

door Hans de Bruijn

Meetbaarheid en maakbaarheid

Over de perverse gevolgen van prestatie meting

De Academische Boekengids 45, juli 2004, pp. 9-11.

Het 'afrekenen op resultaten' bevordert niet altijd de kwaliteit van het werk, noch in het bedrijfsleven noch in de publieke ruimte. Ook op dit terrein staan cijferfundamentalisten tegenover nuchtere waarnemers.

In de afgelopen jaren zijn overal in de publieke sector systemen van prestatie meting geïntroduceerd. Rechters worden afgerekend op aantallen vonnissen, wetenschappers op hun productiescore en de NS op de punctualiteit van de dienstverlening. De Nederlandse politie heeft haar prestatiecontract en de minister van Vreemdelingenzaken wil een systeem ontwikkelen om de mate van integratie van allochtonen te meten.

De stellingen over de wenselijkheid van deze ontwikkeling zijn snel betrokken. Een eerste, veelgehoord, standpunt is dat van de *believers*. Die pleiten voor 'resultaatgerichtheid', 'transparantie' en 'responsiviteit'. De gedachte is: de professional in het publieke domein heeft veel autonomie en dient dus verantwoording af te leggen over de wijze waarop hij die aanwendt. Prestatie meting is een elegante manier van verantwoorden: het maakt glashelder welke producten voor hoeveel geld aan de samenleving worden geleverd. 'Ruimte, resultaat en rekenschap' is een bijpassende oneliner. Wie onvoldoende vonnissen wijst, onvoldoende publiceert of wiens trainen te laat vertrekken, moet de gevolgen hiervan dragen: een onvoldoende resultaat is voor iedereen zichtbaar en kan financiële consequenties hebben. Prestatie meting is zo ook bevorderlijk voor de ontvankelijkheid van professionals voor politieke of maatschappelijke opvattingen, of voor wensen van cliënten.

Een tweede standpunt – dat van de *non-believers* – is dat cijfers onvoldoende zeggen over de aard, de kwaliteit en het innovatieve karakter van de prestatie. Wat zegt het wanneer een maatschap van chirurgen ieder jaar een hoog sterftecijfer moet melden? De maatschap kan een verzameling incompetent chirurgen zijn, maar het kan hier ook gaan om topchirurgen met patiënten die elders zijn uitbehandeld en bij hen hun laatste toevlucht zoeken. Wat zegt het wanneer de rechtbank Amsterdam langere doorlooptijden bij strafzaken kent dan de rechtbank Zwolle? Wellicht zijn Amsterdamse advocaten assertiever dan hun collega's elders, iets waar een rechter niet veel aan kan doen. Prestaties van professionals in de publieke sector zijn nu eenmaal vaak het resultaat van interactie: de chirurg is voor zijn prestatie mede afhankelijk van het type patiënt dat zich bij hem meldt; de rechter is mede afhankelijk van het gedrag van de advocaat en de reclassering; de school is mede afhankelijk van de houding van ouders tegenover hun kinderen. Bovendien: publieke professionals leveren producten die aan verschillende, soms conflicterende eisen moeten voldoen. Een rechter moet snel een vonnis wijzen (kwantificeerbaar), maar ook kwaliteit leveren (niet kwantificeerbaar). Evenzo moet een school leerlingen laten presteren, maar ook een goed sociaal-pedagogisch klimaat creëren. Dat laatste is niet kwantificeerbaar. Wie de chirurg, rechter of leraar slechts afrekent op gekwantificeerde prestaties, doet aan al deze complexiteit geen recht en trekt de professie uit het lood.

'Hoe harder de bestuurder afrekent op prestaties, des te groter de prikkel voor perverserend gedrag.'

Voor beide stellingnamen valt wat te zeggen en is het nodige empirisch bewijs beschikbaar. Een systeem van prestatie meting heeft onmiskenbaar positieve effecten, die in elk van de drie hier besproken boeken aan de orde komen. Het is sterk gericht op resultaten en biedt zo een *prikkel voor presteren*. Dit voordeel spreekt vooral aan in vergelijking met de alternatieven voor prestatie meting. Organisaties kunnen worden aangestuurd – hier is enig managementlingo noodzakelijk – op output, input of throughput. Prestatie meting is een vorm van outputsturing: de professional wordt afgerekend op zijn resultaten. Bij inputsturing is de aandacht gericht op het begin van het productieproces: er wordt een plan gemaakt (bijvoorbeeld: een onderzoeksprogramma voor vijf jaar) en hieraan wordt een budget toegekend. Inputsturing levert een prikkel op om prachtige plannen te maken, veel te beloven en vervolgens de bijpassende financiële middelen te claimen. De vraag of deze plannen vervolgens ook worden waargemaakt, krijgt vaak minder aandacht. Bij throughputsturing staat niet het plan of het resultaat centraal, maar de inspanning die nodig is om van plan tot resultaat te komen: het aantal uren om een vonnis te schrijven of het aantal dagen om een vak te geven. Throughputsturing geeft een prikkel om zo veel mogelijk uren of dagen te maken. Kortom: outputsturing binnen, bijvoorbeeld, een universiteit belooft de wetenschapper die publiceert en goed onderwijs geeft. Inputsturing belooft de wetenschapper die plannen maakt. Throughputsturing belooft de wetenschapper die veel uren maakt.

Prestatie meting levert hiermee een bijdrage aan *de-bureaucratisering van professionele organisaties*. De universiteit die op output stuurt, kan het door velen verfoeide circus van

plannenmakerij-ter-verwerving-van-onderzoeksgelden afschaffen en heeft ook geen tijdschrijfsystemen nodig om te kunnen bepalen hoeveel uren haar medewerkers maken. De prestatie is interessant, niet of ze in het zweet des aanschijns tot stand is gekomen. Prestatiemeting is daarmee een elegante vorm van sturing.

Ten slotte kan prestatiemeting een belangrijke rol spelen bij processen van *verantwoording*. In de afgelopen jaren is in veel westerse landen de politieke agenda volgelopen met vraagstukken die de uitvoering van overheidstaken betreffen. Politieke en maatschappelijke aandacht voor bijvoorbeeld openbare veiligheid, punctualiteit van dienstverlening of snelle afhandeling van strafzaken, betekent dat politie, NS en rechtbanken gedwongen worden verantwoording af te leggen over hun prestatie. Prestatiemeting biedt hier een aantrekkelijke propositie: gekwantificeerde prestaties zijn goed te overleggen en kunnen bijdragen aan een beeld dat de professional gevoelig is voor maatschappelijke opvattingen. De korpschef kan met cijfers aantonen dat zijn stad veiliger is geworden en de NS dat er meer op tijd wordt gereden.

Maar er is ook onderzoek dat tot een minder rooskleurig beeld leidt. Vooral Pollitts *The Essential Public Manager* – het meest evenwichtige boek van de drie – onderbouwt de stellingname van de sceptici. Het eerste inzicht luidt dat prestatiemeting leidt tot *strategisch gedrag*. Productiecijfers laten zien dat het aantal septs bij het OM daalt. Wat is de verklaring? Mogelijkheid 1: het OM vervolgt meer verdachten. Mogelijkheid 2: een medewerker van het OM schraapt op het politiebureau een groot aantal zaken, die dus niet worden doorgeleid naar het OM. Gevolg: minder septs, maar dit is uiteraard slechts een productie-op-papier. Pollitt geeft het voorbeeld van Engelse onderwijzers die de fouten van hun leerlingen corrigeren bij landelijke toetsen om zo de prestaties van de school te verbeteren.

Ook menen sceptici dat prestatiemeting *innovaties blokkeert*. Een indrukwekkende publicatielijst kan erop duiden dat de desbetreffende wetenschapper bestaand onderzoek voortdurend reproduceert. Dit is het type wetenschapper wiens belangrijkste prestatie bestaat uit het voortdurend recyclen van dezelfde noties maar zo toch hoog scoort in bepaalde prestatie-meetsystemen.

En: prestatiemeting *fixeert*. Ze leidt tot wat Pollitt een ‘tunnelvisie’ of *Targititis* noemt. Een organisatie moet in staat zijn kansen te benutten: onverwachte mogelijkheden die zich voordoen om iets te realiseren. Een te sterk accent op het realiseren van vooraf afgesproken prestaties kan een organisatie blind maken voor dergelijke kansen. Bovendien: prestatiemeting *blokkeert ambities*. Ze is namelijk een prikkel voor het optimaliseren van de input. Het selectiecriteria voor de input is dat deze zo min mogelijk throughput vereist om de gewenste output te verkrijgen. In *Evaluating Public Management Reforms* spreken Boyne en anderen in dit verband van het risico van *cream skinning*. Onderzoek laat zien dat scholen die worden beloond voor prestaties, selecteren op input. Zij kunnen zwakke leerlingen weigeren toe te laten. Een organisatie die haar input optimaliseert, levert zo in op haar ambities.

De – empirisch gestutte – litanie gaat verder. Prestatiemeting *verdrijft de professionele habitus*. Prestatie-indicatoren meten kwantiteiten en zullen dus vooral aangrijpen bij meetbare en goed te definiëren aspecten van de prestatie. Onderzoek laat zien dat organisaties die zich jarenlang hebben ingespannen om aan de eisen van prestatiemeting te voldoen, uiteindelijk slechter presteren, omdat ze de niet-definieerbare aspecten hebben verwaarloosd. Prestatiemeting *verdrijft ook de bestelverantwoordelijkheid*. Bestelverantwoordelijkheid houdt in dat de professionele inzichten die organisaties ontwikkelen, beschikbaar worden gesteld aan branchegenoten. Amerikaanse scholen die op prestatie met elkaar concurreren, blijken bijvoorbeeld nauwelijks bereid om hun *best practices* met elkaar te delen. Ten slotte: prestatiemeting *bureaucratiseert*. Organisaties die het goed doen in een outputsysteem, blijken een bureaucratie in het leven te roepen om de productiecijfers op te poetsen.

Dit worden de perverse effecten van prestatiemeting genoemd. Deze ontstaan doordat de professional het systeem van prestatiemeting ontduikt (bijvoorbeeld: strategisch gedrag) of juist doordat hij zich geheel daaraan conformeert en nog slechts in de gekwantificeerde prestatie is geïnteresseerd (bijvoorbeeld: blokkeren van ambitie of innovatie).

Beide standpunten hebben hun eigen institutionele inbedding. Het eerste vinden we – enigszins gechargeerd – in de wereld van de politici, bestuurders, stafafdelingen planning en control, rijksambtenaren, toezichhouders (raden van toezicht, Rekenkamer) en adviesbureaus. Een van mijn studenten aan de Nederlandse School voor Openbaar Bestuur, die van dit verschijnsel studie heeft gemaakt in de Verenigde Staten, sprak van een *performance management industry*: een moeilijk te bestrijden coalitie van voorstanders van prestatiemeting, die soms uit overtuiging handelt en soms uit eigenbelang. De industrie introduceert voortdurend nieuwe instrumenten voor prestatiemeting en belooft de bestuurder dat het gebruik hiervan een belangrijke bijdrage zal leveren aan zijn sturend vermogen.

Het tweede standpunt wordt gedragen door professionals en bestuurders die relatief dicht bij de professe staan (*knowledgeable managers*). De taal in deze wereld is geheel anders. Prestatiemeting kent de taal van de zelfstandige naamwoorden: artikelen, vonnissen, medische behandelingen. De professionele taal is de taal van de werkwoorden, en hierdoor veel rijker: onderzoeken, oordelen, opereren. In de professionele taal is er ruimte voor de ambiguïteit van de professe, die zich in cijfers uiteraard moeilijk laat uitdrukken. Deze professionals en *knowledgeable managers* zijn veel minder goed

in coalitievorming dan de *performance management industry*, maar beschikken wel over de mogelijkheid om prestatie-meetsystemen te perverteren. Overigens wordt boven in de organisatie altijd onderschat hoe intelligent de professional is in het perverteren van prestatie-meetsystemen.

‘Voortdurend wordt de mythe van stuurbaarheid bevestigd en de complexiteit van de professionele of publieke organisatie gereduceerd tot al te eenvoudige denkschema’s.’

Dat deze twee werelden hemelsbreed verschillen, komt ook tot uitdrukking in bovenstaande boeken. Zo worden in *Public Management and Governance* de mogelijke nadelen van prestatie-meting in één enkele alinea afgedaan. Het boek bevat een technisch-instrumentele beschouwing over prestatie-meting en hieraan gerelateerde ontwikkelingen als contractmanagement en de opkomst van allerlei vormen van inspectie en toezicht. Het bevat inleidende overzichten van een groot aantal gerelateerde managementmodes; de voordelen van prestatie-meting krijgen veel aandacht en er is dus in beperkte mate sprake van kritische reflectie. De studie is daarmee een pars pro toto voor veel managementboeken die voortdurend de mythe van stuurbaarheid bevestigen en de complexiteit van de professionele of publieke organisatie reduceren tot al te eenvoudige denkschema’s.

In beschouwingen die een professionele sector meer diepgaand behandelen, krijgt de kritiek op prestatie-meting meer ruimte. Deze literatuur is vaak sectorspecifiek en dus verspreid over verschillende disciplines. Boyne en coauteurs, en – vooral – Pollitt, brengen veel van deze literatuur bijeen. Zij constateren naast positieve ook negatieve effecten. Een studie naar het gebruik van prestatie-indicatoren in het onderwijs, opgenomen in *Evaluating Public Management Reforms*, leidt tot de conclusie dat er geen bewijs is dat deze tot efficiencyverhoging leiden. Wel neemt de responsiviteit (de bovengenoemde ontvankelijkheid voor maatschappelijke voorkeuren) enigszins toe. Ook voor de gezondheidszorg komt in dit boek een kritisch beeld naar voren: er is wel degelijk sprake van een geringe verbetering van de efficiency en van de responsiviteit, maar die gaat waarschijnlijk ten koste van de toegankelijkheid.

Voor de stellingnamen van zowel de *believers* als van de *non-believers* is dus het nodige empirische bewijs voorhanden. De vraag welke de juiste is, is daarmee weinig vruchtbaar. Interessanter is de vraag naar de dynamiek van prestatie-meetsystemen: onder welke condities doen zich processen van pervertering voor en onder welke behouden zij hun elegantie?

Hierbij zijn drie observaties van belang. Ten eerste zal de professional een beoordeling op grond van uitsluitend gekwantificeerde prestaties altijd behoorden als unfair. Immers: prestatie-meetsystemen geven een zeer beperkt zicht op de professionele prestatie aangezien deze in interactie totstandkomt en aan conflicterende criteria moet voldoen. Bijgevolg heeft een gerealiseerd productiecijfer niet alleen weinig betekenis (zie de chirurg hierboven). Het kan zelfs de werkelijkheid in ernstige mate verhullen, en dus bij de sturing slechts een bescheiden rol spelen. Wanneer een gevangenisdirecteur meldt dat het aantal geweldplegingen tegen het personeel in zijn instelling sterk is gedaald, laat dit geheel verschillende conclusies toe. Wellicht heeft hij veel geïnvesteerd in een betere omgang met gedetineerden. Conclusie: goede prestatie, goede cijfers. Maar misschien is hij iemand die problemen afwentelt op anderen en is zijn belangrijkste vaardigheid dat hij lastige gedetineerden weet over te plaatsen. Conclusie: slechte prestatie, goede cijfers. Zijn collega, die het als een professionele waarde ziet om agressieve gedetineerden binnen de eigen instelling te houden, kan met de tegenovergestelde situatie worden geconfronteerd: goede prestatie, slechte cijfers. Cijfers zonder contextuele informatie zijn betekenisloos.

Wanneer gaan nu de negatieve effecten domineren? Dit is – tweede observatie – vooral afhankelijk van de sturingsdrift van de bestuurder. Hoe meer deze op prestatiecijfers stuurt en hoe ingrijpender de sancties die hij aan de gekwantificeerde prestatie verbindt, des te groter de kans dat er negatieve effecten zullen optreden. De professional wordt dan immers geconfronteerd met een zeer frustrerende situatie: hoewel het cijfer een onvolledig of zelfs onjuist beeld geeft, wordt hij er toch op afgerekend. Het behoeft dan ook niet te verbazen dat er voor hem sterke prikkels ontstaan voor het perverteren van de prestatie-meting.

Stel: een bestuurder kondigt aan dat hij de gekwantificeerde prestatie van de professional zal gebruiken voor de allocatie van grote delen van het budget of voor de vaststelling van diens persoonlijke beloning. Het gevolg laat zich raden: de professional zal – zeker na verloop van tijd – strategisch gedrag gaan vertonen, inleveren op ambities en innovatie, een bureaucratisch systeem optuigen om de productie op te poetsen et cetera. Hij is hiertoe gelegitimeerd omdat het cijfer zijn werkelijke prestatie slechts gedeeltelijk weergeeft of zelfs verhuult; een beoordeling op uitsluitend cijfers is dus unfair. Je kunt het wel de Wet van de afnemende effectiviteit noemen: hoe harder de bestuurder afrekenet op prestaties, des te groter de prikkel voor perverterend gedrag. Gevolg: de effectiviteit van prestatie-meting neemt sterk af.

‘De bestuurder-met-sturingsdrift reageert veelal op dit perverterende gedrag door verdere verfijning van het prestatie-meetsysteem.’

In de derde plaats: de bestuurder-met-sturingsdrift reageert veelal op dit perverterende gedrag door verdere verfijning van het prestatie-meetsysteem. Definities van producten worden verfijnd om strategisch gedrag te mitigeren; er worden aanvullende definities geïntroduceerd en dus ook extra cijfers opgevraagd; procedures om prestaties te tellen worden verscherpt, de controle op het cijfermateriaal

wordt geïntensiveerd et cetera. Dit wordt mushrooming genoemd: de informatiehonger neemt alsmaar toe, het prestatie-meetsysteem wordt zwaarder en zwaarder, en verliest hierdoor zijn elegantie (de term is van Geert Bouckaert, een van de auteurs in Bovaird en Löffler). De vraag is of het helpt: nieuwe definities roepen veelal weer nieuw strategisch gedrag op. Bovendien: wanneer aan prestatie-meting zware consequenties worden verbonden, dan moet het systeem op orde zijn. Pollitt geeft een overzicht van de eisen waaraan een goede evaluatie van de prestatie moet voldoen (kwaliteit en tijdige beschikbaarheid van de data, methode van verwerking et cetera). Verfijning van het systeem stelt hogere eisen aan de kwaliteit daarvan. Maar deze eisen zijn niet altijd eenvoudig te realiseren en zeker niet altijd kosteneffectief. Wanneer de bestuurder te veel stuurt met prestatie-meting, is verdere versterking van deze sturing zelden de oplossing. Ze leidt slechts verder af van prestatie-meting als elegant systeem.

Het bovenstaande impliceert dat de remedie juist ligt in een gematigd gebruik van prestatie-meting. Dit gematigd gebruik kan allerlei vormen aannemen: zorg dat niet het gehele budget van een professional afhankelijk is van zijn gekwantificeerde prestaties, investeer in interactie met de professional over zijn prestaties, maak oordeelsvorming over de professional mede afhankelijk van deze interactie, tolereer het idee dat niet over alles verantwoording moet worden afgelegd, span je als bestuurder in om het uitdijen van het systeem te voorkomen, breng niet alle cijfers naar buiten, doe niet moeilijk wanneer organisaties het prestatie-meetsysteem enigszins afdempen. Ook: reken nooit uitsluitend af op gekwantificeerde prestaties, maar besteed eveneens aandacht aan het professionele verhaal achter de prestatie. Gebruik cijfers om professionals kritisch te bevragen en verbeteringen af te spreken; en slechts bij voortdurende wanprestatie om ze erop af te rekenen.' Een gematigd gebruik betekent dat er minder prikkels voor pervertering bestaan en dat de positieve effecten van prestatie-meting dus ook meer kans krijgen.

Prestatie-meting is voor menig bestuurder verleidelijk: ze lijkt veel mogelijkheden te bieden om de professional aan te sturen. Wie echter te veel met dit instrumentarium stuurt, roept de negatieve effecten over zich af. Zo biedt prestatie-meting een lastige propositie aan de bestuurder: ze belooft een versterking van zijn sturend vermogen, maar alleen degene die terughoudend hiervan gebruikmaakt, zal iets van die belofte realiseren. Valt prestatie-meting in handen van de moderne machomanager, met zijn hedendaagse jargon van 'daadkracht', 'afrekenen op prestatie' en dergelijke flinkheden, dan is het systeem van prestatie-meting binnen de kortste keren geperverteerd en verworden tot bureaucratie en zinloos ritueel.

Hans de Bruijn is als hoogleraar Organisatie en Management verbonden aan de Faculteit Techniek, Bestuur en Management van de Technische Universiteit Delft.

Besproken boeken:

Evaluating Public Management Reforms. Principles and Practice
door **door George A. Boyne e.a.**
McGraw-Hill, imprint Open University Press. Buckingham 2003
192 pag., € 37,67

Public Management and Governance
door **door Tony Bovaird en Elke Löffler (red.)**
Taylor & Francis Group. Londen 2003
256 pag., € 40,60

The Essential Public Manager
door **door Christopher Pollitt**
McGraw-Hill. Maidenhead 2003
224 pag., € 34,39