

# Benchmarking en prestatieverbetering in de publieke sector

## Wat is het verband en hoe kan dit worden versterkt?

Sandra Tillema

**SAMENVATTING** In dit artikel wordt beargumenteerd dat het verband tussen *benchmarking* en prestatieverbetering in de publieke sector afhangt van het doel van een bepaalde benchmarkoperatie. Dergelijke operaties kunnen bedoeld zijn om een managementinstrument te bieden, maar ook om externe druk (institutionele dan wel economische druk) te creëren of te versterken. Het artikel doet, voor elk van de bij deze doelen behorende situaties, aanbevelingen ten aanzien van de acties die kunnen worden ondernomen om het verband tussen benchmarking en prestatieverbetering te versterken. Belangrijke aspecten zijn: de selectie van prestatie-indicatoren, de vorm waarin de benchmarkingscores worden gepresenteerd, de belanghebbenden aan wie deze scores bekend worden gemaakt en de vervolgacties die worden ondernomen.

**RELEVANTIE VOOR DE PRAKTIJK** Het artikel biedt een handvat voor degenen die in de praktijk betrokken zijn bij het ontwerp van benchmarkoperaties. Zij worden gestimuleerd meer expliciet te maken via welke weg een bepaalde benchmarkoperatie moet leiden tot prestatieverbetering. Bovendien biedt het artikel aanknopingspunten voor het antwoord op de vraag welke ontwerpkenmerken hierbij aansluiten.

### 1 Inleiding

Benchmarking staat al enkele jaren volop in de belangstelling als techniek om de prestaties van organisaties in de publieke sector te verbeteren. Veel organisaties zijn recentelijk dan ook betrokken geweest bij benchmarkoperaties die expliciet gericht zijn op prestatieverbetering. Voorbeelden van organisaties die hebben deelgenomen aan dergelijke operaties zijn gemeenten, provincies, zorginstellingen en woningcorporaties (voor een overzicht, zie [www.centrumvoorgoodgovernance.nl/benchmarking](http://www.centrumvoorgoodgovernance.nl/benchmarking)). De gedachtegang achter veel van deze benchmarkoperaties is dat benchmarking kan leiden tot een verbetering van de prestaties in de publieke sector, omdat het een alternatief biedt voor concurrentie via de markt (Van Helden en Tillema, 2005, pp. 338-341). Dat wil zeggen, waar organisaties in de private sector een economische druk tot presta-

tieverbetering ondervinden wanneer hun prestaties matig zijn en klanten overstappen naar alternatieve aanbieders, wordt aangenomen dat organisaties in de publieke sector een institutionele druk tot prestatieverbetering ondervinden wanneer zij matige benchmarkingscores hebben. Deze institutionele druk is dan afkomstig van belanghebbenden, zoals bestuursorganen, toezichthouders of klanten.

Deze gedachtegang is aantrekkelijk, omdat ermee wordt gesuggereerd dat benchmarking ertoe kan leiden dat organisaties in de publieke sector meer aandacht gaan besteden aan de wensen van klanten en andere belanghebbenden. Hierbij wordt echter geen rekening gehouden met enkele complicerende factoren die in de praktijk van groot belang kunnen zijn. In dit artikel wordt benchmarking in de publieke sector daarom nog eens kritisch tegen het licht gehouden. Het artikel is geïnspireerd op de resultaten van case-onderzoek bij vier Nederlandse waterschappen (zie Van Helden en Tillema, 2005; Tillema, 2006; Tillema, 2007). Deze waterschappen namen deel aan de eerste sectorbrede benchmarkoperatie op het gebied van afvalwaterzuivering (de zogenaamde 'Bedrijfsvergelijking Zuiveringsbeheer 1999'; zie ook Admiraal en Van Helden, 2003). Uit het onderzoek bleek dat het verband tussen deze benchmarkoperatie en verbeteracties van de waterschappen niet erg sterk was, zelfs niet in situaties waarin de benchmarkingscores suggereerden dat er substantiële ruimte voor verbetering was.<sup>1</sup> De uitkomsten van het case-onderzoek riepen de vraag op hoe het verband tussen benchmarking en prestatieverbetering kan worden versterkt. Niet alleen voor de waterschappen, maar ook voor andere organisaties in de publieke sector. In dit artikel wordt geprobeerd deze vraag te beantwoorden. Meer specifiek wordt gekeken naar de aard van het verband tussen benchmarking en prestatieverbetering en de invloed daarvan op het ontwerp van een benchmarkoperatie. Daartoe wordt een theoretische analyse van drie typen situaties uitgevoerd:

- een situatie waarin benchmarking primair bedoeld is als managementinstrument;

- een situatie waarin de techniek primair bedoeld is om institutionele druk (belanghebbenden dringen aan op prestatieverbetering) te creëren of te versterken; en
- een situatie waarin ze primair bedoeld is om economische druk (klanten stappen over naar alternatieve aanbieders) te creëren of te versterken.

In de praktijk zijn uiteraard ook combinaties van de drie genoemde situaties mogelijk. Zo kan benchmarking in de gezondheidszorg gericht zijn op het versterken van zowel de institutionele druk (die zorginstellingen bijvoorbeeld ondervinden van toezichthouders) als de economische druk (aangezien zorginstellingen in toenemende mate te maken hebben met marktwerking). Om een scherpe analyse te kunnen uitvoeren, zullen dergelijke combinaties van situaties in dit artikel echter buiten beschouwing worden gelaten.

De opbouw van het artikel is als volgt. In paragraaf 2 wordt het verband tussen benchmarking en prestatieverbetering in elk van de drie hierboven geïntroduceerde situaties besproken. Vervolgens wordt in paragraaf 3 gekeken naar enkele vormen van ongewenst en defensief gedrag die kunnen voortvloeien uit benchmarking. Aansluitend wordt in paragraaf 4 aandacht besteed aan de vraag hoe dit gedrag kan worden beperkt. Daarmee worden de bouwstenen aangedragen voor de aanbevelingen die in paragraaf 5 worden gedaan om het gebruik van benchmarking-informatie voor prestatieverbetering in elk van de drie situaties te stimuleren. In paragraaf 6 wordt afgesloten met enkele aanvullende opmerkingen. In het artikel worden de resultaten van eerder onderzoek naar benchmarking in de publieke sector, inclusief de resultaten van het eerdergenoemde case-onderzoek onder Nederlandse waterschappen, gebruikt om de naar voren gebrachte argumenten te illustreren.

## 2 Benchmarking en prestatieverbetering

In de eerste plaats is het belangrijk te begrijpen *hoe* benchmarking kan leiden tot prestatieverbetering. De retoriek die vaak wordt gebruikt om benchmarking in de publieke sector te promoten, suggereert dat benchmarking organisaties helpt bij het vinden, met elkaar delen en implementeren van *best practices* (Northcott en Llewellyn, 2005, pp. 423-424). Daarmee lijkt het alsof organisaties in de publieke sector een vorm van benchmarking gebruiken die gericht is op het opdoen van *ideeën* met behulp waarvan prestatieverbeteringen kunnen worden gerealiseerd. Deze vorm van benchmarking houdt in dat een of meerdere organisaties worden geselecteerd met superieure prestaties op een bepaald gebied en dat vervolgens diepgaand wordt onderzocht welke bedrijfsprocessen ten grondslag liggen aan deze prestaties (Fedor e.a., 1996, pp. 162-163; Northcott en Llewellyn, 2005, pp. 423-424). In de praktijk wordt in de publieke sector echter vaak gebruikgemaakt van een vorm

van benchmarking die gericht is op *indicatoren* in plaats van *ideeën* (Ammons e.a., 2001, p. 101). Deze vorm van benchmarking is gericht op het meten van de prestaties van de eigen en verschillende andere organisaties, zonder dat er aandacht wordt besteed aan de bedrijfsprocessen waarmee deze prestaties worden behaald (Fedor e.a., 1996, pp. 162-163; Northcott en Llewellyn, 2005, pp. 423-424). Benchmarking van indicatoren kan daarmee wel leiden tot een aanscherping van de eigen doelstellingen op een bepaald gebied, maar het geeft geen inzicht in hoe deze aangescherpte doelstellingen kunnen worden gerealiseerd. Dit inzicht zou kunnen worden verkregen door aansluitend een benchmarking van *ideeën* uit te voeren, maar in plaats daarvan kan bijvoorbeeld ook worden gekozen voor het doorlichten van de eigen bedrijfsprocessen.

Vanwege het veelvuldig gebruik van benchmarking van indicatoren in de publieke sector staat deze vorm ook centraal in dit artikel. Hoewel benchmarking van indicatoren niet meteen leidt tot concrete *ideeën* omtrent mogelijkheden om de prestaties te verbeteren, kan deze vorm van benchmarking organisaties wel aanzetten tot het ontwikkelen en vervolgens implementeren van dergelijke *ideeën*. De aard van dit alternatieve verband tussen benchmarking en prestatieverbetering hangt samen met het doel van een bepaalde benchmarkoperatie.

### 2.1 Managementinstrument

Soms is benchmarking van indicatoren primair bedoeld om managers een instrument te bieden, waarmee ze een eerste indruk kunnen krijgen van de terreinen waar misschien prestatieverbeteringen mogelijk zijn. Managers kunnen gemotiveerd zijn om een dergelijk instrument te gebruiken, omdat bepaalde omstandigheden (bijvoorbeeld politici die op bepaalde beleidsterreinen verbeteringen eisen of een zwakke financiële positie van de organisatie) hen dwingen verbeteringen door te voeren. Het is echter ook mogelijk dat managers intrinsiek gemotiveerd zijn om prestatieverbeteringen te realiseren. In beide gevallen is het gebruik van benchmarking-informatie dus een reactie op een interne informatiebehoefte. Echter, nadat de benchmarkingscores bekend zijn gemaakt, moeten andere acties – bijvoorbeeld een benchmarking van *ideeën* – worden ondernomen om te ontdekken *hoe* de prestaties kunnen worden verbeterd.

Het gebruik van benchmarking als managementinstrument veronderstelt dus dat de managers van de betrokken organisaties reeds gemotiveerd zijn om de prestaties te verbeteren en om benchmarking-informatie voor dit doel te gebruiken. Als niet aan deze veronderstelling is voldaan, kan benchmarking van indicatoren worden gebruikt voor het creëren of versterken van externe druk om te verbeteren. Een voorwaarde is dan wel dat de benchmarkingscores ook bekend worden gemaakt aan belanghebbenden

van de organisatie. De externe druk die op deze manier kan ontstaan, kan zowel institutioneel als economisch van aard zijn.

## 2.2 Institutionele druk

In de context van dit artikel betekent institutionele druk dat belanghebbenden benchmarkingscores gebruiken om organisaties aan te sporen tot prestatieverbetering. Deze druk kan afkomstig zijn van zowel interne belanghebbenden (bijvoorbeeld bestuursorganen of leden van dergelijke organen) als externe belanghebbenden (bijvoorbeeld burgers, klanten, bekostigingsinstanties, toezichhouders of hogere overheden). Zij kunnen organisaties aanspreken op lage scores en een verklaring hiervoor verlangen, maar zij kunnen daarnaast ook eisen dat scores in de toekomst worden verbeterd. Eventueel kunnen ook sancties worden verbonden aan het niet realiseren van verbeteringen. Zo kan het toezicht worden geïntensiveerd of kan de bekostiging worden teruggebracht. Dergelijke vormen van institutionele druk kunnen managers vervolgens motiveren om de benchmarking-informatie te gebruiken als managementinstrument. Met andere woorden, hoewel benchmarking in deze situatie primair bedoeld is om de institutionele druk te versterken, kan een (waarschijnlijk vaak ook bedoeld) neveneffect ervan zijn dat managers de benchmarking-informatie gebruiken om in hun eigen informatiebehoefte te voorzien.

Een vereiste voor een succesvol gebruik van benchmarking als middel om de institutionele druk te versterken, is dat belanghebbenden zich in zekere mate betrokken voelen bij de desbetreffende organisaties en bereid zijn invloed uit te oefenen op de prestaties. Bij de Nederlandse waterschappen – die de inspiratiebron voor dit artikel vormden – bleek de invloed van belanghebbenden, en dan met name de externe belanghebbenden, tamelijk beperkt te zijn (Tillema, 2006, pp. 306-307; Tillema, 2007, pp. 511-512). Dit is één van de verklaringen voor het niet erg sterke verband tussen de eerste benchmarkoperatie op het gebied van afvalwaterzuivering en de door de waterschappen ondernomen verbeteracties. Voor sommige andere segmenten van de publieke sector mag worden verwacht dat de invloed van belanghebbenden veel sterker is. Hierbij kan worden gedacht aan de lokale overheid, het onderwijs en de gezondheidszorg. Omdat organisaties in deze segmenten bekend zijn bij een breed publiek en omdat zij het leven van veel mensen beïnvloeden, mag worden aangenomen dat zij kritischer worden gevolgd, bijvoorbeeld door de media. Onder deze omstandigheden kan het bekendmaken van benchmarkingscores tot aanzienlijke institutionele druk leiden.

## 2.3 Economische druk

Benchmarking in de publieke sector kan ook bedoeld zijn om de economische druk te versterken. In dat geval worden

de benchmarkingscores bekendgemaakt aan klanten van de betrokken organisaties, zodat zij een beter gefundeerde keuze tussen alternatieve aanbieders kunnen maken. In de context van dit artikel betekent economische druk dus dat klanten benchmarkingscores gebruiken bij het kiezen tussen aanbieders. Omdat klanten bij lage benchmarkingscores kunnen besluiten over te stappen naar een concurrerende aanbieder, kan prestatieverbetering in deze situatie een belangrijke voorwaarde zijn voor het behouden of verbeteren van de bestaande financiële positie van de organisatie. Dit kan zelfs gelden wanneer klanten de organisatie niet rechtstreeks betalen voor de afgenomen producten of diensten. Een terugloop van het aantal klanten betekent in die omstandigheden immers vaak dat de bekostiging zal worden teruggebracht. De nadelige financiële gevolgen van een geringer aantal klanten kunnen voor managers een sterke prikkel zijn om de benchmarking-informatie ook als managementinstrument te gebruiken.

Een vereiste voor een succesvol gebruik van benchmarking als middel om de economische druk te versterken, is dat klanten het recht hebben om hun eigen aanbieder te kiezen en om van aanbieder te veranderen. In bepaalde segmenten van de publieke sector hebben klanten altijd al de vrijheid om een aanbieder te kiezen, in ander segmenten is in toenemende mate sprake van keuzevrijheid. Voorbeelden van segmenten met keuzevrijheid voor klanten zijn het onderwijs (Adnett en Davies, 1999; Bowden, 2000; Broadbent en Laughlin, 1997; Broadbent e.a., 1999) en de gezondheidszorg (Ferlie e.a., 2006). Niet alleen hebben klanten in deze sectoren de vrijheid om een aanbieder te kiezen, ze worden ook in toenemende mate gestimuleerd om gebruik te maken van die vrijheid en om daarbij gebruik te maken van benchmarking-informatie (Adnett en Davies, 1999, p. 227; Bowden, 2000; Broadbent e.a., 1999, p. 344).

Benchmarking kan echter alleen effectief zijn als instrument om de economische druk te versterken indien een substantieel aantal klanten de benchmarkingscores inderdaad gebruikt om een aanbieder te kiezen. Het is de vraag of dit in de praktijk altijd het geval is. Klanten kunnen het lastig vinden om benchmarkscores te interpreteren en de wijze waarop de scores worden gepresenteerd sluit niet altijd aan bij hun individuele informatiebehoefte. Bowden (2000) geeft bijvoorbeeld aan dat voor een meerderheid van de aankomende studenten geldt dat ranglijsten van universiteiten hen niet de informatie verschaffen die zij nodig hebben bij de keuze van een universiteit. Daarnaast kunnen er voor klanten substantiële kosten verbonden zijn aan het veranderen van aanbieder. In dit verband merken Adnett en Davies (1999, p. 225) op dat ouders en leerlingen die tussentijds van school willen veranderen, met aanzienlijke kosten geconfronteerd kunnen worden. Bovendien zijn er binnen een redelijke

reisafstand niet altijd alternatieve aanbieders te vinden. Onder dergelijke omstandigheden is het onwaarschijnlijk dat het bekendmaken van benchmarkingscores resulteert in een sterke economische druk.

### 3 Ongewenst en defensief gedrag

In paragraaf 2 werd beargumenteerd dat benchmarking van indicatoren kan worden gebruikt om de externe druk tot prestatieverbetering te versterken. Een voorwaarde was wel dat de benchmarkingscores bekend worden gemaakt aan interne en/of externe belanghebbenden van de organisatie. Er zijn echter ook aanwijzingen dat juist wanneer de benchmarkingscores worden verstrekt aan belanghebbenden, het lastig is een sterk verband tussen benchmarking en prestatieverbetering te realiseren. Dit heeft te maken met een neiging tot ongewenst en/of defensief gedrag.

Ongewenst gedrag houdt in dat organisaties acties ondernemen die leiden tot een verbetering van hun *gerapporteerde* prestaties (hier: de benchmarkingscores), maar die hun *feitelijke* prestaties – in het beste geval – ongemoeid laten. Zo kan een school de gemiddelde score op een landelijke toets (de gerapporteerde prestatie) verhogen door minder goede leerlingen niet te laten deelnemen aan deze toets, maar daarmee is de kwaliteit van het onderwijs (de feitelijke prestatie) nog niet verbeterd. Omdat benchmarking van indicatoren kan worden gezien als een vorm van prestatiemeting, mag worden aangenomen dat het tot vergelijkbare typen van ongewenst gedrag kan leiden als andere vormen van prestatiemeting. Smith (1995) beschrijft een aantal van deze typen, waaronder tunnelvisie, suboptimalisatie, bijziendheid en een fixatie op indicatoren. Wiggins en Tymms (2002, pp. 45-46) signaleren dat benchmarking bij Engelse scholen leidt tot verscheidene typen van ongewenst gedrag. Zo laten hun resultaten zien dat deze scholen geneigd zijn zich te concentreren op die aspecten van hun prestaties die ook onderdeel uitmaken van een benchmarking. Dit gaat ten koste van andere belangrijke aspecten. De auteurs merken op dat leerlingen hier tijdens vervolgoopleidingen de negatieve gevolgen van kunnen ondervinden, omdat dan kan blijken dat zij bepaalde vaardigheden, die niet werden meegenomen in een benchmarking, onvoldoende beheersen.

Het aan belanghebbenden bekendmaken van benchmarkingscores kan ook leiden tot defensief gedrag (Bowerman e.a., 2002, pp. 434-435). Na de publicatie van benchmarkingscores kan een organisatie bijvoorbeeld proberen haar belanghebbenden ervan te overtuigen dat de lage scores haar niet zijn aan te rekenen omdat de eigen organisatie in bepaalde belangrijke opzichten niet goed vergelijkbaar is met de benchmarkingpartners (Northcott en Llewellyn, 2005, p. 426). In plaats daarvan kan de organisatie proberen belanghebbenden te bewegen zich uitsluitend of vooral te

richten op die indicatoren waarop de organisatie goed scoort (Bowden, 2000, p. 58). Het eindresultaat van deze typen van defensief gedrag kan zijn dat de motivatie om de benchmarkingscores te gebruiken als startpunt voor prestatieverbeteringen, zeer gering is. Er mag immers worden verwacht dat wanneer een organisatie heeft 'aangetoond' dat de matige benchmarkingscores haar niet zijn aan te rekenen, of dat deze scores veel minder belangrijk zijn dan andere, betere scores, ze zich minder genoodzaakt ziet om de prestaties te verbeteren. Eén van de waterschappen in ons case-onderzoek reageerde bijvoorbeeld op de publicatie van de benchmarkingscores door een persconferentie te organiseren, waarin verschillende verklaringen voor haar eigen, vaak tegenvallende, scores werden gegeven. Vervolgens werden deze scores echter niet of nauwelijks gebruikt als vertrekpunt voor verbeteracties (Tillema, 2006, p. 305; Tillema, 2007, p. 500).

### 4 Het beperken van ongewenst en defensief gedrag

In organisaties waarin ongewenst en/of defensief gedrag de overhand heeft, zal benchmarking niet het gewenste effect (dat wil zeggen een verbetering van de *feitelijke* prestaties) hebben. Het is daarom belangrijk de neiging tot dergelijk gedrag zoveel mogelijk in te perken. Eén van de bouwstenen voor het vinden van mogelijkheden daartoe, is het besef dat organisaties meer geneigd zijn tot ongewenst en/of defensief gedrag indien belanghebbenden externe druk uitoefenen op basis van een onjuiste interpretatie van de benchmarkingscores. Zo'n onjuiste interpretatie ontstaat wanneer belanghebbenden niet goed doorzien wat het effect op de benchmarkingscores is van fundamentele verschillen tussen organisaties, bijvoorbeeld in termen van doelstellingen of omgevingsfactoren (verg. Smith, 1995, pp. 294-297). Indien belanghebbenden zich niet bewust zijn van onjuistheden in hun interpretatie van de benchmarkingscores kunnen ze onterecht druk op organisaties uitoefenen om de ogenschijnlijke prestatiekloven te verkleinen. In dergelijke situaties mag misschien eerder worden verwacht dat organisaties defensief reageren dan wel ongewenste acties ondernemen.

Een tweede bouwsteen voor het beperken van ongewenst en/of defensief gedrag is het besef dat de mate waarin belanghebbenden benchmarkingscores onjuist interpreteren waarschijnlijk samenhangt met de vorm waarin deze scores worden gepresenteerd. De kans op een onjuiste interpretatie kan bijvoorbeeld groter zijn wanneer benchmarkingscores worden gepresenteerd in de vorm van ranglijsten, waarbij wordt geprobeerd verschillende aspecten van de prestaties van een organisatie te 'vangen' in één enkel getal. In de literatuur wordt veel kritiek geuit op het gebruik van ranglijsten. Belangrijke tekortkomingen die daarbij aan de orde komen, zijn de arbitraire weging van prestatie-indicatoren, het feit dat vaak geen onderscheid



wordt gemaakt tussen input en output, het feit dat regelmatig een vergelijking wordt gemaakt tussen organisaties die fundamenteel van elkaar verschillen en de afwezigheid van een heldere standaard voor een acceptabel prestatie-niveau (Northcott en Llewellyn, 2005, p. 426; Turner, 2005; zie ook Bowden, 2000, pp. 52-56). Ondanks deze kritiek worden ranglijsten regelmatig toegepast in verschillende segmenten van de publieke sector, waaronder de gezondheidszorg en het onderwijs.

Een laatste bouwsteen is het besef dat de mate waarin scores onjuist worden geïnterpreteerd waarschijnlijk verschilt tussen groepen belanghebbenden. Wanneer benchmarkingscores in een vergelijkbare vorm worden gepresenteerd aan verschillende groepen belanghebbenden mag worden aangenomen dat er bij belanghebbenden die geen ingewijden zijn op het werkerrein van de organisatie (bijvoorbeeld individuele klanten en andere burgers), in sterkere mate sprake is van een onjuiste interpretatie van scores, dan bij belanghebbenden die kunnen worden aangemerkt als expert (bijvoorbeeld leden van bestuursorganen of toezichthouders; Smith, 1995, pp. 294-297). Op basis hiervan mag worden verwacht dat de neiging van een organisatie om de eigen benchmarkingscores te verdedigen of om ongewenst gedrag te vertonen, groter is wanneer de benchmarkingscores worden bekendgemaakt aan een breder, gemiddeld minder ingewijd publiek. Wiggins en Tymms (2002, pp. 45-46) vinden bijvoorbeeld dat Engelse basisscholen – vergeleken met hun Schotse tegenhangers – meer geneigd zijn tot ongewenst gedrag. De auteurs suggereren dat dit verschil gerelateerd is aan een verschil in publicatiebeleid: de benchmarkingscores van Engelse

scholen worden gebruikt om ranglijsten op te stellen, die vervolgens worden gepubliceerd, terwijl de scores van Schotse scholen niet openbaar worden gemaakt (zie ook Broadbent e.a., 1999).

## 5 Het versterken van het verband

In deze paragraaf wordt besproken hoe het verband tussen benchmarking en prestatieverbetering kan worden versterkt in elk van de drie situaties die zijn besproken in paragraaf 2. In tabel 1 wordt een overzicht gegeven van deze situaties, de eisen die in elke situatie worden gesteld aan de organisatie en/of de omgeving, de bijbehorende kenmerken van het gebruik van benchmarking-informatie en enkele aanbevelingen voor het ontwerp van een benchmarkoperatie. Ten aanzien van het gebruik van benchmarking-informatie wordt onderscheid gemaakt tussen de verwachte focus (wordt vooral geprobeerd de feitelijke prestaties te verbeteren of gaat de aandacht meer uit naar de gerapporteerde prestaties?) en het verwachte type gedrag (is dit vooral gewenst gedrag, namelijk gedrag dat moet leiden tot een verbetering van de feitelijke prestaties, of spelen ook ongewenst en defensief gedrag een belangrijke rol?). Tabel 1 wordt hierna nader besproken.

### 5.1 Managementinstrument

Indien de managers van een organisatie gemotiveerd zijn om de feitelijke prestaties van de organisatie te verbeteren en bereid zijn hiervoor gebruik te maken van benchmarking-informatie, is het verstandig benchmarkingscores niet bekend te maken aan interne en externe belanghebbenden, anders dan de managers zelf. Op deze manier kan de

**Tabel 1** Overzicht van de drie situaties

Situatie	Vereiste	Wijze waarop benchmarking-informatie naar verwachting wordt gebruikt		Aanbevelingen
		Focus	Type gedrag	
Bedoeld als managementinstrument	Managers zijn gemotiveerd om prestaties te verbeteren	Feitelijke prestaties	Gewenst	<ul style="list-style-type: none"> <li>• Scores niet bekendmaken aan belanghebbenden</li> <li>• Benchmarking van indicatoren vervangen door, of aanvullen met, benchmarking van ideeën</li> </ul>
Bedoeld om de institutionele druk te versterken	Belanghebbenden zijn bereid invloed op organisaties uit te oefenen	Gerapporteerde prestaties	Gewenst, ongewenst en defensief	<ul style="list-style-type: none"> <li>• Scores alleen bekendmaken aan belanghebbenden die kunnen worden aangemerkt als expert</li> <li>• Benchmarking van indicatoren gevolgd door benchmarking van ideeën</li> </ul>
Bedoeld om de economische druk te versterken	Klanten hebben (en gebruiken) recht om aanbieder te kiezen en van aanbieder te veranderen	Gerapporteerde prestaties	Gewenst, ongewenst, en – mogelijk – defensief	<ul style="list-style-type: none"> <li>• Scores bekendmaken aan alle belanghebbenden</li> <li>• Zorgvuldig ontworpen set prestatie-indicatoren (met niet alleen gemakkelijk te meten indicatoren) en zorgvuldige presentatie van benchmarkingscores (waarbij rekening wordt gehouden met verschillen tussen organisaties; geen ranglijsten)</li> <li>• Benchmarking van indicatoren gevolgd door interne of <i>best-in-class</i> benchmarking van ideeën</li> </ul>

neiging tot ongewenst en/of defensief gedrag worden beperkt. Een belangrijke beperking die optreedt bij het gebruik van benchmarking als managementinstrument is echter dat veel benchmarkoperaties in de publieke sector gericht zijn op indicatoren in plaats van ideeën. Hoewel benchmarking van indicatoren een eerste indruk kan geven van terreinen waar misschien verbeteringen mogelijk zijn, is een goed begrip van de bedrijfsprocessen noodzakelijk om die prestatieverbeteringen vervolgens ook te realiseren. Om deze reden is het belangrijk dat benchmarking van indicatoren wordt vervangen door of aangevuld met benchmarking van ideeën. In de praktijk worden bijvoorbeeld wel benchmarkingclubs gevormd, waarbinnen informatie wordt uitgewisseld over de bedrijfsprocessen die mogelijk ten grondslag liggen aan de hogere benchmarkingscores (verg. Northcott en Llewellyn, 2005, pp. 428-429).

## 5.2 Institutionele druk

Indien de managers van een organisatie niet erg gemotiveerd lijken om verbeteracties te ondernemen, kan benchmarking van indicatoren worden gebruikt om de institutionele druk tot prestatieverbetering te versterken. Deze situatie vereist wel dat de benchmarkingscores bekend worden gemaakt aan belanghebbenden van de organisatie. Dit kan echter ook de neiging tot ongewenst en/of defensief gedrag vergroten. Deze neiging kan worden beperkt door de benchmarkingscores alleen aan meer ingewijde belanghebbenden bekend te maken. Hierbij kan worden gedacht aan bestuursorganen, bekostigingsinstanties en toezichhouders (verg. Smith, 1995, pp. 294-297). Omdat deze belanghebbenden beter in staat zijn de benchmarkingscores te beoordelen, en een onjuiste interpretatie dus minder waarschijnlijk is, is de kans op ongewenst en/of defensief gedrag minder groot. Bovendien is dergelijk gedrag richting meer ingewijde belanghebbenden minder effectief, omdat zij er sneller 'doorheen zullen prikken'.

Het op beperktere schaal bekendmaken van benchmarkingscores kan uiteraard betekenen dat organisaties in geringere mate zijn blootgesteld aan institutionele druk tot prestatieverbetering. Het is echter ook de vraag in hoeverre minder ingewijde belanghebbenden over het algemeen in staat zijn om invloed uit te oefenen op de beslissing van een organisatie om al dan niet verbeteracties te initiëren. Er zijn bijvoorbeeld aanwijzingen dat de mate waarin individuele klanten benchmarkingscores gebruiken om institutionele druk tot prestatieverbetering uit te oefenen, beperkt is. Zo vinden Wiggins en Tymms (2002, p. 44) dat – ondanks de media-aandacht die er in Engeland bestaat voor ranglijsten van scholen – Engelse ouders niet geneigd zijn benchmarkingscores te bespreken met de leerkrachten van hun kinderen. Deze auteurs suggereren dat druk uitgeoefend door de onderwijsinspectie veel effectiever is.

Meer in zijn algemeenheid mag worden verwacht dat, in vergelijking tot individuele klanten en het algemene publiek, bepaalde organen op het werkteer van een organisatie (variërend van het schoolbestuur of de gemeenteraad tot een bekostigingsinstantie of toezichhouder) een sterkere invloed kunnen uitoefenen op het gebruik dat deze organisatie maakt van benchmarking-informatie. Meestal geldt dat dergelijke organen niet alleen meer belangstelling hebben voor de benchmarkingscores van een organisatie, maar ook dat zij meer machtsmiddelen hebben om prestatieverbeteringen af te dwingen. Zo kunnen ze bijvoorbeeld hun bemoeienis vergroten met organisaties die er niet in slagen hun benchmarkingscores te verbeteren. Bepaalde organen kunnen soms zelfs afdwingen dat activiteiten op een andere wijze worden georganiseerd en/of aangestuurd (verg. Llewellyn en Northcott, 2005, pp. 558-559).

Wanneer een toename van de institutionele druk managers stimuleert om benchmarking-informatie te gebruiken als managementinstrument, is het ook in deze situatie vaak nodig om de benchmarking van indicatoren aan te vullen met een of meerdere vervolgacties. Hierdoor wordt meer inzicht verkregen in de manier waarop de prestaties kunnen worden verbeterd. Een benchmarking van ideeën kan dan zinvol zijn.

## 5.3 Economische druk

Wanneer benchmarking van indicatoren wordt gebruikt om de economische druk te versterken, moeten de benchmarkingscores bekend worden gemaakt aan minstens één groep van vaak minder ingewijde belanghebbenden, namelijk de klanten van de organisatie. Omdat onder niet-ingewijde belanghebbenden het onjuist interpreteren van benchmarkingscores meer waarschijnlijk is, ligt juist in deze situatie het gevaar van ongewenst en/of defensief gedrag op de loer. Organisaties zullen in deze situatie vooral geneigd zijn hun toevlucht te zoeken tot ongewenst gedrag, omdat het verbeteren van de gerapporteerde prestaties nu van cruciaal belang kan zijn om de financiële positie van de organisatie te behouden of te verbeteren. Om de neiging tot dergelijk gedrag te beperken, is het belangrijk dat de ontwerpers van de benchmarkoperatie veel aandacht besteden aan de selectie van prestatie-indicatoren en aan de vorm waarin de benchmarkingscores zullen worden gepresenteerd.

Bij de selectie van prestatie-indicatoren is het belangrijk dat niet alleen indicatoren worden gekozen die gemakkelijk kunnen worden gemeten (zoals de toetsresultaten van scholen), maar dat ook minder zichtbare (bijvoorbeeld meer op de lange termijn gerichte of meer subjectieve) aspecten van de prestaties van een organisatie (zoals de verhouding tussen leerlingen onderling en met leerkrachten) worden meegenomen (verg. Broadbent en

Laughlin, 1997, pp. 285-288). Bij het ontwikkelen van de vorm waarin de benchmarkingscores zullen worden gepresenteerd is het belangrijk dat klanten meer bewust worden gemaakt van mogelijke verklaringen voor verschillen tussen de benchmarkingscores van verschillende organisaties (vgl. Smith, 1995, pp. 296-297). Dit kan bijvoorbeeld worden bereikt door targets te formuleren die rekening houden met de doelstellingen en omgevingskenmerken van een organisatie. Een belangrijk omgevingskenmerk is bijvoorbeeld het 'niveau van complexiteit' waarmee een klant een organisatie binnenkomt, zoals de startkwalificaties van een leerling van een school of de complexiteit van de ziekte van een patiënt van een ziekenhuis. Een mogelijk eenvoudiger alternatief voor het formuleren van achtergrondspecifieke targets, is het presenteren van de benchmarkingscores per groep van organisaties, waarbij elke groep bestaat uit organisaties met vergelijkbare doelstellingen en omgevingskenmerken (vgl. Bowden, 2000, pp. 57-58; Turner, 2005, pp. 369-370). Beide presentatievormen zouden de neiging tot ongewenst en/of defensief gedrag kunnen reduceren. Bovendien laten ze meer ruimte voor verschillen tussen organisaties voor wat betreft de organisatiedoelstellingen, hetgeen ook meer keuzemogelijkheden voor klanten zou kunnen betekenen<sup>2</sup>. Tot slot lijkt het, gegeven de tekortkomingen van ranglijsten die zijn genoemd in paragraaf 4, verstandig het gebruik hiervan zoveel mogelijk te vermijden.<sup>3</sup>

Indien een toename van de economische druk managers stimuleert om benchmarking-informatie als management-instrument te gebruiken, is benchmarking van ideeën niet meteen een voor de hand liggende vervolgstap. Aangezien de organisaties in het specifieke segment van de publieke sector nu elkaars concurrenten zijn, zullen ze niet snel bereid zijn informatie over best practices met elkaar te delen, bijvoorbeeld via benchmarkingclubs. Dit wil echter niet zeggen dat benchmarking van ideeën in deze situatie geen optie is. Het is echter wel van belang dat de benchmarkingpartners niet afkomstig zijn uit hetzelfde segment van de publieke sector. Mogelijkheden zijn dan interne benchmarking, waarbij vergelijkbare processen van verschillende afdelingen binnen de eigen organisatie met elkaar worden vergeleken, en best-in-class-benchmarking, waarbij benchmarkingpartners niet worden geselecteerd op grond van het feit dat ze vergelijkbare producten of

diensten aanbieden, maar omdat ze een bepaalde functie of een bepaald proces buitengewoon goed uitvoeren (verg. Fedor e.a., 1996, pp. 163-164).

## 6 Afsluitende opmerkingen

In dit artikel werd beargumenteerd dat het verband tussen benchmarking en prestatieverbetering niet recht-toe-rechtaan is. Het verband en de acties die kunnen worden genomen om dit te versterken, verschillen van situatie tot situatie. Een actie die echter in alle situaties in overweging kan worden genomen, is het vervangen of aanvullen van benchmarking van indicatoren met benchmarking van ideeën. De laatstgenoemde vorm van benchmarking leidt tot meer inzicht in de specifieke bedrijfsprocessen die leiden tot betere prestaties. Dit betekent echter niet dat benchmarking van indicatoren niet zinvol kan zijn. In de publieke sector kan deze techniek worden gebruikt om de externe druk tot prestatieverbetering te versterken. Om dit te bereiken, is het echter belangrijk dat het ontwerp van een benchmarkoperatie weldoordacht is. Relevante aspecten zijn daarbij: de selectie van prestatie-indicatoren, de vorm waarin de benchmarkingscores worden gepresenteerd, de belanghebbenden waaraan deze scores worden bekendgemaakt en de vervolgacties die zullen worden ondernomen. In de praktijk zal het een grote uitdaging zijn om een benchmarkoperatie zo te ontwerpen dat de gewenste effecten van benchmarking worden versterkt, terwijl tegelijkertijd de neiging tot ongewenst en/of defensief gedrag wordt beperkt. Hopelijk biedt dit artikel aanknopingspunten voor het succesvol aangaan van deze uitdaging. ■

Dr. S. Tillema is universitair docent Accounting en Financiering bij de Faculteit Economie en Bedrijfskunde van de Rijksuniversiteit Groningen. Haar onderzoek richt zich met name op het gebruik van management accounting in de publieke sector. Dit artikel is gebaseerd op een Engelstalig artikel dat in 2009 verschijnt in *Public Money & Management* ([www.cipfa.org.uk/pt/pmm](http://www.cipfa.org.uk/pt/pmm)). De auteur dankt Jan van Helden en Bob Scapens voor hun commentaar en suggesties naar aanleiding van eerdere versies van dit artikel.

## Literatuur

- Admiraal, R.J. en G.J. van Helden (2003), De balanced scorecard binnen een benchmarkingcontext: Bedrijfsvergelijking zuiveringsbeheer bij Nederlandse waterschappen, *Maandblad voor Accountancy en Bedrijfseconomie*, vol. 77, no. 6, pp. 279-286.
- Adnett, N. en P. Davies (1999), Schooling quasi-markets: Reconciling economic and sociological analyses, *British Journal of Educational Studies*, vol. 47, no. 3, pp. 221-234.
- Ammons, D.N., C. Coe en M. Lombarda (2001), Performance-comparison projects in local government: Participants' perspectives, *Public Administration Review*, vol. 61, no. 1, pp. 100-110.
- Bowden, R. (2000), Fantasy higher education: University and college league tables, *Quality in Higher Education*, vol. 6, no. 1, pp. 41-60.
- Bowerman, M. en A. Ball (2000), Great expectations: Benchmarking for best value, *Public Money & Management*, vol. 20, no. 2, pp. 21-26.
- Bowerman, M., G. Francis, A. Ball en J. Fry (2002), The evolution of benchmarking in UK local authorities, *Benchmarking: An International Journal*, vol. 9, no. 5, pp. 429-449.
- Broadbent, J., K. Jacobs en R. Laughlin (1999), Comparing schools in the UK and New Zealand: Individualizing and socializing accountabilities and some implications for management control, *Management Accounting Research*, vol. 10, no. 4, pp. 339-361.
- Broadbent, J. en R. Laughlin (1997), Contracts and competition?: A reflection on the nature and effects of recent legislation on modes of control in schools, *Cambridge Journal of Economics*, vol. 21, no. 2, pp. 277-290.
- Fedor, D.B., C.K. Parsons en C.E. Shalley (1996), Organizational comparison processes: Investigating the adoption and impact of benchmarking-related activities, *Journal of Quality Management*, vol. 1, no. 2, pp. 161-192.
- Ferlie, E., G. Freeman, J. McDonnell, C. Petsoulas en S. Rundle-Smith (2006), Introducing choice in the public services: Some supply-side issues, *Public Money & Management*, vol. 26, no. 1, pp. 63-72.
- Helden, G.J. van en N. Brouwer (2005), Benchmarking en prestatieverbetering in de publieke sector: Ervaren met de bedrijfsvergelijking zuiveringsbeheer, *Maandblad voor Accountancy en Bedrijfseconomie*, vol. 79, no. 11, pp. 574-583.
- Helden, G.J. van en S. Tillema (2005), In search of a benchmarking theory for the public sector, *Financial Accountability & Management*, vol. 21, no. 3, pp. 337-361.
- Llewellyn, S. en D. Northcott (2005), The average hospital, *Accounting Organizations and Society*, vol. 30, no. 6, pp. 555-583.
- Northcott, D. en S. Llewellyn (2005), Benchmarking in UK health: A gap between policy and practice?, *Benchmarking: An International Journal*, vol. 12, no. 5, pp. 419-435.
- Smith, P. (1995), On the unintended consequences of publishing performance data in the public sector, *International Journal of Public Administration*, vol. 18, no. 2/3, pp. 277-310.
- Tillema, S. (2006), Gebruik van benchmarkinginformatie in de publieke sector: Case-onderzoek onder waterschappen naar het gebruik van benchmarkinginformatie voor prestatieverbetering, *Maandblad voor Accountancy en Bedrijfseconomie*, vol. 80, no. 6, pp. 300-308.
- Tillema, S. (2007), Public sector organizations' use of benchmarking information for performance improvement: Theoretical analysis and explorative case studies in Dutch water boards, *Public Performance & Management Review*, vol. 30, no. 4, pp. 496-520.
- Turner, D. (2005), Benchmarking in universities: League tables revisited, *Oxford Review of Education*, vol. 31, no. 3, pp. 353-371.
- Wiggins, A. en P. Tymms (2002), Dysfunctional effects of league tables: A comparison between English and Scottish primary schools, *Public Money & Management*, vol. 22, no. 1, pp. 43-48.

Geraadpleegde website:  
[www.centrumvoorgoedgovernance.nl/benchmarking](http://www.centrumvoorgoedgovernance.nl/benchmarking) (11 september 2008)

## Noten

**1** Inmiddels hebben ook benchmarkoperaties over de jaren 2002 en 2006 plaatsgevonden. Het gebruik van de informatie die dit heeft opgeleverd, komt niet in dit artikel aan de orde. Zie hiervoor bijvoorbeeld Van Helden en Brouwer (2005).

**2** Binnen een segment van de publieke sector, kan het reageren op externe druk tot het verbeteren van benchmarkingscores ertoe leiden dat de (gerapporteerde) verschillen tussen organisaties kleiner worden (Smith, 1995, p. 295). Wanneer we bijvoorbeeld een organisatie met een differentiatie-strategie vergelijken met een organisatie met een kostenleiderschapsstrategie, dan zal, na de publicatie van benchmarkingscores, de eerstgenoemde organisatie mogelijk worden aangespoord tot kostenreduc-

tie, terwijl de laatstgenoemde organisatie druk kan ervaren om de reeks producten uit te breiden. Als beide organisaties op deze druk reageren, zullen de verschillen tussen beide kleiner worden. Het ironische van dit eindresultaat is dat deze grotere uniformiteit in scherp contrast staat met de marktretoerik die wel wordt gebruikt om hervormingen in de publieke sector (inclusief de introductie van benchmarking) te rechtvaardigen. Deze hervormingen zouden immers leiden tot meer keuzevrijheid en meer keuzemogelijkheden voor burgers (Broadbent en Laughlin, 1997, p. 289). Een segment waar een dergelijke tendens tot meer uniformiteit reeds is waargenomen, is het onderwijs in het Verenigd Koninkrijk (Broadbent en Laughlin, 1997, p. 289).

**3** Soms is het gebruik van ranglijsten niet te vermijden. In sommige segmenten van de publieke sector – bijvoorbeeld universiteiten – worden ranglijsten opgesteld door een externe partij, bijvoorbeeld een krant of een tijdschrift (Bowden, 2000; Turner, 2005). Hoewel zulke ranglijsten geen officiële status hebben, kunnen de effecten ervan vergelijkbaar zijn met die van de publicatie van meer officiële ranglijsten.