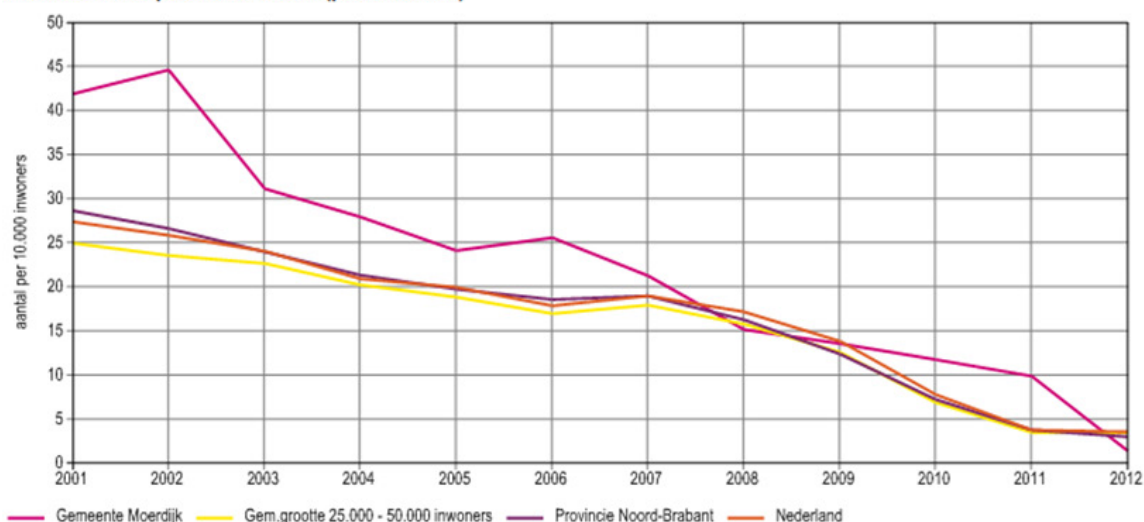
In 2012 waren in de gemeente Moerdijk 5 slachtoffers te betreuen. In vergelijking met gemeenten van gelijke grootte belanden in de gemeente Moerdijk meer slachtoffers in het ziekenhuis.

Slachtoffers in 2012 (abs., rel.)

		Moerdijk	Moerdijk	Gem.grootte 25.000 - 50.000 inwoners	Nederland
		abs.	%	%	%
Verkeersdoden	2012	1	20,0	11,3	9,2
Slachtoffers, ziekenhuisopname	2012	2	40,0	39,2	38,2
Slachtoffers, eerste hulp	2012	2	40,0	49,5	52,6

Het aantal slachtoffers in de gemeente Moerdijk is sinds 2001 met 96,7% afgenomen. De afname bedraagt 40,6 slachtoffers per 10.000 inwoners. Voor de gemeente Moerdijk gaat het hier effectief om een afname naar 5 slachtoffers. Het aantal slachtoffers is sinds 2008 gewijzigd met -91,0%.

Slachtoffers in de periode 2001 - 2012 (per 10.000 inw.)

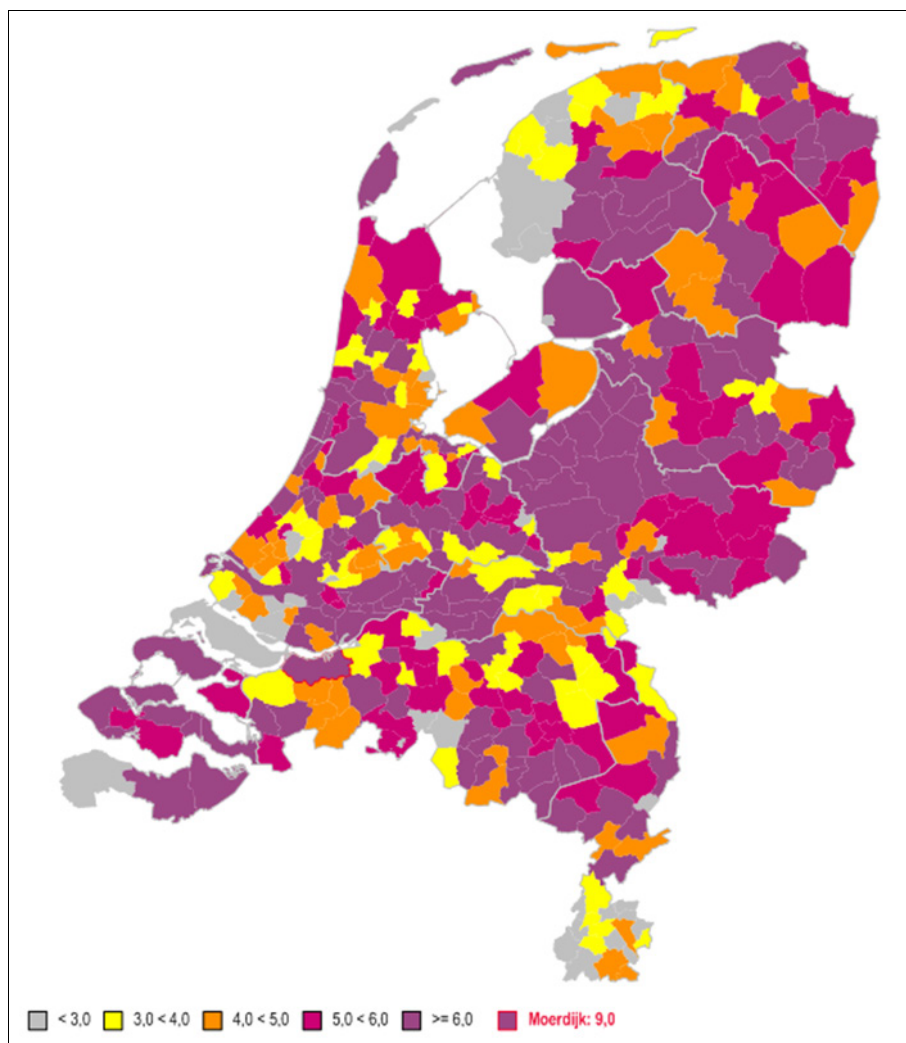


In de gemeente vallen relatief minder slachtoffers onder autogebruikers en minder slachtoffers onder fietsers dan gemiddeld in gemeenten van vergelijkbare bevolkingsomvang.

Verkeersslachtoffers naar vervoersmodus in 2012 (per 10.000 inwoners)

		Moerdijk	Gem.grootte 25.000 - 50.000 inwoners	Provincie Noord-Brabant	Nederland
Voetganger	2012	0,3	0,2	0,1	0,3
Fiets	2012	0,3	0,9	0,9	0,9
Snorfiets	2012	0,0	0,2	0,2	0,2
Bromfiets	2012	0,0	0,4	0,2	0,5
Motorfiets / scooter	2012	0,3	0,2	0,2	0,2
Personenauto	2012	0,5	1,4	1,1	1,3
Bestelauto	2012	0,0	0,1	0,1	0,1
Vrachtauto	2012	0,0	0,0	0,0	0,0
Bus (ov)	2012	0,0	0,0	0,0	0,0
Overig	2012	0,0	0,0	0,0	0,0

Afbeelding B.4. Geregistreeerde slachtoffercijfers inclusief onderverdelingen voor een willekeurige gemeente zoals gepresenteerd in waarstaatjegemeente.nl.



Afbeelding B.5. Presentatie van geregistreeerde ongevallen 2012 per 1000 inwoners in het cijfergedeelte van [waarstaatjegemeente.nl](http://www.waarstaatjegemeente.nl). Een zelfde kaart is ook beschikbaar voor slachtoffers/1000 inwoners.

Verklaringsmodellen

Het KING heeft samen met Atlas voor gemeenten (www.atlasvoorgemeenten.nl) algemene verklarende modellen ontwikkeld voor een aantal dossiers die gemeenten meer inzicht bieden in de achtergronden van hun prestaties op het betreffende dossier. Het gaat hierbij om verbanden tussen algemene en landelijk beschikbare kenmerken van gemeenten. Per gemeente kan een samenvatting worden gepresenteerd. Er is een aparte rapportage beschikbaar met methodologische verantwoording.

Benchmarking en het KING-Keurmerk Benchmarking

Het KING-Keurmerk Benchmarks houdt in dat een keurmerk voor drie jaar wordt uitgegeven aan een product dat aan de volgende voorwaarden voldoet (KING, 2011):

- Eindresultaat en tevredenheid: toegevoegde waarde voor deelnemers die blijkt uit een evaluatie na iedere benchmarkcyclus

- Inhoud: onder andere indicatorenset, omvang van de populatie en betrouwbaarheid van data
- Visie: de benchmark gaat uit van de eigen kracht van gemeenten en wordt samen met hen opgesteld en uitgevoerd (doelstellingen, indicatoren, opzet van het proces, evaluatie); ook wordt toegezien op het verkrijgen van bestuurlijk commitment; tevens wordt een stimulerende leeromgeving geboden en wordt leren van elkaar gestimuleerd
- Proces en organisatie: opzet en professionaliteit waarbij resultaten eigendom van de gemeenten worden; er is een enkelvoudige uitvraag en meervoudig gebruik.

Het keurmerk is niet statisch en proces en resultaat worden tussentijds getoetst. KING toetst zelf niet op het naleven van deze criteria, maar laat deze toetsing uitvoeren door DEKRA certificering (voorheen KEMA). Het keurmerk kost per drie jaar €1.250,-.



Stichting Wetenschappelijk Onderzoek Verkeersveiligheid SWOV
SWOV Institute for Road Safety Research

Postbus 93113
2509 AC Den Haag
Bezuidenhoutseweg 62
2594 AW Den Haag
T 070 - 317 33 33
F 070 - 320 12 61
E info@swov.nl
I www.swov.nl

PO Box 93113
2509 AC The Hague, The Netherlands
Bezuidenhoutseweg 62
2594 AW The Hague, The Netherlands
T 070 - 317 33 33
F 070 - 320 12 61
E info@swov.nl
I www.swov.nl