

Hoofdstuk 14 – Benchmarking van beleid

Sjoerd Keulen en Jedid-Jah Jonker

1. Inleiding

Prestatievergelijking, ook wel benchmarking genoemd, is al jaren bezig met een opmars binnen overheden en diverse beleidsdisciplines. Ook in de media duiken regelmatig lijstjes op waarin landen met elkaar worden vergeleken, van de landen met de meest concurrerende economieën ter wereld (Global Competitiveness Index van het World Economic Forum) tot de ranglijstjes met prestaties van leerlingen per land (PISA onderzoek van de Organisatie voor Economische Samenwerking en Ontwikkeling –OESO). Maar zijn dit soort lijstjes hetzelfde als benchmarking? Wat is benchmarking eigenlijk? Wat kun je ermee?

Kort gezegd gaat het bij benchmarking over vergelijking. En wel vergelijken met als doel om te leren: hoe sta ik ervoor in vergelijking met anderen en wat kan ik doen om mezelf te verbeteren door goed te kijken naar wat anderen beter doen dan ik? Dit lerende element maakt benchmarking ook geschikt als onderdeel van het beleidsproces. Zo stelt benchmarking beleidsmakers en beleidsuitvoerders in staat om (de uitvoering van) beleid te verbeteren. Voor het onderzoek naar doelmatigheid is benchmarking zelfs onontbeerlijk.

In dit hoofdstuk zullen we het begrip ‘benchmarking’ nader verkennen, en het vooral bekijken toegepast op de publieke sector. Eerst (paragraaf 14.2) gaan we in op wat benchmarking is en welke modellen hiervoor bestaan. Vervolgens (paragraaf 14.3) bespreken we de vraag waarom benchmarking zo belangrijk kon worden in de publieke sector. Daarna gaan we in op de belangrijkste aandachtspunten en valkuilen (paragraaf 14.4) voor benchmarking in de publieke sector. Tot slot gaan we in op de ontwikkeling van benchmarking bij de Nederlandse overheid (paragraaf 14.5) en op het internationaal gebruik van benchmarks (paragraaf 14.6).

2. Wat is benchmarking en wat kun je met de verschillende modellen meten?

In de organisatiekunde gaat het er bij een “benchmark” om dat een organisatie haar eigen prestaties vergelijkt met die van andere, met als doel om te kijken of en hoe de eigen prestaties verbeterd kunnen worden (vrij naar Robert Camp, 1989).¹ Oorspronkelijk is het woord “benchmark” (ijkpunt) afkomstig uit de landmeetkunde, en betekent het “afzetten tegen een bepaalde standaard”. In andere woorden: het is een combinatie van vergelijken en op basis daarvan normatief beoordelen.

Het gebruik van het woord “benchmark” in organisatiekunde en bedrijfskunde is opgekomen eind jaren tachtig van de twintigste eeuw. Tegenwoordig is benchmarking niet meer alleen voorbehouden aan bedrijven: allerlei publieke organisaties, van scholen en ziekenhuizen tot gemeenten en zelfs landen, passen het toe om te kijken hoe zij “presteren” in vergelijking met hun zusterorganisaties of evenknieën (in het Engels ook wel aangeduid als *peers*).

Xerox en benchmarking

Het Amerikaanse bedrijf Xerox was in de jaren 70 marktleider in de verkoop van kopieerapparaten. Maar die positie kwam onder druk te staan door Japanse concurrenten die kwalitatief betere apparaten op de markt brachten voor een lagere prijs dan Xerox. Medewerkers van Xerox gingen onderzoeken hoe de concurrentie in staat was om betere prestaties te leveren dan zichzelf voor minder geld: hoe hadden zij hun processen ingericht? Welke aanpakken werden er gebruikt? De inzichten die deze “benchmark” opleverden, stelden Xerox in staat de eigen manier van produceren te optimaliseren en zo haar marktpositie te versterken. Twee medewerkers van Xerox, Robert Camp en David Kearns, hebben hun ervaringen met dit proces beschreven in diverse publicaties. De publicatie *“Benchmarking: the search for industry best practices that lead to superior performance”* van Robert Camp uit 1989 wordt algemeen gezien als de eerste waarin de term “benchmarking” wordt gebruikt in de context van het verbeteren van prestaties.

De belangrijkste keuzes bij het doen een benchmark zijn: wat ga je vergelijken en met wie vergelijk je jezelf?

2.1 Wat ga je vergelijken?

Benchmarking kan op diverse momenten in de beleidsproces gebruikt worden.

Benchmarking van strategieën is een nuttig instrument voor de beleidsbepaling (zie hoofdstuk 6). In de uitvoeringsfase (zie hoofdstuk 7) kan benchmarking gebruikt worden kritische prestatie-indicatoren op te stellen en om de geleverde prestaties te vergelijken met andere organisaties. Dit kan zowel voor de doelmatigheid van de bedrijfsvoering (benchmarking van processen), als voor de doelmatigheid van het beleid (de benchmarking van prestaties). Voor de evaluatie van beleid (zie hoofdstuk 9) kan benchmarkdata gebruikt worden om achteraf de doelmatigheid van beleid vast te stellen. Dit betekent dat er drie varianten van benchmarking mogelijk zijn: (1) benchmarking van strategieën, (2) benchmarking van processen en (3) benchmarking van prestaties.²

Bij de eerste variant, benchmarking van strategieën, gaat het onder meer om de strategische aanpak die is gekozen door succesvolle organisaties. Elementen waar je dan aan zou kunnen denken, zijn onder meer: richten deze organisaties zich op een specifieke doelgroep (of juist niet), hoe ziet hun strategisch personeelsbeleid eruit en welke rol speelt de aandacht voor innovatie?

Bij het benchmarken van een proces wordt gekeken of een bepaald proces, bijvoorbeeld de inrichting van de eerstehulp post in een ziekenhuis of het proces van klachtenafhandeling bij de sociale dienst, verbeterd kan worden op basis van ervaringen van andere ziekenhuizen of andere sociale diensten. Het gaat hierbij meestal om een duidelijk afgebakend deel van de totale organisatie.

Als je tot slot wilt weten hoe de organisatie presteert in vergelijking met andere organisaties, ligt een benchmark van resultaten voor de hand. In dat geval wordt het geheel aan prestaties van bijvoorbeeld een ziekenhuis of school vergeleken met de prestaties van andere ziekenhuizen of andere scholen. Het hoeft hierbij niet alleen te gaan om de prestaties van een hele organisatie: er kan ook worden gekeken naar de prestaties van delen

van een organisatie, zoals een bepaald specialisme in een ziekenhuis of een bepaalde klas van een school.

2.2 Met wie ga je vergelijken?

Bij een benchmark staat het vergelijken met andere organisaties centraal. Een organisatie of bedrijf kan kiezen voor een interne benchmark. Dan worden afdelingen binnen de organisatie met elkaar vergeleken.

Heel gebruikelijk is om een concurrerende benchmark te doen, waarbij een organisatie zich spiegelt aan (directe) concurrenten of zusterinstellingen. In de publieke sector gaat het meestal om organisaties met een zeer vergelijkbare taak, zoals ziekenhuizen, scholen of gemeenten.

Een andere vorm van benchmarking is dat een organisatie onderdelen van haar eigen organisatie vergelijkt met die van organisaties die niet inhoudelijk vergelijkbaar zijn. Dit wordt wel functionele benchmarking genoemd. Een organisatie kan bijvoorbeeld bij heel andere organisaties kijken hoe functies zoals marketing, logistiek of research en development worden gedaan.

Een laatste vorm van benchmarking lijkt op functionele benchmarking, maar gaat over de grenzen van functies heen en wordt wel generieke benchmarking genoemd. Bij generieke benchmarking wordt gekeken hoe organisaties (grotere) processen hebben georganiseerd, zoals klantenservice of de facturering.³

Bij een interne benchmark of een functionele benchmark is men op zoek naar de best presterende organisaties op een bepaald terrein (*best in class*). Bij de concurrerende benchmark, en tot op zekere hoogte ook bij de generieke benchmark, is men juist op zoek naar de organisaties die als geheel de beste prestaties laten zien (*best practice*).

2.3 Resultaten van benchmarking

Benchmarking kan een aantal positieve resultaten hebben. Het biedt allereerst een referentiepunt voor de eigen presentatie. Daarnaast stimuleert benchmarking een leerhouding, draagt het bij aan een externe oriëntatie en kan het werknemers motiveren. Verder kan benchmarking helpen om de beste methoden en werkwijzen over te nemen van andere organisaties. Ook heeft benchmarking een strategische functie: het kan helpen om de doelstellingen van organisaties te bepalen en bijdragen aan de vorming en uitvoering van de strategie.

Verder is benchmarking onontbeerlijk als men uitspraken wil doen over de doelmatigheid van een organisatie, is een benchmark onontbeerlijk. Doelmatigheid is namelijk een relatief begrip en kan niet vastgesteld worden zonder vergelijking: de ene organisatie kan doelmatiger zijn dan een andere. Maar als er maar één organisatie is en je kunt nergens mee vergelijken, dan is het niet mogelijk om te bepalen of deze organisatie doelmatig is.

Er zijn verschillende (econometrische) technieken beschikbaar die voor een set organisaties met een vergelijkbare taak (bijvoorbeeld ziekenhuizen, scholen of universiteiten) kunnen bepalen welke organisaties het meeste produceert tegen de laagste kosten, dit noemen we kostenefficiënt. Deze methoden kunnen ook aangeven welke invloed bepaalde kenmerken, zoals de samenstelling van het personeel of de schaalgrootte van een organisatie, op de

kostenefficiëntie hebben. Deze methoden (*stochastic frontier analysis* en *data envelopment analysis*) zijn een zeer kwantitatieve manier van benchmarking.⁴

In andere vormen van benchmarking wordt vaak gewerkt met een combinatie van kwantitatieve en kwalitatieve gegevens. In het kwantitatieve gedeelte wordt gekeken naar de ingezette middelen en de resulterende productie. Bij de kwalitatieve gegevens gaat het bijvoorbeeld om interviews met betrokkenen en uitkomsten van focusgroepen.⁵

2.4 Algemene aandachtspunten bij benchmarking

Het uitvoeren van een goede benchmark kost veel tijd: de aanpak vraagt een gedegen voorbereiding en met de uitvoering gaat flink wat menskracht gepaard. Als je probeert op een snelle en goedkope manier een benchmark te doen, bestaat het risico dat je op onterechte gronden een bepaalde organisatie aanwijst als 'beste praktijk' of dat er niet genoeg rekening wordt gehouden met verschillen in context, waardoor je appels met peren aan het vergelijken bent.

Er bestaan in de praktijk ook misverstanden over het verschil tussen een monitor en een benchmark. Soms wordt een bepaald onderzoek een benchmark genoemd, terwijl het in feite alleen een monitor is. In een monitor wordt het verloop van bepaalde ontwikkelingen in de tijd gevolgd. In veel gevallen gaat het om indicatoren die op een rij worden gezet. De *Atlas van de lokale lasten* van COELO is een goed voorbeeld van een monitor. In dat onderzoek wordt de ontwikkeling van de lokale lasten per gemeente gevolgd. In een monitor wordt niet gezocht naar verklaringen en wordt anders dan bij benchmarking geen oordeel over prestaties gegeven. Om die reden wordt in een monitor ook niet systematisch gekeken naar achtergrondkenmerken en er wordt niet gepoogd om te leren.⁶ De informatie uit een monitor kan wel een nuttige basis zijn voor een benchmark.

3. Waarom is benchmarking in de publieke sector zo aangeslagen?

De opkomst van benchmarking in de publieke sector is bestuurskundig te duiden aan de hand van vijf thema's: (1) New Public Management, (2) performance management, (3) de opkomst van de vraag naar transparantie, (4) het groeiend belang van verantwoording en (5) (transnationale) governance.

New Public Management is een managementstroming binnen de publieke sector die vanaf halverwege de jaren tachtig opkwam en die managementpraktijken uit de private sector overnam. Via professioneel management zou de publieke sector meer resultaten kunnen boeken tegen lagere kosten.

De bestuurskundige Hood liet zien dat performance management, waarbij prestaties werden gemeten aan de hand van duidelijke doelen en afspraken over output, een essentieel onderdeel is van New Public Management. Concurrentie tussen uitvoerende organisaties vormde een tweede onderdeel in deze managementstroming. Benchmarking kan voor beide zaken goed gebruikt worden. Benchmarking kan daarnaast helpen om de bestedingen te vergelijken en zo inspireren of zelfs voorschrijven hoeveel efficiënter een organisatie in vergelijking met andere gerund kan worden.⁷

In de jaren negentig van de twintigste eeuw nam benchmarking een ware vlucht. Transparantie, de openheid van publieke organisaties en politiek, wordt vergroot als je prestaties tussen organisaties kunt vergelijken. Dit wordt vaak gekoppeld aan de claim dat deze openheid leidt tot betere prestaties, hogere kwaliteit of kostenbesparing vanuit het principe dat vreemde ogen dwingen tot verandering.

De combinatie van afrekenbare doelen vanuit het gedachtengoed van performance management in combinatie met transparantie leidde tot meer aandacht voor verantwoording. Doordat doelen explicieter werden en het inzicht via benchmarking werd vergroot, speelde benchmarking ook een rol in de verantwoording. Benchmarking bleek niet alleen een middel om prestaties van het management van de organisatie te vergelijken, maar ook voor politici zelf bleek het een handig middel om overheidsprestaties inzichtelijk en vergelijkbaar te maken.⁸

Door verdergaande internationale samenwerking bijvoorbeeld in de Europese Unie, maar ook via verdergaande globalisering en internationale handel, werd transnationale governance belangrijk: aansturing op internationaal niveau. Om de samenwerking te stimuleren en als een zacht en daardoor minder weerstand oproepend beleidsmiddel raakten zo de internationale benchmarks in zwang. Er is bijna geen groot internationaal vraagstuk of strategie (duurzame ontwikkelingsdoelen, klimaatverandering of de Covid-19 pandemie) waarvan de voortgang niet gebenchmarkt en/of gemonitord wordt.

4. Waar moet je op letten bij benchmarking?

Alhoewel benchmarking een geaccepteerde managementtechniek is, is er ook kritiek op de methodiek en achtergrond van de techniek. Daarnaast is het goed om bekend te zijn met de valkuilen die er zijn bij het gebruik in de publieke sector. De systematiek van benchmarking is zo algemeen geformuleerd, dat toepassing binnen publieke organisaties goed mogelijk is. Maar omdat het doel van een publieke organisatie anders is dan dat van een private organisatie vooral als die gericht is op winst maken, is het wel belangrijk om oog te houden voor de publieke context en de uitdagingen die hieraan eigen zijn.

De belangrijkste uitdagingen zijn: waardenpluraliteit, coproductie van beleid, externe beïnvloeding van beleid, de bedreiging van transparantie voor de organisatie, legitimering van slechte prestaties en de perverse werking van prestatie-indicatoren zelf. Daarnaast zijn er valkuilen die met name gelden voor internationale benchmarks, maar waarvan het ook nuttig is om die op het netvlies te houden wanneer benchmarking als techniek wordt ingezet.

4.1 Het ontbreken van winstprikkelers en het bestaan van waardenpluraliteit bij publieke organisaties

Het handelen van de overheid is gericht op het realiseren van maatschappelijke doelen zoals welzijn, veiligheid en welvaart. Het beleid dat hiervoor wordt ontworpen, wordt net als in het bedrijfsleven geacht doeltreffend (effectief) en doelmatig (efficiënt) te zijn. Anders dan in het bedrijfsleven bestaat er vaak geen concurrentie die zorgt voor kostenminimalisatie en oog voor de wens van de klant. In de publieke sector opereren – ook vanuit kosten oogpunt – vaak monopolistische organisaties, zoals het Kadaster of de Belastingdienst, of zijn de prijzen

voor de verrichtingen sterk gereguleerd, bijvoorbeeld in de gezondheidszorg. Door het ontbreken van een winstprikkel is het onderling vergelijken door benchmarking een goede optie om tot prestatieverhoging te komen.⁹

Doeltreffendheid en doelmatigheid zijn niet de enige waarden waaraan een publieke organisatie moet voldoen. Zo spelen ook rechtvaardigheid, rechtsstatelijkheid, verantwoording en democratische waarden een grote rol in de afweging van het opstellen en evalueren van beleid. De controleerbaarheid, betrouwbaarheid en verdeling kunnen ertoe leiden dat beleid minder efficiënt kan worden vormgegeven dan wanneer met zulke zaken geen rekening wordt gehouden. Ook kunnen bijvoorbeeld vanwege privacybescherming of vanwege het gelijkheidsbeginsel niet alle data gemeten worden die je nodig hebt om bijvoorbeeld de doeltreffendheid of het bereik van beleid te kunnen meten. Denk hierbij bijvoorbeeld aan de registratie van etniciteit of aan het koppelen van persoonlijke data uit verschillende overheidsadministraties.¹⁰

4.2 Beleid in ketens of door middel van coproductie bemoeilijkt toerekening

Veel overheidsbeleid komt tot stand in ketens waarin verschillende organisaties en departementen samenwerken. Denk bijvoorbeeld aan de strafrechtketen, waarin politie, openbaar ministerie, de rechtbanken en het gevangeniswezen allemaal van invloed zijn op de doorlooptijd van strafzaken. Daarnaast is beleid soms een gedeelde verantwoordelijkheid van verschillende overheidslagen of behoren opdrachtgever en opdrachtnemer niet tot dezelfde organisatie.

Door de opkomst van netwerkmanagement en co-creatie komen steeds meer prestaties tot stand in samenwerking met derden. Dit alles maakt het lastig om prestaties aan een afzonderlijke organisatie toe te rekenen. Benchmarking is daarom lastig, omdat je deze techniek alleen kunt inzetten voor productieprocessen die je als organisatie volledig zelf kunt beïnvloeden.¹¹

4.3 Beleidseffecten lastig te isoleren

Prestatiemeting door benchmarking is lastig, omdat voor beleid in het algemeen geldt dat het lastig is om maatschappelijke invloeden zoals economische recessies of een gespannen arbeidsmarkt die invloed hebben op de effecten van het te benchmarken beleid of productieproces, weg te filteren. De beleidseffecten zijn hierdoor lastig te isoleren. Dit heeft tot gevolg dat benchmarking binnen publieke organisaties vaak het meest informatief is bij het vergelijken van uniforme overheidsproducten en uniforme productieprocessen.¹²

4.4 Perverse werking van prestatie-indicatoren

De inzet van prestatie-indicatoren leidt niet automatisch tot een hogere productie. Een toename van prestatiemetingen kan zelfs leiden tot onbedoelde gevolgen en in negatieve gevallen zelfs tot een lagere productiviteit. Zeker als er een zwakke correlatie is tussen de indicator en de prestatie verliest de prestatie-indicator over een langere waarnemingsperiode zijn waarde. De rapportage zegt hierdoor steeds minder over de daadwerkelijke productie: dit wordt de prestatie-paradox genoemd.

De prestatie-paradox kan vier oorzaken hebben. Ze kan voortkomen uit positieve leereffecten. Doordat de prestatie door leren verbetert, verliest de indicator zijn waarde als gevolg van het feit dat er steeds minder slechte productie wordt gemeten. De tweede

oorzaak is pervers leren. De organisatie leert wat er gemeten wordt en gebruikt die kennis om de meting te manipuleren. Zo kan de gemeten productie omhoog gaan zonder dat er feitelijk iets verbeterd is. De derde oorzaak is selectie. Doordat slechte prestaties door goede worden vervangen, boet de prestatie-indicator aan onderscheidende kracht in. Tot slot kan hetgeen wat gemeten wordt, ook genegeerd worden.¹³

Voorbeeld van prestatie-paradox

Boetes voor verkeersovertredingen zijn bedoeld om het verkeer veiliger te maken. Als de politie wordt afgerekend op het aantal boetes dat ze uitschrijft voor verkeersovertredingen, ontstaat er een prikkel voor de politie om op die plekken boetes uit te schrijven waar veel overtredingen worden begaan. Dit hoeven niet de plekken te zijn die het meest onveilig zijn. Als de onveilige plekken daardoor minder worden gecontroleerd dan voorheen, kan het werken met het aantal uitgeschreven verkeersboetes als indicator voor de verkeersveiligheid dus zelfs leiden tot een lagere verkeersveiligheid.

Zo kunnen prestatie-indicatoren een perverse werking hebben. Wanneer gemakkelijk haalbare prestatie-indicatoren worden geselecteerd zijn de prestaties op papier succesvol maar verbetert de feitelijke productie niet. Dit kan zelfs zo ver gaan dat wanneer uit benchmarking blijkt dat een organisatie bovengemiddeld goed presteert, deze organisatie lagere doelen stelt.

Tot slot zijn prestatiemeting en benchmarking zijn niet gratis. De tijd die het kost om gegevens te verzamelen of te registeren, gaat ten koste van de primaire productie en kan extra overhead met zich meebrengen. Daarom is het van belang om de kosten van de meting altijd af te zetten tegen de opbrengsten van meer inzicht door benchmarking.

4.5 Valkuilen bij internationale benchmarks

Met name over internationale benchmarks is de laatste jaren kritiek geuit. Deze kritiek is toepasbaar op benchmarking in de publieke sector in den brede. De kritiek draait erom dat in benchmarks complexe fenomenen worden versimpeld en geëxtrapoleerd en op symbolische wijze veroordeeld, bijvoorbeeld door lijstjes van top tien landen met de meeste slavernij, corruptie, etc. Deze vertaalslag, die normatieve en omstreden agenda's verhult en tot ogenschijnlijke neutrale en technocratische beoordelingen verengt, krijgt juist door deze presentatievorm geloofwaardigheid en zeggingskracht. Wezenlijk betwiste begrippen en complexe sociaal-economische, politieke en culturele kwalitatieve fenomenen, zoals discriminatie, democratie en stabiliteit, worden radicaal versimpeld tot een getal dat dikwijls zijn oorsprong vindt in percepties van geënquêteerden schattingen op basis van gissingen.¹⁴

Voor internationale benchmarking worden in de bestuurskundige literatuur drie soorten valkuilen omschreven: samenwerkingsproblemen, problemen over de keuze van criteria en indicatoren en transferproblemen.¹⁵

Internationale benchmarks kunnen alleen slagen als er sprake is van samenwerking, maar lang niet al het te vergelijken beleid is onomstreden. Zo was er in de betrokken landen veel ophef over de benchmark die de Wereldgezondheidsorganisatie in 2000 had uitgevoerd naar zorgsystemen van verschillende landen, onder meer omdat landen vonden dat de inrichting

van het zorgsysteem een zaak is van soevereine nationale regeringen.¹⁶ Ook kan er sprake zijn van internationale competitie, wat samenwerkingsmechanismen hindert.

Omdat complex internationaal beleid vaak lastig in output en outcomefactoren is te vertalen en omdat politieke systemen en nationale contexten sterk verschillen, is het vaak moeilijk om de juiste criteria en indicatoren te bepalen. Benchmarking is vaak gericht op het identificeren en daarna overnemen van de best practice. In de praktijk blijkt de overdraagbaarheid, het overzetten van beleid, vanwege de verschillen in context en politieke systemen erg lastig of sluit het niet aan bij lokale voorkeuren.

5. Wat is de ontwikkeling van benchmarking in Nederland?

New Public Management, de managementfilosofie die de overheid meer als een bedrijf wil gaan runnen, heeft grote invloed gehad op de introductie en promotie van benchmarking in Nederland. Deze managementtechniek kon immers goed gebruikt worden om de overheid bedrijfsmatiger te laten werken. Er kwam meer nadruk op prestaties, doelmatigheid en doeltreffendheid en op het management zelf.

Naast benchmarking kwamen er ook vergelijkingstechnieken als monitoring en visitatie in zwang. Ze vonden hun ingang in grote publieke operaties van de jaren negentig als 'Marktwerking, Deregulering en Wetgevingskwaliteit'. Ook werkte deze benadering door bij de opzet van een nieuw begrotingsstelsel dat opgebouwd werd rond beleidsindicatoren (het VBTB-proces, dat staat voor 'Van Beleidsbegroting tot Beleidsverantwoording').

5.1 De start van benchmarken bij de Nederlandse overheid: een leermiddel

Begin twintigste eeuw begon de Rijksoverheid benchmarking op tal van terreinen te promoten. Zo zetten de ministeries van Binnenlandse Zaken en Onderwijs in 2001 een kennisbank op om ervaringen uit te wisselen. Een jaar later waren er al 21 cases beschreven. In hetzelfde jaar vroeg het kabinet de Raad voor Openbaar Bestuur (ROB) advies over de toepassing van benchmarking als middel voor beleidsverantwoording door gemeenten en provincies aan het Rijk. De ROB meende dat benchmarking een nuttig instrument was om als overheidsorganisatie te leren van andere organisaties, maar dat benchmarking ook ongewenste effecten kon hebben. Daarom moest benchmarking een leermiddel blijven en geen informatiefunctie krijgen, laat staan een toezichtsfunctie. In de kabinetsreactie nam de regering daar afstand van: het kabinet zag benchmarking ook als middel om verantwoording af te leggen, toezicht te houden en organisaties op prestaties af te rekenen.¹⁷

In de jaren daarna werden vele benchmarks gepresenteerd. Aanvankelijk ontstonden die vrijwillig op terreinen van dienstverlening, zoals de uitgifte van identiteitsdocumenten of het ophalen van grofvuil. Sommige van deze benchmarks bestaan nog steeds, zoals die naar de kwaliteit van de riolering (sinds 2003) en het functioneren van gemeentelijke sociale diensten (sinds 2000).¹⁸

Benchmarking werd ook binnen de Rijksoverheid opgepakt. In 2002 werd de Rijksbrede Benchmarkgroep opgezet na discussie over de hoogte van de automatiseringskosten van de Informatie Beheer Groep (nu Dienst Uitvoering Onderwijs). Om in te kunnen schatten of die kosten hoog waren in vergelijking tot die van andere uitvoeringsorganisaties ontstond het idee om een benchmark van uitvoeringsorganisaties op te zetten. De eerste maal deden elf

uitvoeringsorganisaties mee, met een gezamenlijke omvang van 23.000 medewerkers die tezamen jaarlijks ruim €50 miljard aan subsidies en uitkeringen verstrekten. Het idee achter de vergelijking was dat alhoewel de organisaties sterk verschillende, financiële diensten leverden, ze in hun bedrijfsvoering ook veel vergelijkbare processen hadden, zoals telefonische bereikbaarheid en de afdoeningssnelheid van bezwaarschriften. Als basis werd het in de overheid toen veel gebruikte INK-model voor kwaliteitsmanagement gebruikt van het Instituut Nederlandse Kwaliteit. Dat model kenden vele organisaties al als middel om verbeteringsprocessen mee vorm te geven. Deze benchmark was een initiatief van de uitvoeringsorganisaties zelf: het leren van elkaars *best practices* stond nadrukkelijk voorop. In de jaren daarna is Rijksbrede Benchmarkgroep uitgegroeid tot een netwerkorganisatie van 41 organisaties in de publieke sector waar ongeveer 200.000 medewerkers werken.¹⁹

5.2 Van leren naar verantwoord: benchmarken als transparantie

In de jaren daarna raakte benchmarking binnen de Nederlandse overheid steeds verder weg van de leerfunctie en kwam er, zoals de bestuurskundige De Bruijn dit noemt, een opklimmende mate van dwang in de toepassing van het instrument: van leren en transparantie naar oordelen en afrekenen.²⁰

Gemeente- en provinciefinanciën: het Besluit Begroting en Verantwoording (BBV)

In 2002 en 2003 zijn provincie- en gemeentebesturen gedualiseerd, waardoor er meer afstand kwam tussen volksvertegenwoordiging enerzijds en het dagelijkse bestuur anderzijds. Hiermee kregen de provinciale staten en gemeenteraden nadrukkelijker de rol om kaders te stellen en het bestuur te controleren. Financiën spelen hierin een grote rol. Daarom is in 2003 het Besluit Begroting en Verantwoording provincies en gemeenten (BBV) ingegaan. Hierin stelt de minister van Binnenlandse Zaken eisen aan de vormgeving van de financiële rapportages.

Tot meer betrokkenheid van staten en gemeenteraden heeft dit niet geleid. Begrotingen bleven het privilege van financieel deskundigen. Na een kritisch rapport van de commissie-Depla van de Vereniging van Nederlandse Gemeenten en tegen een achtergrond van verdere Europese standaardisatie in de vorm van Europese publieke sector accountantstandaards (EPSAS) en de doorvoering van de decentralisaties in het sociale domein, introduceerde de minister van Binnenlandse Zaken uniforme beleidsindicatoren voor alle gemeenten en provincies. Het idee achter deze verplichte set van indicatoren is dat dit het voor gemeenteraden makkelijker maakt om de beleidseffecten van het inzetten van publiek geld te volgen.²¹

Zo ontstond op basis van het Besluit begroting en verantwoording provincies en gemeentes (BBV) een set beleidsindicatoren die in alle begrotingen en jaarverslagen van gemeenten en provincies moet worden gebruikt. Hierdoor kregen enerzijds gemeenteraadsleden beter inzicht in de prestaties van hun gemeenten en werden anderzijds prestaties van gemeenten onderling vergelijkbaar. Deze vergelijkbaarheid moest het debat in gemeenteraad, media en bij burgers en bedrijven stimuleren.²² Maar de Rijksoverheid zag het ook als middel om beter toezicht te kunnen houden op de prestaties en financiën van gemeenten en provincies.

Waarstaatjegemeente.nl

In de loop der jaren is het aantal beleidsindicatoren en het aantal beleidsterreinen van gemeenten waarover gerapporteerd wordt vergroot. Deze zijn terug te vinden op

Waarstaatjegemeente.nl, een website die wordt beheerd door de Vereniging van Nederlandse Gemeenten (VNG) en waarin gegevens over tal van beleidsterreinen per gemeente gebundeld worden. Zo zijn gemeenten onderling of ten opzichte van bepaalde kenmerken (inwonertal, verstedelijking, regio) te vergelijken in overzichtelijke dashboards.

Waarstaatjegemeente richt zich primair op bestuurders en ambtenaren, maar is ook voor burgers goed toegankelijk. Gemeenten zijn alleen verplicht om cijfers over de Wet Maatschappelijke Ondersteuning (WMO) aan te leveren. Tachtig procent van de ingevoerde gegevens komt van andere bronhouders zoals het Centraal Bureau voor de Statistiek, het Kadaster, Gemeentelijke Gezondheidsdiensten (GGD) en het Bureau Halt. De frequentie en actualiteit van de cijfers verschillen. Daarom is een beleidsterrein als gemeentelijke financiën waar de gemeenten jaarlijkse uniforme indicatoren moeten gebruiken makkelijker onderling te vergelijken dan bijvoorbeeld het dashboard jeugd, waar data uit diverse bronnen komen, die minder frequent worden gerapporteerd en meer contextgevoeligs zijn.²³

5.3 Benchmarken om te oordelen en af te rekenen.

Een goed voorbeeld dat de verandering van benchmark als vrijwillig leerinstrument tot wettelijk voorschrift illustreert, is de jaarlijkse benchmark van de Vereniging van Waterbedrijven in Nederland (VEWIN). Deze benchmark ontstond in 1997 als vrijwillig instrument om de efficiëntie van waterleidingbedrijven te vergroten. Zo was het de bedoeling dat deze monopolisten een prikkel kregen om efficiënter te opereren. In een terugblik op twintig jaar benchmarken tussen waterleidingbedrijven vertelde een directeur dat hij uit de eerste benchmark leerde dat zijn bedrijf behoorlijk duur was. Na een reorganisatie verdween de helft van het personeel. Sinds 2007 werd de benchmark wettelijk verplicht. Dat werd gedaan om de doelmatigheid te borgen, maar minstens even belangrijk waren transparantie en de mogelijkheid voor de politiek om via benchmarkverplichtingen te sturen.²⁴

Benchmarken in het onderwijs

De druk om wettelijk verplicht te benchmarken en om transparantie aan sturing te koppelen is breder zichtbaar. Zo nam de minister van Onderwijs, Cultuur en Wetenschappen een pleidooi van de Onderwijsraad uit 2018 over om schoolbesturen via een verplichte benchmark hun uitgaven te laten verantwoorden. De keuze voor deze benchmark die door de koepelorganisaties uit het onderwijs wordt opgesteld, volgt op jarenlange kritiek van de Onderwijsinspectie over de gebrekkige transparantie van bestedingen en het beeld dat sommige scholen te hoge reserves aanhouden. Dit komt deels door de zogenaamde lumpsum, een vaste hoeveelheidgeld waarvan de besturen zelf bepalen hoe ze die inzetten voor personeel en materieel. Dit maakt (bij)sturing vanuit het Rijk lastig.²⁵

In de nieuwe onderwijsbenchmark keren de inspectienormen voor financiën terug. De Inspectie van het Onderwijs maakt, sinds in 2002, toen de Wet op het Onderwijstoezicht in werking trad, steeds vaker onderwijsdata openbaar. Het stimuleren van scholen om betere prestaties te leveren is één van de fundamenten van deze wet. Via portals als scholenopdekaart.nl en onderwijsincijfers.nl zijn op basis van data van de inspectie en de Dienst Uitvoering Onderwijs prestaties van scholen met elkaar te vergelijken.²⁶

Recente fraudes met financiële benchmarks

Dat benchmarks een grote rol spelen en dat de objectiviteit en integriteit van de benchmarks daarbij eveneens van groot belang zijn, blijkt uit de recente fraude met de vaststelling van de EURIBOR rente op basis van financiële benchmarks voor prijsstelling op de geldmarkten. Aan deze benchmarks was in Europa €120.000 miljard aan hypotheek, bedrijfsleningen en rentederivaten gekoppeld. Om de kans te verkleinen dat er met deze private benchmarks kan worden geknoeid, heeft de Europese Commissie in een verordening kwaliteitseisen aan deze financiële benchmarks gesteld. De Autoriteit Financiële Markten houdt in Nederland toezicht op de naleving van deze eisen.²⁷

6. Internationaal gebruik van benchmarking

De universiteit van Warwick houdt de Global Benchmark Database bij, waarin wereldwijde benchmarks worden geregistreerd. Van de 304 wereldwijd uitgevoerde benchmarks die de database in 2020 bevatte, ontstonden er slechts 20 vóór de jaren negentig. In de jaren negentig werden er 34 wereldwijde benchmarks geïntroduceerd. Het eerste decennium van de 21^e eeuw – op het hoogtepunt van de interesse voor benchmarking – werden 147 uitgevoerd, waarna er tussen 2010 en 2019 nog eens 104 nieuwe benchmarks werden uitgevoerd. Veel van deze benchmarks hadden overigens een kort of zelfs eenmalig bestaan.

Deze benchmarks zijn onder te verdelen in vier typen: staatsmanschap, internationale governance, governance voor de zakelijke markt en als laatste internationale belangenbehartiging. Elk van deze typen benchmark wordt ideaaltypisch door een ander soort organisatie uitgevoerd.²⁸

Staatsmanschap

Benchmarks op het gebied van staatsmanschap worden opgesteld door organisaties van nationale overheden, vooral door ministeries van Buitenlandse Zaken of voor Ontwikkelingssamenwerking. Via deze benchmarks, die gaan over zaken als duurzame ontwikkeling, transparantie van begrotingen of bestrijding van mensensmokkel, worden nationale agenda's en waarden verspreid. Ze kunnen ook dienen om verdergaande diplomatieke acties zoals sancties te legitimeren. Het aantal van deze benchmarks is overigens beperkt.

Internationale governance

Dat geldt niet voor benchmarks met betrekking tot governance, die veelal worden opgesteld door intergouvernementele organisaties. Ze gebruiken benchmarking om lidstaten van elkaar te laten leren, maar ook vooral om druk te houden op hun leden om bepaalde doelen gericht na te blijven streven. De meest bekende, en wellicht ook belangrijkste, benchmark is op dit moment die van de Duurzame Ontwikkelingsdoelen van de Verenigde Naties. Aan deze inmiddels vijftien hoofddoelen zoals de uitbanning van honger en armoede in 2030, liggen 169 onderliggende acties ten grondslag die met 232 indicatoren worden gevolgd.²⁹ Andere bekende intergouvernementele organisaties, zoals de Europese Unie en de OESO maken ook intensief gebruik van benchmarking.

In 2000 stelden de Europese regeringsleiders de Lissabonstrategie vast. Hiermee wilde de Europese Unie in tien jaar tijd “de meest concurrerende en dynamische kenniseconomie van de wereld worden, die in staat zou zijn om duurzame groei, met meer en betere banen en een grotere sociale cohesie” te bereiken. Deze ambitieuze strategie zou niet top-down met

richtlijnen en verordeningen uit Brussel worden vormgegeven, maar via een nieuw open coördinatiemechanisme. Deze nieuwe methode moest er voor zorgen dat beleidsmakers internationaal informatie uitwisselden en van elkaars ervaringen konden leren. Dit moest ertoe leiden dat de nationale overheden zelf het eigenaarschap voelen en tegelijkertijd *peer pressure* zouden voelen om te blijven streven naar verbetering. Om dit te bereiken, greep de Europese Commissie naar het middel benchmarking, dat in de woorden van een evaluatie van de Europese Centrale Bank 'de steunpilaar voor deze coördinatie' werd. Binnen korte tijd ontstond er een grote set van indicatoren voor tal van beleidsterreinen. Tien jaar later bleek uit diverse evaluaties dat de ambitieuze doelen, zoals 70% werkgelegenheid en 3% van het BBP naar investeringen in onderzoek en ontwikkeling niet gehaald werden. Ook zonder de recessie die plaatsvond na de kredietcrisis van 2008, was het onwaarschijnlijk dat de doelen bereikt zouden worden. Wat bleef, was de aandacht voor benchmarking. Deze methode was na tien jaar gebruik niet meer weg te denken.³⁰

De OESO werd in 1961 opgericht als economische tegenhanger van de NAVO om economische groei, vrijhandel en werkgelegenheid in de lidstaten te bevorderen. Na de Koude Oorlog vond de organisatie zich opnieuw uit, breidde het ledental zich uit tot nu 36 landen en startte de organisatie een samenwerking met meer dan 100 zich ontwikkelende landen. In de jaren negentig ontwikkelde de OESO zich tot een organisatie die ook goed bestuur en goede beleidsproductie promootte. Dit deed ze bovendien op veel meer beleidsterreinen dan daarvoor. Gezondheidszorg, duurzaamheid en meest in het oog springend onderwijs vielen daar nu ook onder. De OESO kent geen verbindende regelgeving en werkt daarom vooral op basis van overtuiging door middel van adviezen van technische experts. De 200 verschillende commissies zijn het kloppend hart van de OESO. De 250 jaarlijkse publicaties maken de OESO één van 's werelds grootste uitgeverijen op het gebied van beleid en economie. De commissies volgen veelal de "OESO-werkwijze" die bestaat uit een peer review van een beleidsterrein door internationale experts. Benchmarking is hierbij een veel gebruikt instrument om lidstaten te stimuleren. Het bekendste voorbeeld hiervan is het sinds 2000 regelmatig uitgevoerde *Programme for International Student Assessment* (PISA), waarin de leesvaardigheid, rekenvaardigheid en kennis van natuurwetenschappen wereldwijd worden getoetst onder 15-jarigen. De presentatie van de ranking leidt internationaal tot veel debat in politiek en media.³¹

Governance voor de zakelijke markt

De derde categorie internationale benchmarks is heel breed. Dit zijn de benchmarks op het gebied van de zakelijke markt. Zij worden veelal opgesteld door internationale bedrijven in de zakelijke dienstverlening. Dit kunnen consultancybureaus zoals McKinsey zijn die het gebruiken om hun expertise te tonen en bedrijven in de financiële dienstverlening die het vestigingsklimaat, financieringsmogelijkheden of geldmarkten in kaart brengen.

Ook in de financiële pers groeide de aandacht voor benchmarking snel. Met name het Britse blad *The Economist* kreeg hierin een sterke reputatie. Dit blad publiceert al sinds 1986 de Big Mac-index, waarin de prijs van deze hamburger de ludieke indicator is om koopkracht tussen landen te vergelijken. Twee jaar later volgde de kwaliteit van leven-benchmark, de "where-to-be-born" index. Vanaf de start van de 21^e eeuw kreeg het aantal benchmarks vleugels. Een speciale afdeling, The Economist Intelligence Unit, publiceert er vele, van de meest leefbare stedenindex (sinds 2002) tot de Democracy Index (sinds 2006) die de staat van de

democratie in een land probeert te meten. The Economist staat met dertien benchmarks in de Global Benchmark Database.³²

Zo ontstond er een als maar groeiend aanbod van benchmarks, waarbij overheid en pers elkaar soms verstevigden. Denk aan de opkomst van de bekende universiteitsrankings. In 2003 ontstond dit idee nadat de *Treasury*, het Britse ministerie van Financiën, de Britse universiteiten had laten vergelijken om meer inzicht te krijgen in de kwaliteit van hun onderwijs en onderzoek. Zo ontstond in 2004 de eerste universiteitsranking van het Britse bedrijf QS en het Times Higher Education magazine, die sinds 2010 elk hun eigen ranking publiceren. Deze initiatieven groeiden uit tot een volwaardige industrietak. Niet in de laatste plaats dankzij de vele rankings van business schools en managementopleidingen, zoals de ranking van de *Financial Times* die sinds 2006 jaarlijks verschijnt.³³

Internationale belangenbehartiging

Internationaal groeide de aandacht voor benchmarking in de jaren negentig van de twintigste eeuw. Niet voor niets wordt dit het decennium van de transparantie genoemd. Eén van de bekendste benchmarks die in deze jaren ontstond, was de corruptie-perceptie-index van Transparency International (1995), een non-gouvernementele organisatie (NGO) die zich inzet om corruptie te bestrijden door transparantie te promoten. Deze is typerend voor de benchmarks voor internationale belangenbehartiging die door NGO's worden uitgevoerd.

Er zijn inmiddels veel van deze benchmarks. Zo introduceerde het Wereld Natuurfonds bijvoorbeeld de *Living Planet Index* (1997) waarin de biodiversiteit wordt onderzocht en Freedom House onderzoekt sinds 1995 de staat van de democratie wereldwijd. Voor NGO's zijn benchmarks een vorm van belangenbehartiging waarbij de organisatie met een schijn van rationaliteit hun zaak overzichtelijk kan bepleiten.

7. Samenvatting

Het uitgangspunt van een benchmarks is vergelijken om te leren. Benchmarken kun je op verschillende manieren doen: je kunt processen, prestaties of strategieën vergelijken. Dit kan zowel binnen één organisatie als tussen organisaties onderling. Benchmarken in de publieke sector is populair geworden door de opkomst van New Public Management en de roep om meer transparantie en de groei van transnationale governance.

Binnen de publieke sector moet je rekening houden met een aantal valkuilen: binnen de publieke sector hebben organisaties geen winstprikkel en spelen verschillende publieke waarden een rol. Ook wordt veel beleid in ketens gemaakt, mede hierdoor is effecten aan één organisatie toeschrijven lastig. Er kunnen perverse uitkomsten ontstaan wanneer er geen rekening wordt gehouden met de beperkingen van een benchmark.

In Nederland zien we een ontwikkeling van het gebruik van benchmarks als leerinstrument naar de toepassing als verantwoordings- en toezichtsinstrument. Internationaal neemt het gebruik van benchmarks toe. Er zijn vier typen internationale benchmarks te onderscheiden, namelijk met betrekking tot: staatsmanschap, internationale governance, governance voor de zakelijke markt en als laatste internationale belangenbehartiging.

Relevante literatuur

- Broome, A., & Quirk, J. (2015). Governing the world at a distance: the practice of global benchmarking. *Review of International Studies*, 41(5), 819.
- De Bruijn, H. (2007). *Managing performance in the public sector*. Routledge.
- Camp, R. C. (1989). *Benchmarking: The search for industry best practices that lead to superior performance*. Milwaukee : ASQ Quality Press.
- Groenendijk, Nico. (2011). EU and OECD Benchmarking and Peer Review Compared. In: Laursen, Finn (ed.) *The EU and Federalism. Politics and Policies Compared*. Ashgate: Routledge.
- Groot, de, H. (2004). *Benchmarking in de publieke sector: Vergelijken van prestaties als managementinstrument*. Den Haag: Sdu Uitgevers.
- Korsten, A. F. A., Abma, K., & Meer, A. D. (2019). *Het gras bij de burens: Benchmarking de maat genomen*. Amsterdam: BoomBestuurskunde.
- Raad voor het openbaar bestuur. (2002). *Presteren door leren: Benchmarks in het binnenlands bestuur*. Den Haag: Raad voor het openbaar bestuur.
- Thiel, Van, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance & Management Review*, 25(3), 267-281.

Tien opgaven of vragen

1. Welke vormen van benchmarking kun je onderscheiden?
2. Hoe voorkom je bij benchmarking dat je appels met peren vergelijkt?
3. Wat is het verschil tussen benchmarking en monitoring?
4. Kun je drie potentiële voordelen en drie potentiële nadelen van benchmarking noemen?
5. Noem drie oorzaken die benchmarks in publieke organisaties anders maakt dan benchmarking in private organisaties?
6. Welke drie oorzaken zien bestuurskundigen voor de opkomst van benchmarking in de publieke sector?
7. Doen de valkuilen voor internationale benchmarks zich in jouw ogen ook voor wanneer je samenwerkende gemeenten willen benchmarken?
8. Wat is de prestatie-paradox en op welke vier verschillende manieren kan deze zichtbaar worden?
9. Welke voorbeelden voor de opklimmende mate van dwang kun je onderscheiden in de toepassing van benchmarking door de Nederlandse overheid?
10. Welke vier typen van internationale benchmarks zijn er?

Biografie

Suggesties voor het trefwoordenregister

Benchmarking

monitor

New Public Management

performance management

prestatie-paradox

transnationale governance

transparantie waardenpluraliteit

-
- ¹ R.C. Camp 1989. *Benchmarking: The Search for Industry Best Practices that Leads to Superior Performance*, Milwaukee: ASQC Quality Press.
- ² Federico Del Giorgio Solfa, *Public Benchmarking: Contributions for subnational governments and Benchmarking Design*, Buenos Aires 2017; Tim Stapenhurst. *The benchmarking book. A How-to-Guide to Best Practice for Managers and Practitioners*. Elsevier Amsterdam, 2009.
- ³ Philip Waalewijn, Andre Hendriks & Ruben Verzijl, 1996, Benchmarking van het benchmarkingproces, Ribes rapport 9611/M, Erasmus Universiteit Rotterdam en M. de Haan en K.R.E. Huizingh, 1997, Benchmarking: een methode om te leren, *Maandblad voor Accountancy en Bedrijfseconomie*, 71 (4), p. 153-163
- ⁴ Peter Bogetoft. *Performance benchmarking: Measuring and managing performance*. Springer New York 2013.103-126.
- ⁵ K. Alexander, L. Elke, K. Helmut, & K.K. Nada, (1999). Benchmarking and performance measurement in public sectors. Towards learning for agency effectiveness. *The International Journal of Public Sector Management*, 12(2), 121-144.
- ⁶ Zie: www.coelo.nl
- ⁷ C. Hood. (1991). A public management for all seasons? *Public Administration*, 69(1), 3-19.
- ⁸ Anna van der Vleuten, & Mieke Verloo. (2012). Ranking and benchmarking: The political logic of new regulatory instruments in the fields of gender equality and anti-corruption. *Policy & Politics*, 40(1), 71-86.
- ⁹ Hans de Groot, 'Waarom benchmarking', in: Hans de groot (eds.) *Benchmarking in de publieke sector. Vergelijken van prestaties als managementinstrument*. Sdu, Den Haag, 2013, 9-15.
- ¹⁰ J.L.M. Hakvoort & H.L. Klaassen, *Bedrijfsvoeringstechnieken voor overheid en non-profitorganisaties* Sdu Uitgevers, Den Haag 2013, pp. 113-126.
- ¹¹ A. F. A. Korsten, K. Abma, & A.D. van der Meer. *Het gras bij de burenen: Benchmarking de maat genomen*. BoomBestuurskunde Amsterdam 2019. Een gratis versie is te downloaden via www.arnokorsten.nl
- ¹² Idem.
- ¹³ Sandra van Thiel & Frans Leeuw, (2002). The performance paradox in the public sector. *Public Performance & Management Review*, 25(3), 267-281.
- ¹⁴ André Broome & Joel Quirk. (2015). Governing the world at a distance: the practice of global benchmarking. *Review of International Studies*, 41(5), 819-841.
- ¹⁵ Nico Groenendijk. EU and OECD benchmarking and peer review compared. In: Finn Laurensen ed., *The EU and Federalism: Politics and Policies Compared*, Ashgate London 2010, 181-202.
- ¹⁶ N. Daniels, J. Bryant, R.A. Castano, O.G. Dantes, K.S. Khan & S. Pannarunothai. (2000). Benchmarks of fairness for health care reform: a policy tool for developing countries. *Bulletin of the World Health Organization*, 78, 740-750.
- ¹⁷ Raad voor het openbaar bestuur, *Presteren door leren: Benchmarks in het binnenlands bestuur*. Den Haag: Raad voor het openbaar bestuur 2002.
- ¹⁸ Zie: Maarten de Geest, De riolering in Nederland is op orde, *H2O/O2* (2011), 11-13; Benchmark sociale diensten 2001. SGB0 VNG Den Haag, 2003.
- ¹⁹ Peter Noordhoek (2005). Benchmarking en publieke verantwoording, *Bestuurskunde* 14:1, pp.33-39. Voor de huidige activiteiten van de Rijksbrede Benchmarkgroep: www.rbbgroep.nl.
- ²⁰ Hans de Bruijn, H. *Managing performance in the public sector*. Routledge Basingstoke 2007, p. 8.
- ²¹ Vereniging Nederlandse Gemeenten, Vernieuwing van de begroting en verantwoording van gemeenten, VNG Den Haag 2014.
- ²² www.commissiebbv.nl
- ²³ <https://www.waarstaatjegemeente.nl/content/over>
- ²⁴ M. de Goede, B. Enserink, G.I.M. Worm, & J.P. van der Hoek, (2016). Drivers for performance improvement originating from the Dutch drinking water benchmark. *Water Policy*, 18(5), 1247-1266 en z.a. Een benchmark van de benchmark *Helder* (2007) 7:2, . <https://doi.org/10.2166/wp.2016.125>
- ²⁵ Zie het pleidooi in: *Inzicht in en verantwoording van onderwijsgelden. Naar meer eenvoudige bekostiging en betere verantwoording van besteding van publieke middelen*. Onderwijsraad, Den Haag 2018.

²⁶ www.scholenopdekaart.nl en www.onderwijsincijfers.nl

²⁷ Zie: Benchmarks - Regulation (EU) 2016/1011 via www.ec.europa.eu

²⁸ André Broome & Joel Quirk. (2015). Governing the world at a distance: the practice of global benchmarking. *Review of International Studies*, 41(5), 819-841.

²⁹ Zie voor een kritische reflectie": S. Fukuda-Parr, A.E. Yamin, & J. Greenstein, (2013). Synthesis Paper-The power of numbers: A critical review of MDG targets for human development and human rights. Working Paper Series, Cambridge MA Harvard, 2013. Via:
http://fxb.harvard.edu/wpcontent/uploads/sites/5/2013/09/Synthesis-paper-PoN_Final1.pdf.

³⁰ Caroline De la Porte, Philippe Pochet & Graham Room. (2001). 'Social Benchmarking, Policy-Making and the Instruments of New Governance in the EU'. *Journal of European Social Policy*. 291-307; Nico Groenendijk. EU and OECD benchmarking and peer review compared. In: Finn Laurensen ed., *The EU and Federalism: Politics and Policies Compared*, Ashgate London 2010, 181-202.

³¹ Zie hiervoor: Christian Ydesen (Ed.). *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*. Palgrave Macmillan London 2013.

³² De Global Benchmark Database van de universiteit Warwick is te vinden op:

<https://warwick.ac.uk/fac/soc/pais/research/researchcentres/csgr/benchmarking/database/>

³³ Het rapport dat het idee van universiteitsrankings plantte, was: R. Lambert. *Lambert review of business-university collaboration*. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship. HM Treasury Londen 2003.