

In de ban van beoordelingen

**Kwaliteit van visitaties
in
het wetenschappelijk onderwijs**

Arno F.A. Korsten

en

Jean M.L.R. Schutgens

150812

Voorwoord

Nederland kent een traditie van enkele decennia in het beoordelen van opleidingen in het wetenschappelijk onderwijs (w.o). Deze beoordelingen geschieden sinds ongeveer 1985 voor opleidingen op eenzelfde vakgebied. Steeds licht een visitatiecommissie van onafhankelijke vakgenoten een opleiding uit een reeks op hetzelfde vakgebied door: alle scheikundeopleidingen, alle bestuurskundeopleidingen, enz.. De beoordeling geschiedt onder de vlag van 'visitatie'. Opleidingen worden dus gevisiteerd, zoals het in de wandeling heet. Visiteren komt ook op andere terreinen voor, zoals van hbo- opleidingen, van politiekorpsen, van gemeenten. Visitaties zijn bedoeld om zich te verantwoorden en om te verbeteren. Meestal zijn beide doelen aan de orde. Waar het accent op ligt, moet steeds opnieuw goed bekeken worden. In het w.o. ligt het accent op openbare oordelen en 'verbeteren waar nodig'. Waar openbaar geoordeeld wordt, is sprake van verantwoording, in dit geval naar de inspectie, de minister en de samenleving.

Kernelementen bij visitatie zijn doorgaans: een startmoment waarop alle opleidingen er klaar voor zijn, een beoordelingskader en protocol (volgorde, fasering, werkwijzen), een zelfstudie op te stellen door de gevisiteerde, een visitatiecommissie die de regie voert en ter plaatse komt 'inspecteren' en de zelfstudie tegen het licht houdt en een blik van buiten binnen haalt, een conceptrapport (visitatierapport genoemd) en de mogelijkheid tot het geven van commentaar, een openbaar eindrapport met bevindingen en doorgaans ook aanduidingen van de verbeterrichting.

De beoordeling is steeds integraal. Opleidingen worden integraal bekeken. Dat betekent dat visitaties van alleen het studierendement of een ander onderdeel van het onderwijssysteem uitgesloten zijn.

In het w.o. geschiedt visiteren niet in het luchtledige maar onder toezicht van de onderwijsinspectie die beoordeelt of een visitatie aan de maat is en natrekt of verbeteringen worden doorgevoerd. Visitaties zijn in het w.o. sinds 1985, als gevolg van besluiten van alle colleges van bestuur gezamenlijk, verplicht.

Visiteren valt te omschrijven als het organiseren van een integrale inspectie en beoordeling van prestaties door onafhankelijke vak- of branchegenoten die zijn opgenomen in een visitatiecommissie. Deze beoordeling resulteert in een schriftelijke openbare rapportage van de visitatiecommissie. Visitaties zijn dus op te vatten als een vorm van resultaatmeting en -beoordeling van de kern van het bestuur. Daarmee zitten we in het studiegebied van 'performance measurement'.

Door te rapporteren organiseren visitatiecommissies een brede terugkoppeling op het handelen van een organisatie of onderdeel daarvan, waaruit te leren valt hoe de organisatie zich kan verbeteren. Visitaties zijn daarmee een onderdeel van het kwaliteitsbeleid van een organisatie en soms van een hele sector of branche. Daarmee leveren visitaties in principe een bijdrage aan het leervermogen van de publieke sector.

Dit visiteren komt sinds een aantal jaren voor in de sector hoger onderwijs (w.o. en hbo) maar ook in het grotestedenbeleid, bij provincies en gemeenten, zelfstandige bestuursorganen, ziekenhuizen, de sector geestelijke gezondheidszorg en gehandicaptenzorg, omroepmedia, politiekorpsen, woningcorporaties.

Visiteren kan verschillende vormen aannemen. We kunnen onderscheiden: verplichte of vrijwillige visitaties; visitaties die gericht zijn op verantwoorden, op leren of op beide; visitaties die retrospectief gericht zijn of prospectief, of beide.

Hoe visitaties op zich weer te beoordelen? Deze studie behandelt het onderwerp kwaliteit van visiteren in de sector waar volgens ons het langste ervaring met visiteren is opgedaan en het meest bezinning daarop heeft plaatsgevonden: het wetenschappelijk onderwijs (w.o.). In die sector is een heel bestel ontstaan, waarin visiteren een gekend en erkend onderdeel is. Vragen die aan de orde zijn, luiden:

- wat is visiteren eigenlijk?;
- wat is kwaliteit van visiteren?;
- hoe is die kwaliteit te meten?;
- wat blijkt uit metingen van die kwaliteit?; treden neveneffecten op?;
- kan het resultaat van metingen vertekend zijn of niet?;
- beantwoord het visiteren aan doelen die actoren in het betreffende veld ermee hebben, zoals de onderwijsinstellingen, het ministerie van OC&W, de onderwijsinspectie,?

Daarbij ging het uiteindelijk om de vraag of sprake is geweest van succes of falen van visitaties in het w.o. en wat bijdroeg aan dat succes of falen.

Als sprake is geweest van succes is daaruit dan af te leiden hoe visiteren er idealiter uit zou moeten zien in een andere sector? Deze (congruentie-)vraag mag gesteld worden omdat in de literatuur gesteld wordt dat er essentiële *'beginselen van behoorlijk visiteren'* bestaan (Korsten, 2004) respectievelijk een vaker voorkomend *'stramien'* in het visiteren bestaat (Schillemans, 2008). Maar er zijn ook pleidooien om juist oog te hebben voor verschil in branche of context: elke sector kan eigen visitaties hebben die ook succesvol zijn. Is het gemeenschappelijk stramien klein of juist groot?

Het verder weg liggende doel van deze studie was om te komen tot 'verwachte uitkomsten' (in andere arena's wel hypothesen genoemd) voor verder onderzoek naar visiteren in andere sectoren. Daarom komen ook (aanzetten tot) verwachtingen aan bod. Hoewel de verwachtingen gebaseerd zijn op ervaringen in het w.o. zijn ze mede bezien in het licht van de internationale wetenschappelijke literatuur over *'performance measurement'* en de rol van *'performance information in the public sector'*. Deze literatuur wordt gebruikt om de verwachtingen verder te funderen en te ondergraven.

In de literatuurlijst vindt de lezer dan ook vooral meerdere typen literatuur: over visiteren in het algemeen, over visiteren in het w.o., over performance measurement & management; over het gebruik van onderzoek, benchmarking en visitaties in het openbaar bestuur en in het wetenschappelijk onderwijs; over 'de relatie tussen wetenschap en beleid'.

Voor het uitvoeren van dit uitgebreide onderzoek is gebruik gemaakt van een aantal bronnen. Ze worden elders vermeld. Speciaal moet gewezen worden op de (indirecte) betekenis van eigen directe ervaringen met visiteren van een van ons: een onderzoeker (Schutgens) kreeg als gemeentesecretaris zelf te maken met de voorbereiding op twee visitaties van 'zijn' gemeente. De ander (Korsten) heeft zelf ervaringen als iemand die onderwijs- en onderzoeksvisitaties op het gebied van bestuurskunde en van recht onderging, die visitaties leidde (visitatie van de hbo-opleidingen in de bestuurskunde en overheidsmanagement; midterm review na een visitatie van wo-opleidingen in de bestuurskunde aan de EUR), visitaties uitvoerde van twee provinciebesturen (Limburg, Zeeland) en van negen Vlaamse grootstedelijke gemeentebesturen (Antwerpen, Gent, Leuven, enz.).

Deze uitgebreide studie, waaraan maanden gewerkt is, is weliswaar verricht door twee auteurs, maar deze zal uiteindelijk een onderdeel zijn van een onderzoeksrapportage van J.Schutgens naar visiteren in het algemeen en in de gemeentelijke sfeer in de vorm van 'meting' van bestuurskracht in het bijzonder. De studie zal eerder worden gepubliceerd op de site: www.arnokorsten.nl. Deze werkwijze vertoont gelijkenis met de publicatie van artikelen van

meerdere auteurs, die ook wel verschijnen in tijdschriften en naderhand een rol vervullen als hoofdstuk in een dissertatie.

Geen kwaliteit zonder 'review'. Voor instemmend commentaar op een eerdere versie van dit essay en enkele adviezen gaat onze dank uit naar prof.dr. F. (Frans) L. Leeuw, voormalig Hoofdinspecteur Hoger Onderwijs. Hij is sinds mei 2003 directeur van het Wetenschappelijk Onderzoeks- en Documentatiecentrum (WODC) van het Ministerie van Veiligheid en Justitie en deeltijd-hoogleraar Recht, Openbaar bestuur en Sociaal-wetenschappelijk onderzoek aan de Universiteit Maastricht (FdR).

Arno Korsten/Jean Schutgens
Heerlen/Maastricht, 15 augustus

Inhoud

1 Inleiding

2 Argumenten pro en contra visiteren

3 Visiteren in het wetenschappelijk onderwijs: eerste verkenning

4 Breukpunt in hoger onderwijsbeleid: streven naar universitaire autonomie

5 Contouren van het visitatiestelsel in het wetenschappelijk onderwijs

6 De kwaliteit van visitaties in het wetenschappelijk onderwijs

6.1 Beoordelingscriteria

6.2 Hoofdbeeld van de kwaliteit van visiteren in drie rondes

6.3 Nadere analyse van visitatiekwaliteit in de eerste ronde: 1988-1993

6.4 Nadere analyse van visitatiekwaliteit in de tweede ronde: 1994-1999

6.5 Nadere analyse van visitatiekwaliteit in de derde ronde: 2000-2003

7 Terug naar de positieve en kritische beelden van visiteren

8 Enkele algemene conclusies

9 Enkele verwachtingen voor verder onderzoek naar visitatiekwaliteit

10 Benutting van onderzoeksbevindingen nader bezien

11 Vier kwaliteiten van visitaties

Literatuur

Over de auteurs

1 Inleiding

Visitaties van wetenschappelijke opleidingen en van wetenschappelijk onderzoek worden wel gezien als de paradedepaardjes van de universitaire externe kwaliteitszorg. Maar ze passen niet bij elk interbestuurlijk systeem van sturing, dus van bevoegdheden, verhoudingen en betrekkingen tussen de centrale overheid en de onderwijsinstellingen. Ze passen nauwelijks bij hiërarchische sturing en juist wel in een systeem waarin zelfregulering, autonomie en kwaliteit betekenis hebben. Dat leggen we uit en daarna volgt de probleemstelling van dit stuk. We zullen laten zien dat visiteren inderdaad in het hoger onderwijs van Nederland van de grond kwam toen een omslag naar meer zelfregulering ('sturen op afstand') werd gemaakt.

Visiteren past bij 'sturing' maar welke?

Het overheidsbestuur kan ten behoeve van sturing van de samenleving gebruik maken van de creatie van voorzieningen en van inzet van verschillende *sturingsinstrumenten*, zoals juridische, economische en communicatieve sturing en combinaties daarvan. Bij juridische instrumenten is over het algemeen sprake van geboden en verboden, bij economische sturing van financiële prikkels (subsidies) en sancties (belastingen) en bij communicatieve sturing van propaganda, op dialoog gerichte interactie of voorlichting. Deze instrumenten kunnen verruimend of beperkend worden ingezet. Wie via beleidsinstrumenten beperkend optreedt, voorkomt, remt, snoert in of kanaliseert maatschappelijk initiatief. Wie initiatiefname in de samenleving verruimt, doet het tegendeel (Van der Doelen, 1991). In de werkelijkheid van alle dag komt vaak op een bepaald terrein een *'policy mix'* voor, een combinatie van genoemde sturingsinstrumenten. Hier en daar kan een beleid zowel bepaalde verruimende als beperkende kenmerken in zich dragen (Bressers, De Jong, Klok en Korsten, 1993).

Achter een concreet beleidsprogramma, waarin een beleidsinstrumentenmix is opgenomen, gaat doorgaans een *sturingsconcept* (of filosofie) schuil. Dat is een abstract idee over hoe overheid, organisaties en samenleving met elkaar omgaan op een bepaald gebied (Bressers e.a., 1993; Bekkers, 2007: 92). Denk daarbij bijvoorbeeld aan concepten als centrale regulerende sturing, marktwerking, sturen op afstand, netwerksturing of prikkelsturing. In dit verband wordt ook wel gesproken van een *beleidsparadigma*.

Dat mag zo zijn, het is niet zo dat een heel kabinet een en hetzelfde sturingsconcept hanteert. Het is onmogelijk Nederland uitsluitend te besturen vanuit bij voorbeeld netwerksturing of marktwerking. In de praktijk komt dat de afgelopen decennia ook niet voor, hoewel het zetten van conceptuele accenten als nastrevenswaardig kan gelden. Dat wil zeggen: een kabinet kan wel bepaalde (eenheid van) sturing nastreven, maar de eigenheid van problemen, oplossingen en bestaand beleid en de aanwezigheid van koepels en organisaties kaderen een vrije keuze in en beperken die doorgaans. Regeerakkoorden, regeringsverklaringen en tal van beleidsanalyses getuigen daarvan (Hill, 2009). Niet overal is bijvoorbeeld marktwerking mogelijk; waar geen markt is, geen marktwerking. Een markt waar lucht - een collectief goed - verhandeld wordt, is bijvoorbeeld uitgesloten. Het is ook gebleken dat de sturingsfilosofie achter gezondheidszorg verschilt van die in het hoger onderwijs en dat ook *binnen* zowel de gezondheidszorg als binnen het onderwijsbeleid verschillende sturingsarrangementen gegroeid zijn (In 't Veld, 2001a; Huisman & Theisens, 2001; Maarse, 2011). Maar er zijn wel omslagen in beleidsprogramma's te maken (De Vries, 2002).

In het midden van de jaren tachtig is op het terrein van hoger onderwijs een omslag in gang gezet van centralistische regelgerichte sturing naar een ander sturingsconcept, namelijk naar *sturing op afstand* (Noordegraaf en Kickert, 1993; Knippenberg en Van der Ham, 1994; Algemene Rekenkamer, 2008a). Onderwijsinstellingen als universiteiten en hogescholen kregen door die beleidswijziging meer autonomie op het gebied van onderwijs. Die omslag -

een vernieuwing genoemd - vinden we terug in de HOAK-nota uit 1985, die ten tijde van het kabinet-Lubbers I onder verantwoordelijkheid van minister Deetman verscheen. Hoe tracht het kabinet-Lubbers I zicht te houden op de kwaliteit van hoger onderwijs en verantwoording over onderwijsprestaties en bezinning bij betrokkenen op kwaliteit te bevorderen? Door in het kielzog van de HOAK-nota een *visitatiestelsel* (of *visitatieregime*) in te richten waarin verantwoord, leren (verbeteren) en communicatie een belangrijke rol vervullen en *visitatie* ook verbonden is met andere beleidsinstrumenten zoals toezicht (op die *visitatie*) en bekostiging van de gevisiteerde opleiding. Belangrijke elementen uit dat stelsel zijn: de zelfevaluatie van een opleiding; *visitatie* van de instelling door een bezoek van een *visitatiecommissie* ter plaatse en gesprekken met betrokkenen bij en studenten van een opleiding; een *conceptvisitatie*rapport dat wordt voorgelegd aan de instelling, en een (overall-) *visitatie*rapport over de kwaliteit van alle gevisiteerde opleidingen op eenzelfde terrein.

Vraagstelling

Bij 'sturen op afstand' in het hoger onderwijs is voorzien in beoordeling van reguliere opleidingen uit *eenzelfde* vakgebied door *visitatiecommissies*. Die *visitaties* staan hier centraal, in het bijzonder *visitaties* in het universitair onderwijs, dat wil zeggen: het wetenschappelijk of academisch onderwijs (verderop ook wel: 'het w.o.' genoemd).

De *vraagstelling* hier luidt:

Hoe is de *visitatie* van w.o.-opleidingen ingericht en wat is achteraf de kwaliteit van de *visitaties* geweest? Is *visiteren* een positief proces geweest waardoor het *visiteren* jarenlang is voortgezet of is het *visiteren* geëindigd in een echec?

Op deze vraag en subvraag zijn twee voorlopige antwoorden te geven: een positief en een negatief antwoord. In het positief beeld is *visiteren* een paradepaardje, in het kritisch beeld ligt dit volkomen anders. In deze analyse wordt nagegaan of er empirische steun is te krijgen voor het positieve of kritische antwoord of een combinatie van beide.

Het verder weg liggende doel is om hypothesen op te stellen voor onderzoek naar de kwaliteit van *visitaties* in andere sectoren, zoals het grotestedenbeleid.

Definitie

We hanteren hier de volgende *werkdefinitie* van 'visiteren'. *Visiteren* valt te omschrijven als het organiseren van een (verplichte of vrijwillige) integrale inspectie en beoordeling van prestaties door vakgenoten, opgenomen in een schriftelijke rapportage van een onafhankelijke *visitatiecommissie*, van een organisatie of organisatiedeel of van een aantal organisaties of onderdelen op eenzelfde gebied of in eenzelfde sector of branche. In deze definitie is 'onderwijs' niet opgenomen omdat *visiteren* in tal van sectoren voorkomt. Kernbegrippen zijn: structureel, prestaties, beoordeling, integraliteit, commissie van vakgenoten en rapportage.

'Structureel' wil zeggen dat er sprake is van afspraken over vorm en frequentie van *visitatie* in een bestel waarin meerdere organisaties met onderscheiden verantwoordelijkheden betrokken zijn (bijvoorbeeld: een ministerie, onderwijsinspectie, VSNU, instelling, enz.).

'Integrale beoordeling en inspectie' verwijst naar een beeld van de resultaten van de prestatieve kern van de organisatie. Het gaat altijd om de primaire processen en dus niet slechts om de ondersteunende functies of bedrijfsvoering of alleen het strategisch beleid.

Een *visitatie* zoals we die hier opvatten is probleemgericht en multi-aspectueel en volstaat nooit met de analyse van een of enkele elementen of aspecten van een organisatietaak.

Wie tracht het te visiteren object in beeld te krijgen poogt een beeld van 'goed bestuur', dus van 'deugdelijk bestuur' van een gevisiteerde organisatie of organisatiedeel (instelling, opleiding, etc.) in zijn context (van burgers, klanten of studenten, enz.) te krijgen. Het gaat om een basaal beeld, niet om een specialistische analyse van een deel, zoals bij onderwijs (alleen) de examenkwaliteit of het opleidingsrendement.

De te visiteren organisaties is altijd behulpzaam bij het aanleveren van data die passen in en relevant zijn voor het beoordelingskader dat de visitatiecommissie hanteert. Vaak zal sprake zijn van een zelfstudie of zelfevaluatie. Is dat het geval dat bespaart dat de visitatiecommissie veel werk. Een commissie zal overigens de kwaliteit van die zelfstudie beoordelen: op zich, door bezoek ter plaatse en daar dan documenten te raadplegen, observaties te verrichten en gesprekken te voeren en door te spreken met externen.

Een visitatie mondt altijd uit in een rapportage. Er wordt dus niet volstaan met bij voorbeeld een film, een mondelinge overdracht van de beoordeling, een power point presentatie of een ingevulde vragenlijst. Doorgaans zal die rapportage openbaar zijn. Bij visitatie van ziekenhuizen is dat overigens niet het geval.

Als een visitatie plaatsvindt, kunnen nog uitgangspunten worden toegevoegd en moeten nog meerdere inrichtingsbeslissingen worden genomen, zoals over de samenstelling en werkwijze van een visitatiecommissie (zie ook Schillemans, 2009: 166-167). Die komen ook verderop aan bod.

Doel

Het doel van deze beschouwing is tweeledig:

- 1) om zowel te komen tot bevindingen in de vorm van conclusies over de kwaliteit van visitaties in het wetenschappelijk onderwijs in de onderzoeksperiode,
- 2) als om te komen tot hypothesen voor vervolgonderzoek. Die hypothesen noemen we hier in de lijn van Haverland en Yanow (2010) 'verwachtingen', dat wil zeggen: verwachte uitkomsten van vervolgonderzoek (over onderzoek verder: Van Thiel, 2007).

De indeling van deze tekst is daaraan aangepast. Eerst volgen enkele positieve en negatieve veronderstellingen over visitaties in het wetenschappelijk onderwijs, waarna de praktijk van visitaties in drie perioden belicht wordt. Daarna volgen conclusies en wordt terug gekoppeld naar de positieve en negatieve oordelen. Is visiteren in het wetenschappelijk onderwijs een succes of geen succes geweest?

Er is vanaf gezien om een samenvatting op te nemen.

2 Argumenten pro en contra visiteren

Wat houdt nu een positief en kritisch antwoord op de kwaliteit van visiteren in het wetenschappelijk onderwijs in? We geven veronderstellingen op basis van literatuurstudie en eigen ervaringen in de academische wereld.

Mogelijk positief antwoord op de kwaliteit van visiteren: argumenten

Eerst de argumenten achter het positieve beeld over visiteren in het wetenschappelijk onderwijs.

a. *Het argument van de brede steun voor opzet en werkwijze.* De visitatieprocedure in het w.o. werd door ministerie, besturen van universiteiten en onderwijsgevendens rond 1985 ontworpen en bediscussieerd en daarom als verantwoord en legitiem ervaren. Dat is een eerste

veronderstelling gebaseerd op de aanname: 'wat je zelf met anderen bedenkt, deugt, zeker als anderen instemmen' (ook het 'invented here'-argument te noemen).

b. *Het argument van de kwaliteit en betrouwbaarheid van de (moment)opname.* De instellingen werkten volledig mee aan de visitaties en de uitkomsten werden over het algemeen door de betrokkenen - zoals instellingsbesturen, decanen en onderwijscoördinatoren - herkend ('dit is een goede foto van ons'; geldigheid van de meting) en erkend ('deze foto accepteren we als waarheidsgetrouw'; betrouwbaarheid van de visitatiemeting).

c. *Benuttingspotentie.* De uitkomsten kunnen als opstap voor verbetering van opleidingen dienen.

d. *De behoorlijkheid van de procedure.* Van Vught en Westerheijden (1994) stellen dat nationale kwaliteitszorgstelsels, zoals structurele visitatie in het wetenschappelijk onderwijs en hoger beroepsonderwijs, vaak gebaseerd zijn op zelfde beginselen, zoals een coördinerend agentschap, de indiening van een zelfevaluatierapport door de onderwijsinstelling, de visitatie door vakgenoten en een openbaar rapport met bevindingen, alsmede betrokkenheid van velen, het lering trekken uit de bevindingen waardoor de leereffecten tot een interne dynamiek leiden van kwaliteitszorgsystemen. Korsten (2004a) ziet deze en enkele andere beginselen - zoals het iteratieve karakter dat zich uit in de mogelijkheid voor een instelling om commentaar te leveren op een conceptrapport van de visitatiecommissie - als positieve bijdrage aan de erkenning en herkenning van bevindingen uit het visitatieproces en spreekt daarom over 'beginselen van behoorlijk visiteren'. Volgens hem is visiteren van opleidingen een 'kwaliteitsrijk' proces als aan alle beginselen is voldaan. De volgende beginselen zijn van toepassing op het 'succes' van visiteren van opleidingen in het wetenschappelijk onderwijs.

1. Er is sprake van een initiatiefnemer tot visiteren. In geval van een dekkend visitatiestelsel kunnen immers niet alle visitaties gelijktijdig starten, maar is er een clustering en volgorde.
2. Het coördinerend agentschap (VSNU) voert regie en treedt dus op als regisseur.
3. Het visiteren geschiedt op basis van een bepaalde opzet: het protocol. Daarin is voorzien in een beoordelingskader.
4. De procedure voorziet in de aanwezigheid van een visitatiecommissie met een opdracht (gegeven in het initiatief en het protocol).
5. De visitatiecommissie bestaat uit onder meer uit vakgenoten en is onafhankelijk.
6. De dataverzameling (in het kader van een zelfstudie) vindt plaats op basis van het protocol (in casu het beoordelingskader).
7. Er is sprake van betrokkenheid van de gevisiteerde bij de dataverzameling.
8. De procedure voorziet in interactie met betrokkenen (docenten, studenten, etc.) door bezoek ter plaatse.
9. Er is sprake van contact met externen (zoals bij voorbeeld een werkveldcommissie).
10. Het proces resulteert in een openbaar rapport, dat in conceptvorm wordt voorgelegd aan betrokkenen voor beoordeling en eventuele correctie van ontbrekende of niet juiste feiten.
11. Er vindt enigerlei bezinning plaats bij de gevisiteerde op de uitkomsten en er is zeker 'toezicht' daarop.

e. *Het veerkrachtargument.* Visiteren verschaft het wetenschappelijk onderwijs een zachte vorm van feedback waardoor soepele adaptatie mogelijk is, die respect voor onderwijsprofessionals overeind houdt. Het is een instrument dat een zachte landing maakt. Daardoor kan het instrument populair worden. Er is geen sprake van dat de inspectie 'in de banken aanschuift' of dat een wellicht ongeschikte vorm van kwantitatieve analyse - in de

vorm van een organisatievergelijking ('benchmarking') die de kern van onderwijsprocessen niet raakt - tot verbeteren dwingt.

f. *Internationale vergelijking*. Het laatste verondersteld positief argument ter verdediging van de visitatie luidt dat deze praktijk van externe kwaliteitszorg in internationaal-vergelijkend opzicht goed uit de bus komt.

Tot zover de argumenten pro visitatie. Tot welke onderzoeksvragen geven deze aanleiding?

De vraag is of degenen die de visitatie in het wetenschappelijk onderwijs uit de periode 1985-2002 als een goede praktijk beschouwen bij nader inzien wel of geen gelijk hebben? Dat is hier aan de orde. Daarbij kijken we naar

- 1) of de doelstellingen achter onderwijsvisitaties gerealiseerd worden (verantwoorden en leren),
- 2) of de visitatiepraktijk de beginselen van behoorlijk visiteren bevestigt,
- 3) of er geen ongewenste neveneffecten optreden (Smith, 1995),
- 4) of de visitatiepraktijk in de loop der jaren 'overeind' blijft.

In het algemeen is hier aan de orde: hoe bevalt de visitatiepraktijk in het wetenschappelijk onderwijs? Voor dat aan de orde komt, gaan we eerst nog in op de kritiek op visiteren.

Mogelijk kritisch antwoord op de kwaliteit van visiteren: argumenten

Er is ook een meer kritisch beeld mogelijk over verantwoordings- en verbeterpraktijken van externe kwaliteitszorg in het wetenschappelijk onderwijs. Deze kritiek is over het algemeen breder en betreft meerdere verantwoordingsystemen in verschillende sectoren, zoals van milieuzorg in bedrijven. De kritiek is dus niet uitsluitend of voornamelijk gericht op verantwoordingsprocessen in de vorm van visitatie in het wetenschappelijk onderwijs maar betreft ook andere sectoren, zoals ziekenhuizen, rechtbanken en woningbouwcorporaties en slaat meer in het algemeen op verantwoordingsmechanismen, zoals jaarverslaggeving. Het gaat hier om een internationaal debat over 'accountability' en 'performance' & 'performance information' (o.a. Behn, 2001, 2003; De Bruijn, 2001, 2007; Halachmi, 2002a, 2002b; Noordegraaf & Abma, 2003; De Vries, 2005; Dubnick, 2005; Hood, 2007; Bemelmans-Vidéc e.a., 2007; Van Dooren & Van de Walle, 2008; Swierstra & Tonkens, 2008; Tonkens, 2008; Bovens en Schillemans, 2009; Frey & Osterloh, 2009; Van Dooren, Bouckaert & Halligan, 2010; Noordegraaf, 2011).

De bezwaren richten zich niet tegen beoordeling en verantwoording als zodanig. Beoordeling van prestaties en transparante verantwoording is verdedigbaar. Het kernbezwaar is dat de feitelijke evaluatiepraktijken eenzijdig en anderszins te beperkt zijn. Het roer moet om als serieuze verantwoording wordt nagestreefd, meent De Vries (2005).

Wat zijn nu de kritische kanttekeningen bij visiteren vanuit een *vertaling* van de discussie in de verantwoordingsliteratuur en literatuur over prestatie management naar de kwaliteit van visiteren in het wetenschappelijk onderwijs? De volgende argumenten zijn te noemen. Zet u schrap want er is sprake van een bombardement.

a. *Het argument van de beoordelingsziekte*. Gaan universiteit en wetenschap ten onder aan de beoordeelziekte 'evaluitis?', zo vragen Frey en Osterloh (2008, 2009) zich af. Of, in termen van Power (1994, 1999, 2005) aan een 'audit explosion'? Vooral Frey en Osterloh (2008, 2009) zijn kritisch. Ze brengen op basis van onderzoek naar academische prestaties de volgende argumenten in.

Ten eerste, universitair personeel krijgt niet alleen te maken met beoordeling van reguliere onderwijsprogramma's maar ook met beoordeling van andere onderwijsprogramma's zoals voor deeltijdstudenten, met evaluaties van afzonderlijke cursussen (door studenten), met evaluaties van onderzoeksprogramma's, met beoordeling van individuele aanvragen om financiering van studeerbaarheidsprojecten of onderzoeksprojecten; en dat allemaal naast de functioneringsgesprekken. Dat wordt veel, heel veel. Velen beoordelen en velen ondergaan beoordelingen. De auteurs onderkennen dat beoordelingen nodig kunnen zijn maar ze vragen zich af of het resultaat van al die beoordelingen steeds een verbetering is. Ze plaatsen daar forse kanttekeningen bij. Dat ze spreken van een beoordelingsziekte duidt er op dat hier sprake is van een ongewenste doorgeschoten praktijk. Hun tweede argument houdt in dat prestatiebeoordelingen meestal meer kosten met zich mee brengen dan gedacht. Er is ook volop sprake van *verborgen kosten*. Ten derde gaan instellingen zich verweren en het risico op een zwakke score beperken. Daarom is begrijpelijk dat betrokken universitair personeelsleden hun gedrag systematisch aanpassen en contrastrategieën ontwikkelen. Ten vierde wordt de intrinsieke werkmotivatie van universitaire medewerkers beperkt. Daardoor ontstaat er als vanzelf ook een vervolgeffect: een *'lock in'-effect*; van gedwongen meedoen aan visitaties. Veel universitaire medewerkers, zelfs verantwoordelijken voor opleidingen, hebben geen interesse in deelname maar kunnen zich er niet aan onttrekken en doen dan toch maar mee. Dat is het vierde argument, maar daar blijft het niet bij. Het vijfde argument houdt in dat de voordelen van prestatiebeoordelingen volgens hen twijfelachtig zijn. Beoordelingen leveren te weinig relevante informatie op voor de besluitvorming over verbeteringen. Er moet in feite na een rapportage van een visitatiecommissie nog een 'slag' worden gemaakt om tot verbeteringen te komen. Achter dat verbeteringsstreven gaat dan ook nog een maakbaarheidsdenken schuil alsof alles manipuleerbaar is, alsof je medewerkers met enige prikkels tot verbeteringen in onderwijs of onderzoek aanzet.

Ziedaar een forse kritiek op visiteren waartegen anderen zich wel teweer stellen. Bij voorbeeld met het argument dat er publiek debat mogelijk is over visiteren, om te zoeken naar mogelijkheden om de nadelen te beperken.

De argumentatie van beide auteurs is weliswaar indrukwekkend maar ze heeft een groot nadeel; beiden richten zich vooral op beoordeling van onderzoeksprestaties, nauwelijks op visitatie van opleidingen. En de argumentatie is vooral kwalitatief; cijfers ontbreken. Wat de feitelijke verborgen kosten zijn, wordt niet duidelijk. Wat de afname van arbeidsmotivatie is, is nog niet duidelijk. Daarmee krijgt het betoog van Frey en Osterloh iets van een onderzoeksprogramma, een analyse die vooral verwachtingen wekt. Dat neemt niet weg dat enkele argumenten ook voor externe kwaliteitszorg van opleidingen kunnen opgaan.

b. *Het argument van uitlokkend pervers gedrag*. Zijn er nog andere kritische geluiden? De Nederlandse socioloog Engbersen zegt in 2010: *'de huidige visitatie-industrie dient te worden ontmanteld'*. Hij voegt argumenten toe: *'Ze is te eenzijdig en te oneerlijk. Daarbij is zij een voorbeeld van een fatale remedie'*. De kwaliteit moet door visitaties worden bevorderd maar het omgekeerde dreigt te worden bewerkstelligd (Engbersen, 2010: 164). Zijn oplossing: maak een alternatief beoordelingssysteem waarbij aan een minimumeis in het presteren wordt voldaan.

Wat is het probleem? De stap om voortdurend niet alleen onderwijs maar ook wetenschappelijk onderzoek te willen meten in visitatieprocedures, is een *overreactie* waar tegen gewaarschuwd moet worden omdat vaak *onbedoelde gevolgen* van visitatie uitgaan die ongewenst zijn, zo waarschuwen Power en Engbersen. Wie onderzoek meet en alleen let op aantallen geproduceerde artikelen en het gemiddelde per onderzoeker en niet op het uitbrengen van boeken is eenzijdig bezig, aldus Boomkens (2008) en Engbersen (2010). Daardoor wordt onrecht gedaan aan degenen die tijden bezig zijn een meesterwerk te schrijven. Van dergelijke grote werken zijn tal van goede voorbeelden te geven. Een dergelijk

goed boek krijgt soms ook zeer veel lezers, veel meer dan er artikelen in al of niet obscure buitenlandse bladen gelezen worden (Engbersen, 2010). Wie bovendien slechts op aantallen publicaties let, moet rekening houden met *het salami-effect*: onderzoekers gaan meer en dunnere artikelen produceren alsof ze dunne plakken snijden van de salami. Zo halen ze meer uit een worst, lees een onderzoek. Of ze publiceren vaker hetzelfde maar in een iets andere vorm.

En wie op aantallen let moet weten dat het *co-auteursverschijnsel* gaat optreden: hele groepen auteurs samen gaan publiceren om elkaar te helpen om de individuele (onderzoek)scores op te hogen (Frey en Osterloh, 2008, 2009). Maar gaat het niet méér om citaties dan om aantallen publicaties? Ook hier is strategisch gedrag mogelijk, zowel op collectief als individueel niveau. Op collectief niveau kan sprake zijn van de keuze van modieuze thema's, waar veel aandacht voor is. Ook wetenschapsgebieden kennen thema's die actueler zijn dan andere. Bovendien is er het fenomeen van de *citatiekartels*: sommige beoordelaars kunnen de neiging hebben om juist die conceptartikelen te accepteren waarin zij zelf vaak voorkomen of ertoe neigen bij de herzieningsprocedure te wijzen op bepaalde 'ontbrekende literatuur' van henzelf (Frey en Osterloh, 2008). Dat betekent dat de beoordelingswijze van prestaties belangengericht gedrag uitlokt dat als pervers is te betitelen. Dat brengt ons verder bij het individuele niveau. Wie weet dat aantallen citaties van een artikel van belang zijn en in visitaties geteld worden (hoe meer geciteerd, hoe beter), weet dat hij vooral een boude, dwaze of extreme bewering moet doen in een artikel, want dan zullen veel andere auteurs daarop teruggrijpen om er afstand van te nemen en scoor je goed in de citatieoverzichten, aldus hoogleraar politieke filosofie Grahame Lock (De Vries, 2008). Het is niet uit te sluiten dat een slecht maar wel gepubliceerd artikel meer citaties oplevert dan een briljant en evenwicht artikel. Nog meer individuele effecten? Zeker. We noemen hier *het entertainmenteffect*. Wie op evaluaties let van studenten van gevolgde colleges, moet er rekening mee houden dat docenten aan - bij studenten goed vallend - entertainment (opleuken) gaan doen om de scores op te krikken, zoals hoogleraren wel tegen elkaar vertellen. Op gedoseerd entertainment is niets tegen maar die houding kan te overheersend worden. Sommige onderwerpen zijn bovendien inherent ongeschikt om 'leuk over te doen'. Metingen moeten niet naar voren brengen dat de populairste docent goed scoort maar de vakmatig en didactisch beste.

Zo zal het meer in het algemeen gaan bij onderwijsvisitaties, is de veronderstelling van critici. Hier kunnen zich tal van onbedoelde en ongewenste reacties voordoen, waardoor de beoordeling in zijn gewenste effecten niet brengt wat verwacht wordt (Frey en Osterloh, 2008, 2009). Nog een effect? We noemen dit het *vakbeschermend effect*. Als de onderwijsdecanen invloed kunnen uitoefenen op de samenstelling van de visitatiecommissie zullen ze niet een overkritische vakgenoot willen voordragen voor aanstelling in die commissie of iemand met een in hun ogen eenzijdige opvatting over het vakgebied. Men wil niet dat het vakgebied als geheel slecht uit een visitatie komt. En gedurende het proces zullen decanen ieder voor zich strategisch gedrag vertonen om goed uit de visitatie te komen, want bij een negatieve uitkomst is er werk voor elk aan de winkel. En wie wil zich zelf veel werk op de hals halen? Kortom, er wordt in de praktijk niet zo maar gemeten wat beoogd wordt om te meten en dat is 'werkelijke kwaliteit' (Jensma, 2008). Ook Engbersen wijst daarop pathetisch als hij het heeft over 'een preoccupatie met permanente evaluatie van onderwijs en onderzoek op basis van infantiele criteria en vooral ook over de Hollandse ijver om daarin veel verder te willen gaan dan bestuurders van universiteiten in andere West-Europese landen en de Verenigde Staten' (Engbersen, 2010: 162).

c. *Het argument van de disciplinering*. Boomkens (2008) kijkt in 'Topkitsch en slow science' nog breder. Hij beziet de totale ontwikkeling aan universiteiten van de laatste jaren en spreekt, in de voetsporen van Grit (2000) die promoveerde op een studie naar de bedrijfsmatige herinrichting van universiteiten, over een nieuw opmarcherend vocabulaire met puur

bedrijfsmatige termen als: onderwijsrendement, docentproductiviteit, resultaatgerichtheid, efficiency, vraagsturing, verantwoording. De universiteit en het daar gegeven onderwijs staan volgens deze cultuurfilosoof al jaren in het teken van *'een vrijwel totale disciplinerings'*. Docenten worden geknecht door bestuurders en hun paladijnen. *'Alles dient haast permanent te worden getoetst en gecontroleerd'*, aldus Boomkens (2008: 44). In termen van de eveneens kritische filosoof Grahame Lock (hoogleraar in Oxford en Nijmegen) wordt de ideale doelmatige uitkomst in het universitair onderwijs dan om met een bepaalde 'input', en vooral lage 'throughput' (lees: goedkoop onderwijs), een hoge 'output' (onderwijsrendement) te bereiken (Jensma, 2008). Minder onderwijspersoneel moet bij onderwijs efficiënt afkoersen op hoge rendementen, dus wordt kortdurend(er) onderwijs met veel afgestudeerden en weinig uitvallers de norm. Dat onderwijs ook wat mag kosten en dat een docent in 'real life'-situaties ook zijn best kan doen om studenten echt iets te leren (aan 'Bildung' te doen), desnoods buiten werktijd, geldt als een afwijking uit een voorbij tijdperk. *'Bildung is iets voor het weekend'*, aldus Lock, ditmaal in De Groene Amsterdammer (De Vries, 2008; ook in Trouw, 070411). Het ad hoc geven van onderwijs kan ook niet goed meer. De collegezalen zijn op slot, tenzij de vakgroep een zaal via de interne markt van zalenverhuur heeft gehuurd. Alles wordt 'markt', zo spot Lock ook al is het soms alleen op papier.

Zeggen de metingen dan niets? Lock heeft zelfde kritiek als Engbersen, die de criteria infantiel noemt. Lock spreekt niet van infantiel. Maar volgens Lock *'hebben de indicatoren niet met 'real life' te maken'*. *'Ze hebben alleen met zichzelf te maken'*, zo laat hij in de NRC optekenen (Jensma, 2008). De metingen schieten volgens hem voorbij aan waar het werkelijk in onderwijs om gaat en wat ertoe doet. Hij noemt de criteria waarop onderwijs beoordeeld wordt vaak *tweederangscriteria*, zoals of er genoeg onderwijslocaties zijn, of zalen schoon zijn, of de collegezaal goed verlicht is, of de powerpointpresentaties werken. Het gaat er volgens hem bij kwaliteit om of grote denkers, mensen met echte vakmatige kwaliteit, het onderwijs geven en enthousiasmeren en inspireren.

Metingen leiden niet zo maar tot meer onderwijskwaliteit want de kwaliteit van studenten zou zijn afgenomen, aldus Lock, zelfs internationaal. Visitaties kunnen zijn inziens de 'crisis' in het onderwijs, die samenhangt met de bedrijfsmatigheid op universiteiten, verkorte studieduur en minder mogelijkheden voor studenten om buiten een onderwijsprogramma te treden (wie kan zich permitteren om buiten het programma om zonder straf een half jaar te gaan studeren aan een buitenlandse universiteit?), dus niet keren.

Lock kan overigens wel begrijpen dat velen met minder kennis toe kunnen dan vroeger als gevolg van technologische ontwikkelingen. Een caissière moest vroeger nog leren tellen maar kan nu volstaan met scannen van de artikelen om de prijs vast te stellen; geschoold werk heeft plaatsgemaakt voor bijna ongeschoold (scan)werk. Zo gezien is die internationale ontwikkeling naar economisering en meer bedrijfsmatigheid in het onderwijs en de door hem gestipuleerde niveauperlaging zelfs nog te begrijpen (Jensma, 2008).

Die ontwikkeling ging, volgens Boomkens, aan universiteiten gepaard met de komst van een bureaucratische tussenlaag van ambtenaren die zich inmengen in een wereld waarin ze niet competent zijn maar wel 'formulieren' eisen waarin studenten de satisfactie met hoor- en werkcolleges of andere werkvormen aangeven. Valt er in het verband van 'turven, tellen en toetsen' en 'opgeleukt onderwijs' te objectiveren? Ook daarop heeft Boomkens kritiek. Bij het visiteren van onderwijs legt de toevallige *'make-up'* van een commissie teveel gewicht in de schaal omdat objectieve criteria niet altijd voorhanden zijn, meent hij (Boomkens, 2008: 46). Eigenlijk mondt zijn betoog, waarin Lock hem bijvalt, uit in de (vast)stelling dat sprake is van een *'normaliserende'* en *hiërarchiserende verantwoordingskarikatuur* waarin zich superieur wanende managers met protocollen vakdocenten en hun opleiding de maat nemen. Dit is een sluipende ontwikkeling geweest, waarin terechte wensen - zoals 'slechte docenten moeten

plaats maken voor goede docenten' - zijn verknoopt met naar meer bedrijfsmatigheid strevende politici die minder budget over hebben voor onderwijs en meer voor andere zaken.

Boomkens neemt overigens ook de veranderingsprocessen in het universitair onderwijs op de hak. Volgens hem is de invoering van de bachelor-masterstructuur een '*nephervorming*' geweest omdat de bachelorfase nauwelijks status heeft als eigenstandig traject waarmee men als 'bachelor' succesvol de arbeidsmarkt kan betreden. Bijna iedereen studeert door en gaat de masterfase in. De poging van onderwijsministers om via de 'Bologna'-verklaring de universiteiten te uniformeren door ze te persen in een bama-structuur is volgens hem niet alleen onzinnig geweest, maar ook, afgemeten aan de beleidsmatige veronderstelling van de eigenstandige waarde van bachelordiploma's, een fiasco gebleken (2008: 69).

Conclusie: De opgesomde kritiekpunten zijn een ware aanklacht. Visiteren past in een ontwikkeling naar meer bedrijfsmatigheid die wel te begrijpen is, maar niet zonder meer positief te duiden is, aldus de critici. De echte kwaliteit komt meestal niet in beeld en meetsystemen hebben vaak onbedoelde, negatieve effecten. Deze kritiek viel in de periode 2008-2011 te beluisteren (zie ook Engelen, 2012), maar onze analyse richt zich op een eerdere periode 1985-2003. Daarmee is die kritiek niet vanzelfsprekend ook van toepassing op die eerdere periode. Maar toch is die prikkelend genoeg om hier te vermelden.

Een tweede weerwoord naar de critici houdt in dat de onderwijsprofessionals *zelf* volop betrokken zijn geweest bij de opzet van de onderwijsvisitaties in vakgebieden in de periode 1985-2003. Blijkbaar deelden de meeste verantwoordelijke onderwijsdecanen in die periode de latere kritiek van Boomkens, Lock en Engbersen niet of niet helemaal.

Een derde reactie behelst dat een groot deel van de kritiek veel meer betrekking heeft op onderzoekvisitaties dan op onderwijsvisitaties.

d. *Het argument van de (overdreven) luxe.* Een variant op de in het voorgaande weergegeven kritiek luidt in de periode 1985-2003: visiteren van opleidingen is in macroperspectief te zien als een (overdreven) luxe in een welvaartsmaatschappij waarin de Nederlandse overheid blijkbaar over voldoende budget beschikt om er een '*feedback- en toezicht*'-industrie voor hele branches als het w.o. en hbo op na te houden (Hood, Dixon & Beeston, 2008). De vraag is of er in Nederland überhaupt een visitatiebestel gericht op visitatie van opleidingen nodig is omdat de kwaliteit van het wetenschappelijk onderwijs al boven de streep is, volgens sommigen. Een beoordelende onderwijsinspectie zou kunnen volstaan met een soberder analyse, namelijk of de interne kwaliteitszorg van onderwijsinstellingen adequaat is (zie ook Frey & Osterloh, 2008, 2009). Wie werkelijk wil 'leren' en opleidingen wil verbeteren, zou moeten profiteren van de *interne* kwaliteitszorg en daarover kan soberder externe verantwoording worden afgelegd dan nu het geval is. Er is dus een alternatief voor het visitatiebestel. Een ander bescheidener alternatief is visiteren van faculteiten of instellingen, dus de beoordeling naar naar een hoger aggregatieniveau tillen.

Een tweede argument dat duidt op 'luxe' is dat vanuit een perspectief van visitatie en rankings een aantal regeringen van landen dat de Bologna-verklaring heeft getekend - een verklaring die gericht was op de komst van een bachelor-masterstelsel en meer mobiliteit van docenten en studenten -, toch niet direct zal kunnen voldoen aan een niveau van beoordelen dat gelijk is aan het Nederlandse, met visitaties van opleidingen. Na de ondertekening van de Bologna-verklaring hebben inmiddels meer dan 45 landen zich hierbij aangesloten, waaronder Roemenië, Albanië, Kroatië, Estland, Turkije. In totaal gaat het om meer dan 1500 opleidingsinstituten. Zouden in landen als Roemenië en Albanië individuele wo- opleidingen - if any - allemaal gevisiteerd worden als de Nederlandse? En hoe zit het in Turkije, een land dat veel en veel meer universiteiten heeft dan Nederland?

Een derde argument betreft de vraag of de visitatie-‘industrie’ een omvangrijk proces van leren en verbeteren produceert? Critici als De Vries (2005) betwijfelen of omvangrijke verbetertrajecten in de verschillende onderwijsinstellingen plaatsvinden, want docenten ervaren visitatie als een irritante onderbreking van het eigenlijke werk. Dit argument keert verderop nog in iets andere vorm terug in verband met een verwachte matige benutting van uitkomsten van visitaties en als kosten-batenargument.

Hier staat overigens een argument tegenover, namelijk dat de visitaties in Nederland slechts een fractie zouden zijn van de kosten die in de Verenigde Staten aan externe kwaliteitszorg worden uitgegeven (Hulshof e.a., 2003).

e. Het argument van de *zwakke legitimatie van de starter*. Het is steeds een ander, een derde instantie, die een ander de maat neemt door het visiteren te initiëren, maar zelf een dergelijk regime niet of moeizaam vrijwillig over zich zelf afroept. Laat het departement van onderwijs en wetenschappen, de VSNU of onderwijsinspectie zich zelf ook visiteren op basis van zelfstudies? Dit argument - niet ‘zij zelf’ maar andere organisaties moet de maat genomen worden - gaat ervan uit dat visiteren als *onvrijwillig en paternalistisch* en/of *belastend* en/of als *nutteloos* wordt ervaren door al degenen die bij de opstelling van een zelfstudie betrokken zijn en de visitatie moeten ‘ondergaan’ (De Vries, 2005). Universitaire instellingen die gevisiteerd worden, zijn - ondanks hun status van zelfstandig bestuursorgaan (zbo) - in deze argumentatie eigenlijk slachtoffers van een centrale overheid die toch weer niet kan ‘loslaten’ door vertrouwen te schenken.

f. *Visiteren in het wetenschappelijk onderwijs is politiek speelgoed*. Visiteren moet in deze tijd nu eenmaal omdat de centrale overheid met interne kwaliteitsbeoordeling niet wenst te volstaan en de overheid onderwijs (inclusief kwaliteitsbeoordeling) niet geheel aan de universitaire vakmensen, zoals docenten, wenst over te laten (verg. Noordegraaf, 2008). Universitair onderwijspersoneel moet een inkijk geven, zodat een aangrijping mogelijk wordt om - bij stijgende collegegelden - kwaliteitsverhoging van het onderwijs te eisen of als politieke ruil toe te zeggen. Kwaliteitsbeoordeling wordt zo ‘politiek speelgoed’. Zoals Power zegt: ‘It is part of the politics’ (Power, 2005: 342).

Er is in de ogen van critici nog een ander politiek argument: het initiëren van verantwoordingsprocessen poetst het bord van de rijksoverheid schoon als incidenten in het onderwijs blijken. Aan het ministerie van OC&W heeft het dan niet gelegen want die had immers kwaliteitszorg mede mogelijk gemaakt. Visiteren valt dus ook te zien als een *indekkingsoperatie* in tijden dat media op zoek zijn naar schandalen en het parlementair activisme in de vorm van onderzoeken en parlementaire enquêtes toeneemt.

g. *De zwakte van het integraal beeld*. De echte feiten over het te evalueren object komen niet voldoende naar boven, doordat de visitatie versnipperd raakt in onderdelen en/of van ‘*checklist-denken*’ (afvinken) bij visiterende commissies sprake is. Visitatiecommissies en andere evaluatoren krijgen het integrale beeld van de opleidingskwaliteit niet goed te pakken. Ze bezien onderwijsaspecten niet voldoende in samenhang.

h. *Het argument van de matige toegevoegde waarde*. Organisaties kennen altijd bepaalde zwakten en dat is ook bekend. Of het nu om een krantenredactie, politiekorps, rechtbank, ziekenhuis of universiteit gaat, er is altijd wel iets dat niet helemaal functioneert zoals het zou moeten (*‘ijzeren wet van de suboptimaliteit’*). Dat komt deels door veranderprocessen of andere dynamiek. Universitaire veranderprocessen kunnen inhouden dat de onderwijskundige aanpak verandert of een vak een andere inhoud kreeg en nu voor het eerst wordt gegeven in de nieuwe vorm (kinderziekten?). Of, goed personeel vertrekt en het nieuwe personeel is nog niet

aanwezig wat ten koste gaat van de onderwijskwaliteit. Of, als gevolg van ziekteverzuim worden invallers gezocht en die zijn niet altijd zo goed als de zieken. Al dit soort situaties en processen impliceren dat een visitatie ook altijd wel een suboptimaliteit constateert. Waar het nu om gaat is, dat de rapportages en uitkomsten van visitaties te weinig uitstijgen boven de bij de betrokken onderwijsgeevenden daarover al aanwezige 'common sense' en over onderwijskwaliteit in het algemeen. De nieuwsaarde is voor insiders in het wetenschappelijk onderwijs (te) beperkt. Zij hebben de visitaties zelf niet nodig om zwakten te kennen en te verbeteren. Er bestaat immers al interne kwaliteitszorg. Goede interne kwaliteitszorg drijft externe uit.

i. *Weinig leereffecten.* Zelfs als er enige toegevoegde waarde uit visitatierapporten naar voren komt, dan nog: het vermogen om te leren als gevolg van 'feedback', argumentatie en dialoog is beperkt (Frey en Osterloh, 2008, 2009). Er gaan te weinig leereffecten uit van evaluaties, aldus Van der Knaap (1997); de benutting van visitatierapporten (of het gebruik) is in het algemeen te gering of beperkt tot onderdelen. Dat komt mede omdat vakgroepen onderwijs claimen om te overleven en de samenstelling van onderdelen van een opleiding een compromis is, waarbij elke vakgroep een deel van het werk krijgt. Het goede of beste onderwijs bieden is dus niet de enige prikkel bij de bouw van onderwijsprogramma's;

j. *De adresseringskwestie.* De verantwoording door visitatie is betrekkelijk betekenisloos omdat de verantwoordelijke voor zwakheden zelden kan worden aangewezen. Verantwoordelijkheid is in het wetenschappelijk onderwijs versnipperd over een college van bestuur van de universiteit, faculteitsbesturen, decanen, opleidingscoördinatoren, examencommissies, commissies van bezwaar en beroep; hoogleraren en andere onderwijsgeevenden.

k. Voorgaande argumenten cumuleren in *het kosten-batenargument*. Visiteren van opleidingen in het wetenschappelijk onderwijs zou tot bureaucrativering in het onderwijs leiden of een groot risico hierop (verg. Onderwijsraad, 2004; Elte, 2007).

Bureaucrativering heeft drie betekenissen. Bureaucratie als overdreven papierwerk, traagheid, ondoelmatigheid, het ontbreken van resultaat. Bureaucratie als macht van grote organisaties, zoals een onderwijsinspectie of VSNU. Of bureaucratie als formeel systeem met visitatieregels die kunnen leiden tot het blokkeren van de bekostiging van opleidingen of accepteren van positief toezicht.

De eerste en derde betekenis komen hier samen. Bij de VSNU zijn ambtenaren actief om de visitaties te ondersteunen en bij de instellingen zelf kost het opstellen van zelfstudies tijd en energie die onttrokken wordt aan onderwijs en onderzoek. En er is sprake van betrokkenheid op allerlei momenten bij het visiteren: bij de samenstelling van de commissie, het bezien van het referentiekader, het te woord staan van de commissie, het becommentariëren van een conceptrapport (Ehren, 2007: 201).

Wat is nu de kritiek? Het eerste argument is een macroargument: de 'verantwoordings- en controledrift' zuigt tijd, menskracht en budget weg uit de primaire processen van onderwijs en onderzoek. Wat aan tijd opgaat aan verantwoording kan niet ingezet worden voor andere doelen.

Het tweede argument betreft meer een microargument: een concrete visitatie is over het algemeen te duur ('het kost veel tijd om zelfstudie per opleiding te maken') ten opzichte van wat het oplevert (kosten-batenargument).

Deze argumentatie vinden we in varianten in publicaties of interviews van de filosofen Boomkens (2008) en Lock (2008, 2009). Verantwoording is wellicht beter mogelijk, zo luidt de reactie van critici, door een instelling door te lichten in plaats van tal van visitaties per faculteit

of opleiding uit te voeren. Bij de keuze van een hoger aggregatieniveau van analyse ('visitatie') zullen de visitatiekosten vermoedelijk aanzienlijk dalen.

1. *Het tijdelijkheidsargument.* Tenslotte is er nog een curieus tegenargument. Dat houdt in dat visiteren nog eens over de toch altijd aanwezige dynamiek heen komt. Zelfs als er geen visitatie is, zal onderwijs nooit blijven wat het is door bewegingen in de studentenpopulatie (o.a. door verschuiving in marktaandeel), technologische ontwikkelingen (opkomst ICT in het onderwijs) of uitvindingen (bij voorbeeld de rol die de opkomende biotechnologie in bepaalde studierichtingen had en heeft) of ontwikkelingen op wereldschaal in handel en industrie (denk aan globalisering). De argumentatie luidt verder als volgt. Degenen die een lange termijnperspectief kiezen en op de hoogte zijn met *beleidserfenissen* en *slijtage van beleid* benadrukken dat een beleidsimpuls als het visiteren van afzonderlijke opleidingen slechts een tijdelijke eruptie in onderwijsbeleid is. Misschien is visiteren wel nuttig of juist niet nuttig, het verdwijnt vanzelf. Vijftien jaar na de start van een beleid zal een beleidstraject heroverwogen zijn of heroverwogen worden of, als gevolg van allerlei omstandigheden, afgeschaft of aangepast zijn. Dan is er wellicht wel Europees hogeronderwijsbeleid dat het Nederlandse beleid deels heeft opgeslokt. Of het ene beleid is opgevolgd door het andere en heeft dus een soort evolutie ondergaan (Van der Knaap, Korsten, e.a., 2004). Beleidsopvolging ('Policy succession') komt heel vaak voor, zowel binnen instellingen als hele stelsels, constateren Hogwood & Peters (1983). Beleidsopvolging kan impliceren dat een beleid opgaat in een ander beleid, de opvolger. Anders gesteld, wat van waarde is, verdwijnt en wat niet van waarde is, verdwijnt ook in de oorspronkelijke vorm. 'Niets is vast, alles beweegt'. De benadrukking van het tijdelijke, is het cynisme van de langetermijndenkers. Aan het adagium 'behoud het goede' of 'verbeter steeds' hebben deze cynici geen boodschap.

Tot zover de betrekkelijk forse (nog *niet* empirisch gefundeerde) kritiek op visiteren in het wetenschappelijk onderwijs die wij mede baseren op de meer algemene verantwoordings- en feedbackliteratuur. Deze kritiek is zelden zo scherp in deze vorm aan te treffen in de literatuur over hoger onderwijs (zie het blad TH&MA). Maar we ontlopen die kritiek niet. We lieten die in volle glorie de revue passeren omdat die de analyse direct op scherp zet.

Centrale aanname

De positieve en negatieve beschouwingen hebben als gemeenschappelijke beperking dat ze beknopt zijn ('kort door de bocht') en vooral dat ze niet breed gefundeerd zijn op basis van onderzoek naar de kwaliteit van visitaties in het wetenschappelijk onderwijs en hoger beroepsonderwijs. In dit stuk gaan we daarom in op de kwaliteit van visiteren in het wetenschappelijk onderwijs, door de bezinning hierop breder te funderen. Wat blijkt zoveel jaar nadat de contouren van deze uitspraken zichtbaar werden? Er moet inmiddels meer informatie voorhanden zijn om gefundeerde uitspraken te doen.

De centrale aanname is dat, mochten visitaties in het wetenschappelijk onderwijs (uit de periode 1985-2003) qua doelstelling, opzet, beginselen en uitvoering positief zijn te beoordelen, ze daarom in andere gebieden, zoals bij visitatie van gemeentebesturen, navolging verdienen. Als de kritiek daarentegen overheerst, kunnen visitaties in het wetenschappelijk onderwijs geen referentiepunt zijn voor visitaties in andere sectoren. Een eerste reflectie hierop is overigens in dit stuk zijdelings aan de orde omdat we kijken of visiteren in het hbo qua beginselen van behoorlijk visiteren spoort met visiteren in het w.o.

Inperking en waarnemingsbronnen

Hier komt het (oude) landelijke visitatiestelsel vóór de invoering van de accreditatie aan de orde. Het is ons erom te doen te achterhalen hoe die zuivere visitaties van een aantal jaren terug - uit grofweg de periode 1985-2003 - 'werkten' en wat de kwaliteit en doorwerking

daarvan is geweest. We gaan ervan uit dat daarover inmiddels enige inzichten zijn verkregen. We spreken van een stelsel. Blijken zal waarom. De evolutie van visitatie in een uitwerking naar accreditatie laten we hier terzijde omdat we accreditatie niet kunnen gebruiken voor onze analyse van visitatiestelsels in andere sectoren (Douma, 2004). We onderzoeken de pure vorm van visiteren omdat accrediteren elders, zoals bij visitatie van gemeentebesturen of van politiekorpsen, niet voorkomt.

Om antwoord te geven op de vraagstelling is gebruik gemaakt van de volgende bronnen.

- 1) *studies naar onderwijsbeleid en sturing*. Deze bron dient om de achtergrond van visitatie in het wetenschappelijk onderwijs te begrijpen;
- 2) *gegevens van de VSNU*. Denk raadpleging van de basisprotocollering van de VSNU en van de inhoud van een aantal zelfstudies en visitatierapporten op het gebied van economie, bedrijfskunde en bestuurskunde. Deze bronnen zijn van belang om de stappen in de procedure te leren kennen en controleren;
- 3) *participerende observatie*. benutting van ervaringen met visitatie in het wetenschappelijk onderwijs en participatie in de periode 1986-2006 in de voor bestuurskunde en bedrijfskunde-opleidingen relevante disciplineoverlegorganen. Het doel is om vast te stellen of en in hoeverre sprake is geweest van betrokkenheid van decanen en andere verantwoordelijken bij de start van de visitatie en de opzet, en bij de samenstelling van de visitatiecommissie. Daartoe zijn vergaderverslagen geraadpleegd en is informatie van enkele leden (vanuit de OU) gebruikt;
- 4) *informatie uit 'een belendend perceel': het hbo*. De uitvoering van een visitatie in het hoger beroepsonderwijs als voorzitter van een visitatieteam (Korsten) leverde indirect informatie op om zo te zien of procedure-elementen uit het w.o. (beginselen) en oordelen over de w.o.-aanpak ook bevestiging vinden in het hbo;
- 5) *raadpleging van jaarverslagen van de Inspectie van het Onderwijs*. Deze bron kan ons inzicht verschaffen in de beoordeling van de kwaliteit van de visitaties volgens de inspectie;
- 6) *raadpleging van specifieke empirische studies*. Denk aan studies naar aspecten van visitaties, zoals van de Algemene Rekenkamer en duiding daarvan door enkele gesprekspartners. In dat verband zijn in 2012 ook enkele korte gesprekken gevoerd met Theo Douma (o.a. oud-ambtenaar bij OC&W; directeur ROC Nijmegen; auteur) en Wienke Blomen (directeur Hobeon) om te achterhalen welk beeld zij hebben van de kwaliteit van visitaties en om te zoeken naar additionele bronnen die zij van belang achten.

De indeling van dit stuk impliceert dat nu eerst teruggedaan wordt naar de omslag in het wetenschappelijk onderwijs en de opzet van de visitatie om eerst daarna in te gaan op ervaringen met visiteren in de drie rondes van visiteren in het w.o. uit de periode 1985-2003.

3 Visiteren in het wetenschappelijk onderwijs: eerste verkenning

Vóór 1985 was onderwijskwaliteit wel een aandachtspunt van universiteiten en hogescholen, maar was er volgens ons geen sprake van een systematische aanpak door middel van kwaliteitszorg. Er was volgens Kwikkers e.a. (2011: 30) sprake van een informele bekommernis 'binnen de vanouds internationale academische gemeenschap, zonder binding met institutioneel beleid of reguleringsmechanismen, zoals overheidserkenning'. In die jaren werden er dus alleen interne kwaliteitszorgsystemen gebouwd. Kwaliteitszorg stond in die tijd los van vormen van regulering. Er was geen consequentie en geen rechtsgeldig gevolg

verbonden aan vroege vormen van evaluatie, onder meer van opleidingsprogramma's of cursussen. Dat veranderde rond 1985. De kwaliteitszorg in het hoger onderwijs kende vanaf die tijd een fors groeitempo.

Nederland beschikt vanaf 1985 in het academisch onderwijs, hier gelijk te stellen aan wetenschappelijk onderwijs (w.o.) of universitair onderwijs, over een visitatiestelsel van afzonderlijke, geregistreerde opleidingen; ook van onderzoek, maar dat blijft hier terzijde. Commercieel gerichte (niet-reguliere) onderwijsprogramma's bleven bij die visitatie buiten beschouwing.

Dit visitatiestelsel, dat ook in het hoger beroepsonderwijs (hbo) bestond en bestaat, is een stelsel van externe kwaliteitszorg per opleiding, dus van wo-opleidingen op het gebied van studierichtingen als bij voorbeeld scheikunde, biologie, rechtsgeleerdheid, bouwkunde, aardwetenschap, tandheelkunde of bestuurskunde. Deze kwaliteitszorg houdt in dat *elke* studierichting uit het wetenschappelijk onderwijs sinds 1985 om de vier of vijf jaar beoordeeld wordt. Daarvoor wordt een *visitatiecommissie* ingesteld die als bron voor de beoordeling gebruik maakt van een *zelfstudie* en van *bezoeken ter plaatse*. De commissie brengt een *verslag* uit *per instelling* en een *openbaar overallverslag* uit van bevindingen over de kwaliteit van het regulier onderwijs in het vakgebied.

Het visiteren is dus niet per universiteit of per faculteit maar per vakgebied georganiseerd en wel door de branche- en belangenorganisatie van de instellingen, de Vereniging van Samenwerkende Nederlandse Universiteiten (de VSNU). Dat geschiedt 'volgens een zelfontwikkeld, tamelijk gedetailleerd protocol', aldus Drooglever Fortuijn (2003: 241) in zijn boek *Onderwijsbeleid*.

Een centrale unit bij de visitatie is dus de VSNU, die als *procescoördinator* bij de organisatie van visitaties optreedt (Laman, 2003: 38; Douma, 2004). Ook *de inspectie* heeft een rol. Het is de Inspectie van het Onderwijs die na afloop beziet of de visitatie deugdelijk is geweest. De onderwijsinspectie vervult ook nog een andere rol in dit stelsel. 'Een publiek verslag met bevindingen en aanbevelingen vormt voor de inspectie van het hoger onderwijs de basis om binnen een jaar bij elke instelling na te gaan wat er met de aanbevelingen is gedaan. Als een visitatiecommissie ernstige tekortkomingen heeft geconstateerd en de inspectie van oordeel is dat er te weinig is gedaan om verbeteringen aan te brengen, kan zij de minister adviseren bekostiging van de opleiding in te trekken. Zover is het in 2003 nog niet gekomen, de dreiging is meestal voldoende om een instelling tot verbetering aan te zetten', aldus Drooglever Fortuijn (2003: 241). Het visiteren in het w.o. is dus *geenszins vrijblijvend*.

Dit visitatiestelsel is overgegaan in een accreditatiestelsel. De aanzet daartoe vinden we in de nota *Keur aan kwaliteit* uit 2000, die door toenmalig minister Hermans is aangeboden aan de Tweede Kamer. Daarop is later wetgeving gevolgd om accreditatie ook wettelijk te verankeren. In het accreditatiestelsel, dat op 1 januari 2002 is ingevoerd, is ook visitatie voorzien maar dan niet volgens een landelijk stamien van gelijktijdige beoordeling van alle 'gelijke' opleidingen op een gebied. Ging bij de vakgebiedsgerichte visitatie oorspronkelijk het (start)initiatief uit van de VSNU namens de gezamenlijke colleges van bestuur van de universiteiten, bij accreditatie gaat het initiatief uit van de instelling zelf die een opleiding tegen het licht wil laten houden. Instellingen kunnen kiezen uit enkele erkende beoordelende organisaties (VBI's) om een opleiding te laten accrediteren. Tegenwoordig komen overigens ook varianten op opleidingsaccreditatie voor. Een variant is de instellingsaccreditatie of een getrapte vorm van beoordeling.

Positieve accreditatie van een universitaire opleiding is tegenwoordig voorwaarde voor het recht van de onderwijsinstelling op bekostiging, het verlenen van getuigschriften waaraan de wet civiel effect verbindt, voor aanspraken van studenten op studiefinanciering en het voeren van titulatuur. Nieuwe opleidingen worden vooraf aan een toets onderworpen (Hulshof e.a., 2003: 31). Non-accreditatie zal in vrijwel alle gevallen leiden tot intrekking van de bekostiging. Slechts op enkele gronden kan de minister de non-accreditatie negeren (Hulshof e.a., 2003: 31).

Is visitatie vóór de invoering van accreditatie van een opleiding gericht op *verantwoording* van de opleidingskwaliteit en *kwaliteitsverbetering*, met de invoering van accreditatie wordt de kwaliteitsverbeteringsdoelstelling van een opleiding losgelaten (Hulshof e.a., 2003: 31)

Er bestaan met de invoering van de accreditatie meer vrijheidsgraden per instelling. Er wordt in dat beoordelingsproces meer rekening gehouden met strategische doelstellingen van de instelling respectievelijk de opleiding, die onderscheidend zijn ten opzichte van doelstellingen achter opleidingen met dezelfde naam aan andere universiteiten. Onderscheid is mogelijk naar bij voorbeeld de taal waarin het onderwijs gegeven wordt en de onderwijsdoelgroepen. Accreditatie geschiedt overigens onder auspiciën van de Nederlands-Vlaams Accreditatie Orgaan (NVAO) (Dittrich, 2010).

We gaan het hier hebben over de kwaliteit van visitatie en verder niet over accreditatie. De reden daarvoor is dat deze analyse gericht is op verder onderzoek naar visiteren in andere sectoren en niet op accreditatie.

4 Breukpunt in hoger onderwijsbeleid: streven naar universitaire autonomie

Een breukpunt in het denken over wetenschappelijk onderwijs, het w.o., was - zoals gesteld - de HOAK-nota van het ministerie van O&W uit 1985. Minister Deetman was voor de nota de verantwoordelijke minister. Deze nota handelde over autonomie en kwaliteit in het hoger onderwijs, dus inclusief het hoger beroepsonderwijs (hbo). Het hbo viel tot dan nog onder het schooltoezicht voortgezet onderwijs, terwijl in het w.o. het bijzonder deel van het zgn. Academisch Statuut het middel was om de kwaliteit van het onderwijs te garanderen. Dit deel bevatte voorschriften over de inhoud van onderwijsprogramma's in het w.o. Het vaststellen van nieuwe inhoud of een nieuwe studierichting was een starre zaak; wijziging vergde een lange procedure. Instellingen moesten een hele tijd wachten alvorens er goedkeuring was verkregen voor een verandering (een aanpassing aan veranderde inzichten) of een nieuwe richting van start kon gaan (Hulshof e.a., 2003: 22). Van inspectie in het w.o. was toen eigenlijk geen sprake en van visitaties en zelfevaluaties (ook wel zelfstudies te noemen) evenmin.

De beweging die gemaakt werd door middel van de HOAK-nota was dus dat instellingen *meer autonomie* kregen. Verondersteld werd dat aan hun deskundigheid en de gewenste flexibiliteit meer ruimte geboden moest worden.

Om toch enige sturing van ministeriële zijde mogelijk te maken ('sturen op afstand' genoemd) die zich niet teveel zou richten op kleine onderwijseenheden werd een hoger aggregatieniveau gekozen, namelijk dat van de sturing op negen onderwijssectoren, waarin opleidingen geclusterd waren. Die besturing had daarmee veel weg van - zoals gesteld - meer autonomie aan de instellingen, maar ook van meer deregulering en decentralisatie. Argument: het hoger onderwijs moest meer kunnen inspelen op omgevingsontwikkelingen, dus op veranderingen in onder meer de samenleving, in het beroepenveld, in de economie en in de technologie. Het werd niet meer mogelijk geacht dat onderwijsprogramma's jarenlang gelijk bleven. Er moesten bovendien prikkels voor innovatie mogelijk worden (Goedegebuure e.a, 1990; Lips, 1996).

Het oude sturingsmodel voor w.o.-onderwijs wordt met de HOAK-nota dus ingeruild voor een nieuwe sturing van planning vooraf en controle achteraf. Deze nota heeft geleid tot de Wet op het hoger onderwijs en wetenschappelijk onderzoek (WHW) die in 1993 van kracht werd. Daarmee werd het aangrijpingspunt van sturing verschoven van condities naar prestaties. De sturing werd minder gericht op de voor universiteiten interne proceskarakteristieken en meer op het resultaat ervan (Huisman en De Vijlder, 1999: 358).

De grote man achter dit HOAK-beleid van sturen op afstand was Directeur-generaal Hoger Onderwijs en Wetenschappelijk Onderzoek Roel in 't Veld, kenner van sturingsmodellen en eerder werkzaam op de afdeling planning van de Universiteit van Leiden en als hoogleraar aan de Katholieke Universiteit Nijmegen (Korsten, 2005a; Bekker, 2012; Knippenberg en Van der Ham, 1994: 689 e.v.). Op de HOAK-nota en het vervolg gaan we hier niet in algemene zin in. We verwijzen daarvoor naar studies over sturing op afstand als van Mirco Noordegraaf en Walter Kickert (1993), Roel in 't Veld (2001a, b) en naar de dissertaties over sturing en autonomie van Wil Foppen (1989), Victor Bekkers (1994), Theo Camps (1995), Miriam Lips (1996) en Olav McDaniël (1997). Deze studies laten zien dat toch sprake is geweest van worsteling op rijksniveau om bij sturen op afstand daadwerkelijk los te laten. Ministers neigen ertoe om als er vragen komen uit de Tweede Kamer niet te zeggen dat de minister er niet meer over gaat, maar om toch weer in de instellingen informatie te gaan halen. Die informatievoorziening neigt tot centralisme, zowel tussen ministerie en instellingen en binnen instellingen zelf (Bekkers, 1994). Dit leidde nogal eens tot strijd tussen de minister en de VSNU.

5 Contouren van het visitatiestelsel in het wetenschappelijk onderwijs

In het nieuwe sturingsmodel was plaats voor visitatie in het w.o. en hbo (door de sector zelf) waarbij de rol van de overheidsinspectie er een is van 'op de achterhand zitten'. 'De rol van de overheid bij de kwaliteitsbewaking veranderde van preventief naar sanctionerend: 'Van prescriptie vooraf naar controle achteraf', aldus Hulshof e.a. (2003: 22). De verhouding overheid-veld veranderde van unicentrisch naar pluricentrisch, met name met betrekking tot de kwaliteitszorg.

Opvallend aan de aanzet tot visiteren in het w.o. is dus dat er een initiatief was aan de kant van het ministerie (HOAK), een optiek over interbestuurlijke sturing tussen nationale overheid en onderwijsinstellingen in het w.o. (sturen op afstand) en er druk op de ketel bleef doordat de inspectie zou toezien op de kwaliteit van visitaties van w.o.-opleidingen. Aan door de inspectie erkende 'goede' visitaties konden consequenties worden verbonden in termen van verbetering of afschaffing van opleidingen. Er was dus in principe voorzien in sanctiemogelijkheden mocht een instelling de visitatie-uitkomsten niet serieus wensen te nemen.

Na het uitkomen van de HOAK-nota in 1985 werd het visitatiestelsel opgebouwd en uitgewerkt. Dit stelsel werd in de kern een zaak van 'peer review' (door vakgenoten) bij opleidingen. Het zijn de externe deskundigen die in de vorm van een externe visitatiecommissie naar 'hun' opleidingen kijken, waarvoor ze kunnen beschikken over een zelfevaluatie (of zelfstudie) per instelling. Een visitatiecommissie met hoofdzakelijk deskundige scheikundigen kijkt dus naar alle w.o.-opleidingen scheikunde in den lande en spreekt daarover een kwaliteitsoordeel uit.

'Doordat de externe kwaliteitszorg werd georganiseerd door het veld zelf, kon het visitatiestelsel rekenen op (tenminste enige) acceptatie in het veld', aldus Hulshof e.a., 2003: 23).

Het doel van de visitatie in het w.o. is vooral geweest: (externe) verantwoording en kwaliteitsverbetering (Hulshof e.a., 2003: 23). Door middel van openbare rapportages leggen instellingen verantwoording af over de onderwijskwaliteit. Als de interne kwaliteitszorg goed op peil is, zal een instelling waarschijnlijk goed kunnen scoren bij visitaties en zijn er ook voorwaarden aanwezig om later verder te verbeteren.

De toepassing van visitaties in het w.o.

In 1986 maakte de minister met de instellingen de afspraak voor een systeem van kwaliteitszorg dat 'flexibel, dekkend en doorzichtig' zou moeten zijn (Laman, 2003: 37). De universiteiten gaven in 1987 de eerste 'Gids voor de Onderwijsvisitatie' uit en startten een pilot. In 1988 volgden de eerste visitatieprojecten en er zou een periode volgen van meer dan tien jaar visiteren, zonder dat het accreditatiesysteem bestond. De periode 1988-1993 is te beschouwen als de eerste visitatieronde (Laman, 2003) omdat toen alle w.o.-opleidingen tegen het licht gehouden zijn.

Dakpansgewijs en cyclisch

Dit visitatiestelsel van afzonderlijke opleidingen bestond uit globale afspraken tussen de colleges van bestuur van alle (reguliere) universiteiten (en technische hogescholen) om de kwaliteitszorg te laten uitvoeren via de VSNU, het samenwerkingsverband van universiteiten, en alle w.o.-opleidingen dakpansgewijs te beoordelen: elk jaar een aantal te beoordelen opleidingen zodat elke opleiding ongeveer eens per vijf jaar tegen het licht wordt gehouden. Hierdoor ontstaat ook een cyclus van beoordelingen. Het beoordelen van opleidingen is cyclisch en eindigt in beginsel niet; na vier of vijf jaar is een opleiding opnieuw aan de beurt en zo verder.

Procescoördinator

De VSNU bestaat bestuurlijk gezien in de kern uit de voorzitters van de universiteitsbesturen samen: het algemeen bestuur onder leiding van een boven de partijen staande voorzitter (tegenwoordig). De VSNU wordt eind jaren tachtig en in de jaren negentig ambtelijke ondersteund door een staf met medewerkers.

De VSNU is dan ook het kader voor overleg tussen decanen van opleidingen. Die zijn in disciplinaire commissies verenigd en worden ambtelijk ondersteund door medewerkers van de VSNU. Heel veel beleid maken die decanen niet gezamenlijk want ze staan onder leiding van de faculteitsbesturen en de colleges van bestuur van hun eigen universiteit. Ze hebben dus weinig vrijheidsgraden voor zelfstandig handelen en moeten al snel terug naar hun besturen voor rugdekking, wat af en toe ook gebeurt.

Protocol, commissie en verder

In het kader van de VSNU is een protocol opgesteld van de gang van zaken voor alle visitaties, een kaderstelling dus over de start, de sturende rol van een onafhankelijke visitatiecommissie, de fasering van het gehele proces, de zelfevaluatie per opleiding en de toets op volledigheid van die zelfevaluatie, de 'site visits', de terugkoppeling, het conceptrapport per instelling en de mogelijkheid tot feedback richting commissie, een eindrapport en de openbaarmaking daarvan. Dat protocol staat vast, de inspectie erkent dit en de disciplinaire commissies van decanen kunnen dat niet veranderen. Met eventuele bijzonderheden van het soort opleiding kan tijdens de toepassing nog wel rekening gehouden worden. Zo zal tijdens site visits bij een scheikundeopleiding op enkele andere aspecten gelet worden, zoals bij voorbeeld op ruimten voor het doen van proeven, dan bij een opleiding rechtsgeleerdheid, die ook een oefenrechtbank kent.

De VSNU is het scharnierpunt in het aanzoeken van leden van onafhankelijke visitatiecommissies, die op basis van zelfevaluaties van instellingen ter plaatse gaan kijken. De disciplinaire commissies met decanen doen suggesties voor de samenstelling van de visitatiecommissie en waken voor belangenverstremming tussen elk lid en de te visiteren instelling (er mogen geen banden zijn).

Het is steeds een visitatiecommissie die zich met een aantal opleidingen op het zelfde gebied inlaat. Elke commissie bestaat doorgaans uit vakgenoten die afstand hebben tot de te visiteren instellingen, alsmede uit een onderwijskundige, een student, een vertegenwoordiger van het beroepenveld en een secretaris.

Ondersteuning

De visitatie kent in essentie dus een onafhankelijke visitatiecommissie, die overigens niet wordt ondersteund door een adviesbureau maar door iemand van de VSNU. Een rol voor meer mensen dan een ambtelijk secretaris hoeft ook niet want instellingen verrichten zelf op basis van een protocol (dat de VSNU heeft opgesteld en dat zo nodig wordt gespecificeerd via de disciplinecommissie van de opleidingen - dus het landelijk verband - die worden gevisiteerd) een uitgebreide zelfevaluatie. Veel werk wordt dus door de instellingen zelf verricht. In de praktijk is het de decaan die de zelfevaluatie opstelt met medewerking van stafleden. De zelfevaluatie is daarmee doorgaans een gedragen product. Belangrijke betrokkenen in het onderwijsproces kennen en erkennen de inhoud.

Zelfstudie

Voor de opstelling van een zelfstudie of zelfevaluatie geldt een inhoudelijk kader ('een frame') waarin staat wat aan bod moet komen. Daarbij geldt een datum waarop de dan geldende situatie moet worden beschreven en van kwaliteitsoordelen voorzien. Wat zo'n protocol dan verwacht dat in een zelfevaluatie aan bod komt? Er moet een *integraal* beeld van elke opleiding ontstaan. Dat beeld betreft een hoofdbeeld en niet een optelsom van alle afzonderlijk beleid in de instelling. De commissie is bij voorbeeld evenmin geïnteresseerd in het beleid ten aanzien van ziekteverzuim, de medezeggenschap, het energiebeleid en het onderhoudsbeleid aan gebouwen als in het facultaire personeelsbeleid. De visitatie is geen evaluatie van het instellingsbeleid en evenmin van het faculteitsbeleid.

De opleiding, dus het primaire proces, staat wel in het middelpunt. Waar gaat de concrete interesse per opleiding naar uit? We geven een algemeen beeld zoals dat blijkt uit gesprekken met de decanen aan enkele universiteiten op het gebied van rechtsgeleerdheid, bestuurskunde, economie en bedrijfskunde en blijkt uit de rapporten van de visitatiecommissies. Een visitatiecommissie keek destijds vanuit een *integraal kader* naar onder meer de doelen van de opleiding en eindtermen; de programmatische opbouw met vakken; de ondersteuning en fasering van de opleiding; de intake, doorstroming, de uitval en het rendement van de opleiding; het mentoraat en de studieadvisering; de kwaliteit van examinering; de kwaliteit van in een opleiding centrale afstudeerwerkstukken of scripties; de bibliotheek; waar de afgestudeerden (alumni) terecht komen na het afstuderen en wat hun oordeel is over de opleiding; de kwaliteit van de staf; de kwaliteit van de infrastructuur; de kwaliteit van de hulpmiddelen voor studenten; de informatieverstrekking aan studenten over de opleiding en onderdelen daarvan (stagereregeling enz.); de nationale of internationale dimensie van de opleiding (uitwisseling?); de organisatie en het management van de opleiding. Dit zijn eigenlijk allemaal voor de hand liggende onderwerpen die samen een integraal beeld geven.

De visitatiecommissie gaat vanuit dit protocol, een zelfevaluatie en een groot aantal documenten van de instelling zich ter plaatse, gedurende een periode in de orde van grootte van drie dagdelen, oriënteren. Ze spreekt met de decaan, de docenten en de verschillende

commissies, studenten uit verschillende leerjaren en afgestudeerden en eventueel met een beroepenveldcommissie of ander adviesorgaan en bekijkt in een strak georganiseerd programma de onderwijsprogrammering, de vakinhoud, de scripties, de onderwijsorganisatie en ondersteuning, de bibliotheek of andere mediatoepassingen. De commissie bezint zich op de eerste bevindingen en geeft bij monde van de voorzitter en onder voorbehoud een eerste terugkoppeling naar de betrokkenen.

De ambtelijke secretaris geeft vervolgens een voorzet voor een concepteindrapport per instelling, dat het oordeel van de visitatiecommissie weergeeft, en de commissie legt dat voor commentaar voor aan de decaan van de gevisiteerde instelling. Er kunnen correcties op als onjuist beschouwde feiten gegeven worden. Daarna stelt de commissie het visitatierapport vast. De commissie maakt ook een tweede rapport, dat is een landelijk overalrapport met algemene bevindingen over het wetenschappelijk onderwijs in de discipline.

Tijdens de eerste visitaties waren visitatiecommissies zeker voldoende concluderend maar nog enigszins terughoudend in het doen van aanbevelingen of suggesties. De richting van de verbeteringsnoodzaak zoals de commissie die aangaf, was doorgaans overigens wel duidelijk.

Vervolgens zijn de instellingen aan bod om dat wat verbeterbaar is aan te pakken. Meestal is de decaan dan als eerste aan zet. Dat gebeurt onder toezien oog van het eigen college van bestuur en de inspectie hoger onderwijs. Die inspectie beoordeelt de visitatie op deugdelijkheid en kijkt vervolgens ook of er consequenties verbonden moeten worden aan de visitatie. De deugdelijk is meestal al gegarandeerd als de zelfevaluatie aan de richtlijnen voldoen want op de juistheid, correcte invulling en volledigheid wordt toegezien.

Inspectie

De inspectie oordeelt of de visitatie kwalitatief voldoet. Opleidingen die op veel fronten niet voldoen (ernstige tekortkomingen vertonen), kunnen op basis van het visitatierapport en het vervolg daarop met opheffing bedreigd worden. Dit heet *'license to kill'*; de inspectie kan dr. No zijn (Scheele, 2003:67). De inspectie voert daarover, maar ook over de gewenste verbetering van opleidingen die wel 'boven de streep' zijn, een aantal maanden na het uitkomen van het visitatierapport, overleg met het college van bestuur. Er is dus zeker sprake van een visitatiestelsel, waarbij veel van onderop kan geschieden maar een inspectie is in beeld om *de meta-evaluatie* te verrichten (de evaluatie van de visitatie, die ook een evaluatie is) op de doorwerking van de bevindingen toe te zien. De inspectie werkt in deze op basis van vertrouwen en let op de overtuigingskracht van de reactie van het college van bestuur. Daarbij heeft de inspectie het voordeel dat als een verbetering uitblijft de volgende visitatie een evident zwak punt weer in beeld kan krijgen. Bij visitatie in het wetenschappelijk onderwijs is er dus druk op de ketel om te verbeteren.

Dynamiek

Rond de externe kwaliteitszorg in het w.o. heeft zich enige dynamiek voorgedaan. Die laat zich vatten in een figuur (Meijerink, 1999).

Figuur 1: *Ontwikkeling in kwaliteitszorg op gebied van onderwijs v.a. 1985 en enkele andere markante momenten*

1985	HOAK-nota: Hoger Onderwijs - Autonomie en Kwaliteit.
1987	Start eerste ronde onderwijsvisitaties van w.o.-opleidingen.
1988	Hoger Onderwijs en Onderzoek Plan 1988 (HOOP 1988) (ook latere)
1992	Start eerste ronde erkenning onderzoeksscholen.
1993	Start door VSNU georganiseerde kwaliteitszorg wetenschappelijk onderzoek.
1994	Start tweede ronde onderwijsvisitaties waarbij onder meer is vastgelegd dat in visitatiecommissies tenminste een student moet zitten.
1994	Uitbreiding van rol inspectie; behalve meta-evaluatie ook beoordeling van de

	bestuurlijke hantering van aanbevelingen uit visitatierapporten.
1998	Afronding herordeningsoperatie opleidingsaanbod: sanering door universiteiten zelf van de in het CROHO ingeschreven opleidingen
1999	Minister tekent de Bolognaverklaring, waarin voornemen wordt geuit naar een hoger onderwijs in twee fasen over te gaan.
2000	Start derde ronde onderwijsvisitaties en tweede ronde onderzoeksvisitaties. In beide systemen wordt de beoordeling aangescherpt en de internationale dimensie versterkt.
2000	Invoering Prestatie Bekostigings Model: budgetverdeling gaat in kader van financiële verhoudingen met universiteiten vooral afhangen van de resultaten, te weten aantallen diploma's e.d..
Rond 2003	Start accreditering van opleidingen.

Samenvatting

Het voorgaande kan als volgt kort samengevat worden.

Visitatie in het w.o. is opgenomen in een stelsel van visitaties. Het doel van de visitatie van opleidingen was in de periode van onderzoek 1985-2003 om te verantwoorden en te verbeteren. Visiteren impliceert in het w.o. altijd een legitimatie van de betrokken 'partners', zoals van het ministerie, de instellingen en inspectie; een initiatiefnemer die de volgorde bepaalt en aftrapt; een visitatiecommissie die is samengesteld na consultatie van het onderwijsveld en die onafhankelijk is; een protocol (opzet e.d.); een zelfevaluatie of zelfstudie; een site visit; een concepteindrappport per instelling; de mogelijkheid tot commentaar op een conceptrapport; een eindrapport voor de instelling; een eindrapport over alle gevisiteerde opleidingen van een soort heen; openbaarheid van het rapport; de inspectie die een visitatieproces moet goedkeuren.

6 De kwaliteit van visitaties in het wetenschappelijk onderwijs

We komen nu bij de kern van deze analyse: wat is de kwaliteit van visitaties in het wetenschappelijk onderwijs in de periode 1985-2003, vóór de introductie van de accreditatie?

6.1 Beoordelingscriteria

De eerste vraag op weg naar beantwoording is dan: wat zijn beoordelingscriteria om kwaliteit van de visitaties vast te stellen? In feite is al een aantal criteria genoemd bij de weergave van de positieve en negatieve kijk op visiteren in het w.o. Maar om zicht te krijgen op de relevante criteria is uiteraard meer literatuur geraadpleegd. Daartoe behoren overheidsdocumenten over de doelstellingen achter de visitatie (wat wil de overheid ermee bereiken?) en gangbare criteria die door onderwijskundigen en anderen beschouwd worden als relevante aspecten van de kwaliteit van externe kwaliteitszorg (o.a. Van Vught & Westerheijden, 1994; Algemene Rekenkamer, 2000; Douma, 2004).

Doelen die achter visiteren in het w.o. schuil gaan, betreffen verantwoording en leren. Wie zich wil verantwoorden moet dat voor *het forum* van vakgenoten doen en in *openbaarheid*. Openbare gegevens zijn niet genoeg want de bevindingen moeten ook de onderwijskwaliteit *integraal* in beeld brengen (en dus niet slechts enkele elementen of aspecten). Selectiviteit is dus gewenst. Bij visiteren van een organisatie of opleiding zijn niet alle gegevens relevant maar vindt selectie op grond van een integraliteitsfocus, een 'governance'- kijk of perspectief op het te visiteren object, plaats.

De visitatie moet ook een voor derden begrijpelijk beeld opleveren

- dat *geldig* is (daadwerkelijke meting van onderwijskwaliteit; geen vertekening door strategisch gedrag van onderwijsgeevenden?; De Bruijn en Ten Heuvelhof, 2005; De Bruijn, 2008),

- dat *betrouwbaar* is (bij herhaling dezelfde uitkomsten zou moeten opleveren),
- en *bruikbaar* is voor benutting door de inspectie (die een kwaliteitsoordeel over procedure en rapport uitspreekt) en door de onderwijsgeevenden van de instellingen zelf.

We willen ook de *procedurele kwaliteit* kennen: dragen genoemde 'beginselen van behoorlijk visiteren' wel of niet krachtig bij aan de visitatiekwaliteit, in casu geldigheid, betrouwbaarheid en bruikbaarheid? (Korsten, 2004a)

Daarmee is niet te volstaan. Immers, er zijn positieve argumenten pro visitatie en negatieve argumenten genoemd. We zullen dus ook moeten letten op de (non)acceptatie van het kwaliteitszorgsysteem als geheel (de vraag of de instellingen van harte meewerken aan de visitatieprocedure of dit als een loden last dragen); op de samenstelling en werkwijze van de commissie; op de aan- of afwezigheid van meerwaarde (of uitkomsten meerwaarde hebben of slechts 'commen sense' bieden?); op de kosten-batenbalans (meer lasten dan baten of omgekeerd?). Uiteraard ontkomen we niet aan aandacht voor mogelijke onbedoelde (on)gewenste effecten van visiteren.

Het wetenschappelijk onderwijs moet geplaatst worden in de context van een bestuurlijk meerlagensysteem ('multi-level governance') waarin een rol wordt vervuld door de minister (die onder meer voor bekostiging zorg draagt), de (toezichhoudende) inspectie, de vereniging van samenwerkende universiteiten (VSNU) die het visitatiestelsel coördineert, de instelling en opleidingsdecanen. We ontkomen er daarom niet aan om te kijken naar de beoordeling van de visitatie door de verschillende betrokkenen. Hier en daar zal gespecificeerd worden wat het oordeel *van wie* is.

In het navolgende volgt eerst een hoofdbeeld van de bevindingen over de gehele periode 1985-2003 waarna de focus gericht wordt op elk van de drie rondes van visiteren in het wetenschappelijk onderwijs.

6.2 Hoofdbeeld van de kwaliteit van visiteren in drie rondes

De kwaliteit van de visitaties in het w.o. kan - zoals gesteld - onder meer bepaald worden aan de hand van de betrouwbaarheid en de validiteit van de visitaties, menen ook Hulshof e.a. (2003: 24). Hulshof was toen ze dit als een van de auteurs schreef coördinerend inspecteur hoger onderwijs bij de Inspectie van het Onderwijs.

Validiteit wil zeggen: meet de visitatiecommissie de goede dingen? Hier gaat het om geldigheid van wat wordt vastgesteld. Heeft het oordeel van de visitatiecommissie ook echt betrekking op de onderwijskwaliteit? En betrouwbaarheid houdt in: meet de commissie 'de goede dingen' zo dat bij herhaling hetzelfde blijkt? Zou opnieuw een onafhankelijke visitatiecommissie gaan visiteren op basis van dezelfde zelfevaluatie en dezelfde gesprekspartners dan zou een zelfde kwaliteitsoordeel uitgesproken moeten worden om van betrouwbaarheid te kunnen spreken. Herhalingsonderzoek om de betrouwbaarheid te bevestigen en analyse via alternatieve data die iets zeggen over geldigheid bestaat echter niet. Het vaststellen van betrouwbaarheid en geldigheid moet dus op een andere manier plaats vinden. De toets van de inspectie kan ons hier als 'next best'-middel dienen bewijzen.

Om maar direct met de deur in huis te vallen: gezien die inspectieoordelen kunnen we vaststellen dat visiteren in de periode 1985-2003 deugt. Immers, voor zover bekend zijn visitatieprocedures nooit uit onvrede afgebroken en zijn visitatierapporten nooit door de onderwijsinspectie, die onder andere tot taak heeft om visitaties te beoordelen, vanwege gebrekkige kwaliteit naar de prullenmand verwezen (Laman, 2003: 45). Dat is bij nader inzien logisch. Om de validiteit en betrouwbaarheid te borgen, zijn al in de visitatieprocedure in het

w.o. diverse maatregelen genomen die geen ontsnappen toelaten en die ertoe dwingen om te komen tot betrouwbare uitspraken en tot geldige uitspraken, althans in intersubjectieve zin.

Wat die maatregelen zijn? Ze passeerden al de revue maar we noemen nog even een aantal argumenten (verg. Van Vught en Westerheijden, 1994).

Ten eerste, de colleges van bestuur hebben via hun orgaan, de VSNU, een visitatieprotocol opgesteld. Ze hebben dus gekozen voor zekere standaardisatie waaraan de colleges zich individueel en collectief binden.

Ten tweede, voorzien is dat de visitatie wordt uitgevoerd door gezaghebbende professionals op het vakgebied. De disciplineoverlegorganen konden zelf meedenken over wie geschikt is om daarin zitting te nemen. Deze inbreng is niet altijd een op een gehonoreerd (beoordelaars moeten ook beschikbaar zijn en tegen de vergoeding willen werken) maar de invloed was zodanig dat disciplineoverlegorganen zich doorgaans in de samenstelling konden vinden. Ten derde is er de zelfevaluatie waarvan de relevante vragen en de inhoudsopgave vooraf gegeven zijn. Betrokkenen beantwoorden de vragen met het volle verstand en de decaan, opleidingscoördinatoren en medewerkers binden zichzelf hieraan. De visitatiecommissie toetst die evaluatie op volledigheid en aanwezigheid van de relevante feiten alvorens de inhoud tot zich te nemen.

Ten vierde mag de instelling commentaar leveren op het conceptrapport van de visitatiecommissie en dus vaststellen of gegevens of inzichten voor de oordeelsvorming van de commissie ontbreekt. Deze gegevens moeten dan alsnog worden geleverd.

Deze argumenten maken dat het negeren van een visitatie door de onderwijsinstelling en de inspectie nagenoeg onmogelijk is en ook niet plaatsvond.

Heeft alleen de inspectie zich in algemene termen uitgesproken? Ook de Algemene Rekenkamer kwam met een analyse. De Rekenkamer toont zich in 2000 positief over de visitatiepraktijken. De VSNU op haar beurt was verheugd over die conclusie van de Rekenkamer 'dat de externe kwaliteitszorg in grote lijnen goed functioneert'. En de minister van Onderwijs, Cultuur en Wetenschappen uit het kabinet-Kok II onderschreef in dat jaar de conclusies van de Rekenkamer (AR, 2000: 6). De minister stelt in het 'Hoger Onderwijs en Onderzoek Plan 2000': *'Het Nederlandse stelsel van kwaliteitszorg staat internationaal hoog in aanzien'* (TK 26 807, nr. 19: 2). Daarmee is een deel van de negatieve kritiek op visiteren in het wetenschappelijk onderwijs impliciet van tafel geveegd, maar we moeten toch nog preciezer kijken, want oordelen van ministers moeten geconfronteerd worden met machtsvrije analyses.

Wordt de veronderstelde 'goede praktijk' bij een analyse van de opeenvolgende ronden (of cycli) van visiteren van w.o.-opleidingen bevestigd of genuanceerd? We zetten enkele gegevens op een rij, die hoofdzakelijk ontleend zijn aan jaarverslagen van de inspectie alsmede aan analyses van onder meer Frederiks (1996), Mertens (1999); Algemene Rekenkamer (2000); Leeuw (2001); Goossens, Laman & Fröling (2001); Hulshof e.a., (2003), Laman, (2003), Jeliaskova en Westerheijden (2000, 2004). Het gaat om een algemeen oordeel per visitatieronde en daarmee om oordelen over geldigheid en betrouwbaarheid, de samenstelling van de commissie, de onafhankelijkheid, de werkwijze, de kwaliteit van de zelfstudies en de eindrapporten en andere eerder, in het begin van deze paragraaf, genoemde onderwerpen. Er komt ook ruimte voor nuancering en kritiek omdat we bezien of er grond is voor het eerder genoemde kritisch beeld over visiteren in het w.o.. Verderop komen ook de 'beginselen van behoorlijk visiteren', die al aangekondigd zijn, aan bod.

6.3 Nadere analyse van visitatiekwaliteit in de eerste ronde: 1988-1993

Beoordeling 1: Het algemeen oordeel over de eerste visitatieperiode 1988-1993

Een onderzoeker van het - in Enschede gevestigde en aan de Universiteit Twente gelieerde - onderzoekscentrum CHEPS concludeert in 1996 dat er onder opleidingen, faculteiten en instellingsbestuurders een relatief grote *tevredenheid* over het visitatiestelsel bestond. De *acceptatiegraad* van het visitatiebestel is dus meer dan voldoende (Frederiks, 1996: 215). De stelling dat bestuurders van onderwijsinstellingen, onderwijsdecanen en andere betrokkenen vol chagrijn of met gepast cynisme aan de luxe 'feedback- en toezichtindustrie' meegewerkt hebben, moet verworpen worden. Ook instellingen beseffen dat als er kwaliteitszorg bestaat in alle sectoren van het onderwijs universiteiten zich niet aan kwaliteitszorg kunnen en moeten onttrekken. De instellingsbesturen hebben in feite alle kritiek op het visitatiestelsel vooraf verworpen door zelf proactief mee te bouwen aan het stelsel.

Daar bleef het niet bij. De onderwijsinspectie publiceerde over de eerste visitatieronde in 1994 het rapport '*Kwaliteitszorg wetenschappelijk onderwijs*'. De visie van de inspectie over de eerste ronde werd gekenmerkt door '*tevredenheid, ja hier en daar zelfs ingehouden enthousiasme*' aldus Laman (2003: 38), senior onderzoeker hoger onderwijs bij de Inspectie van het Onderwijs. Twee positieve kwalificaties dus.

Beoordeling 2: Strategisch gedrag niet afwezig maar moeilijk aantoonbaar

De inspectie voelt zich wel op afstand staan. Voorzichtig wordt gewag gemaakt van mogelijk strategisch gedrag van instellingen, een argument dat we nog niet onder de negatieve kritiek scharden maar wel van belang is. Instellingen kunnen er belang bij hebben om een niet in alle opzichten volledig waarheidsgetrouw beeld te geven van de onderwijsrealiteit, aldus het inspectierapport. Waaraan dan te denken? We geven enkele zelf onderkende mogelijkheden tot strategisch gedrag en ook mogelijkheden, die de visitatie kunnen detoneren (De Bruijn en Ten Heuvelhof, 2005; De Bruijn, 2008; Schillemans, 2009).

Specifiek in relatie tot visitatie in het w.o. is het niet overdreven te spreken van een *hoera-effect*: verantwoordelijken voor opleidingen willen in het algemeen gunstig uit visitaties te voorschijn komen. Een gunstige uitkomst geeft immers weinig interne discussie en verbeternoodzaak en dus weinig werk, en biedt voor een onderwijsdecaan of opleidingscoördinator wellicht de mogelijkheid de aandacht naar andere thema's te verleggen. Een opsteller van een zelfstudie kan daarom het benoemen van zwakheden van de opleiding in een zelfstudie zoveel mogelijk trachten te vermijden. De vrees kan immers bestaan dat het tonen van kwetsbaarheid in de zelfstudie in de rapportage van de visitatiecommissie misschien zelfs uitvergroet terugkeert. De commissie moet tenslotte ook met bevindingen komen. Verantwoordelijken uit de gevisiteerde praktijk zullen zwakheden dus 'klein' willen houden of onvermeld laten. Zie daar het strategisch gedrag dat de geldigheid en betrouwbaarheid van de visitatie kan aantasten omdat de zelfstudie een belangrijke bron is voor de commissie. Daar staat tegenover dat visitatiecommissieleden hiermee bekend zijn. Als ze ook maar een begin van 'strategische handigheden' bespeuren, zullen ze - ook al is de tijd voor 'site visits' beperkt - nog strijdbaar op zoek gaan naar de waarheid, zo vertelden enkele leden van deze voormalige commissies ons.

Een andere reactie die een opsteller kan uitvoeren, is om te komen met zo veel plannen, nota's, cijfers, evaluaties, voortgangsrapportages dat de oren van de commissieleden tuiten. We noemen dit het *evaluatiestroom-effect*: visiteren roept het verschijnsel op om werkelijk alles tevoren te evalueren. Als de commissie vraagt naar informatie kan de opleidingsdecaan over bijna alles wel een document overleggen. Zo kunnen verantwoordelijken voor opleidingen trachten kritiek te smoren of te laten zien dat 'dit item' al vóór de visitatie is gesignaleerd en opgepakt om zo absorptievermogen van kritiek te etaleren. De inschatting kan bovendien zijn dat door het verstrekken van veel informatie de kans toeneemt dat de ordening en duiding van de opleidingsdecaan door de commissie gevolgd wordt. Maar een visitatiecommissie kan dit

ook weten en daarom juist van betrokkenen verwachten dat ze er niet omheen draaien. Een visitatiecommissie kan argwaan tonen en het gevecht om de waarheid (blijven) aangaan. Er is bij strategisch gedrag dus altijd een tegenkant: de commissie wil zich niet in de luren laten leggen.

Een derde vorm van strategisch gedrag die op het eerste gezicht onschuldig oogt, is de benadrukking van lokale kennis en omstandigheden (*het lokale omstandigheden-effect*). De commissie heeft moeite om tot een inhoudelijk oordeel te komen over tal van aspecten, juist vanwege de specifieke situatie, waarmee de commissie niet of nauwelijks vertrouwd is. Benadruk het specifieke omdat dit een verontschuldiging kan zijn waarom de opleiding toch net even afwijkt van andere opleidingen die de commissie bekijkt (De Bruijn, 2008: 103).

Er is nog een vierde vorm van strategisch gedrag: *het bestuurderseffect*. Dat is een variant op het omstandigheden-effect. De opleidingsdecaan benadrukt dan dat gedrag van bestuurders van invloed is geweest op de onderwijsprestaties, zoals de wisseling in een college van bestuur of faculteitsbestuur. Het is ook mogelijk dat een verandering in ministeriële kaders op onderwijsterrein is opgetreden.

Een vijfde vorm van strategisch gedrag houdt in dat een *strategische visie* wordt gepresenteerd. Die visie leent zich doorgaans voor toelichting en discussie waardoor de ogen afgehouden worden van de in de zelfstudie opgenomen prestatiegegevens, het primaire issue voor de commissie. De commissie krijgt op een strategisch document doorgaans minder goed grip waardoor de discussie naar plannen voor de toekomst verlegd worden. De redenering die een strategisch opererende decaan bij de commissie ingang wil doen vinden is: het heden moet met een korrel zout genomen worden, want de toekomst zal een ander heden brengen. Zo wordt de opleiding minder kwetsbaar, minder kritiseerbaar gemaakt (De Bruijn, 2008: 103).

Een zesde vorm van strategisch gedrag betreft het *meebuigeffect*. Een opleidingsdecaan weet dat de commissie altijd met enige kritiek wenst te komen en daarom komt hij ook daadwerkelijk zelf met kritiek; en wel met bepaalde kritiek. De decaan probeert de commissie voor te zijn. Degene die strategisch gedrag vertoont, benoemt zelf de kritische punten die hij benadrukt wil zien: omdat ze geen kwaad kunnen en/ of ze hem nauwelijks aangewreven kunnen worden. Zoals bijvoorbeeld de staat van de gebouwen. Dat is immers een zaak van het college van bestuur en niet van de onderwijsprofessionals. Opname van een dergelijke punt is vanuit een gereserveerde decaan gezien gunstig.

Een andere mogelijkheid is dat een decaan anders meebuigt, namelijk door in het beraad met de commissie te komen met '*dilemma sharing*' of een *cascade aan toezeggingen*. Ook hier gaat het om meebuigen. Dilemma sharing verwijst naar het voorleggen van dilemma's aan de commissie. De decaan kan ook concessies doen en zeggen dat hij 'de zaken al aan het oppakken was en is'. De gevisiteerde gaat er hierbij vanuit dat de commissie ergens halverwege al tevreden is en ziet dat de decaan ontvankelijk is voor argumenten en suggesties. De decaan kan erop speculeren dat dit een vertrouwenwekkende houding oplevert bij de commissie zodat die niet doorzet om alle kritische punten te noemen (De Bruijn en Ten Heuvelhof, 2005: 43).

Een zevende vorm van strategisch gedrag is het *verantwoordelijkheidseffect*. Dat houdt in dat de opleidingsrepresentanten de commissieleden huiverig maken voor het geven van een hard oordeel over de opleiding. Door te benadrukken dat een zwaar kritisch rapport de opleiding lokaal van de kaart kan vegen, accentueren deze opleidingsvertegenwoordigers welke zware taak op de schouders van de commissie rust. De commissie kan terugdeinzen voor een eventueel groot effect van haar rapport op een aantal opleidingen in den lande. De onderwijsinspectie onderkent dat een visitatiecommissie op haar beurt ook strategisch kan

opereren door het vakgebied, waartoe de leden zelf behoren, in bescherming te nemen (Laman, 2003). Een inschatting maken van de reacties en effecten op het rapport, aannemend dat persmedia graag een kleinigheid uitvergroten, kan een commissie milder stemmen dan ze in werkelijkheid is. Zoals commissieleden wel zeiden: 'Je hoeft niet alles op te schrijven, je kunt ook discreet enkele meningen achterlaten bij de opleidingsdecaan als je denkt dat dit verstandig is'.

Tot zover enkele mogelijke reacties die naar voren komen wanneer visiteren vanuit een (strategische) spelmetafoor gezien wordt. De ruimte voor strategisch gedrag is in zelfstudies niet groot omdat voor de inhoud daarvan regels gelden. Bij 'site visits' lijkt daarentegen meer ruimte voor strategisch gedrag aanwezig te zijn, tenzij de commissie zich strak houdt aan het programma en de vragen die ze op basis van de zelfstudie beantwoord wil zien.

Hoe begrijpelijk het ook is om strategisch gedrag te veronderstellen, onderzoek naar de visitatiekwaliteit toont over het algemeen geen strategisch gedrag aan.

Filteringseffect

Reageert de inspectie zoals ze ook doet in andere sectoren van het onderwijs waar ze van oudsher juist zelf in de onderwijskeuken gaat kijken? Kan de inspectie niet loslaten?

De zelfstudies vormen een belangrijke kern van het visitatiebestel en daar heeft de overheid geen greep op, en de onafhankelijke inspectie ook niet; evenmin op het visitatieproces zelf en op de 'site visit'. We noemen dit verschijnsel *het filteringseffect*: door de visitatieprocedure van zelfstudie en site visit kan een beeld van de opleidingswerkelijkheid bij de inspectie ontstaan dat tweemaal gefilterd is, aldus Laman (2003: 39). Anders gesteld: de inspectie krijgt een beeld dat gebaseerd is op wat zichtbaar gemaakt is in de zelfstudie en in het commissierapport. Is dat het hele verhaal of is sprake van een *façade* waarachter ook wel dingen schuilgaan die niet gezien worden of niet gezien mogen worden? Het zou kunnen maar er is geen onderzoek dat daarop wijst.

Conclusie: positieve hoofdbevinding blijft overeind

Ondanks deze vragen en hier benoemde effecten (zoals onder meer het hoera-effect; evaluatiestroomeffect; het lokale omstandighedeneffect; het meebuigeffect; het verantwoordelijkheidseffect; het filteringseffect) weerstaat de inspectie de aandrang om zelf achter de schermen, de façades, te kijken. Immers, achter de HOAK-nota ging de wens tot sturen op afstand schuil en dus moet de inspectie terughoudend zijn.

De inspectie gaat er in beginsel vanuit dat er '*een redelijk getrouw beeld van de werkelijkheid wordt verkregen en kwaliteit en kwaliteitsverbetering zichtbaar kunnen worden*', aldus Laman (2003: 39). Het positief beeld staat dus nog steeds. Maar laten we verder kijken naar enkele specifieke aspecten van de externe kwaliteitszorg, zoals de kwaliteit van de commissies en werkwijzen.

Beoordeling 3: De samenstelling, onafhankelijkheid en werkwijze van de commissie

De commissiesamenstelling is positief beoordeeld. Slechts in een op de vijf visitaties uit de eerste ronde van visitaties (1994-1999) had de inspectie aanmerkingen op de samenstelling van de visitatiecommissie en in vier op elke vijf, de overgrote meerderheid, dus niet.

Aan de onafhankelijkheid van de leden werd zelden getwijfeld (Goossens, Laman & Fröling, 2001; Hulshof e.a., 2003: 27).

En de werkwijze dan? De commissies moesten de protocollen hanteren die vooraf zijn vastgesteld en geaccepteerd zijn in disciplineoverlegorganen van de VSNU. De commissies

hebben door vast te houden aan de protocollen over het algemeen dan ook een adequate werkwijze aan de dag gelegd.

In de praktijk hadden visitatiecommissies zekere speelruimte en die benutten ze ook. Wie vervolgens de blik richt op visitaties van geheel verschillende studierichtingen constateert wel enige (beperkte) verschillen tussen grofweg een visitatie van de studie Letteren versus die van Aardwetenschap, Maritieme Techniek, Tandheelkunde of Sociologie. Die verschillen hebben met name betrekking op de signalering van tekortkomingen in het onderwijs (Hulshof e.a., 2003). Neem de toetsing van onderwijsprestaties. Van inspectiezijde wordt opgemerkt dat er bijvoorbeeld niet steeds sprake is van een consistentie tussen bevindingen en oordelen van de visitatiecommissies over toetsing. Blijkbaar zijn sommige commissies in hun bevindingen (feiten, kwalificaties door betrokkenen) kritisch maar milder in hun oordeel of omgekeerd: kritisch in het oordeel, terwijl de bevindingen daarvoor wat 'dun' zijn en de onderbouwing dus niet steeds sterk is (Hulshof e.a., 2003: 30). Dat is dus een kritisch punt, tegen de achtergrond van het eerdere gememoreerde positieve hoofdbeeld.

Beoordeling 4: De kwaliteit van de zelfevaluatie

We komen bij een vierde criterium voor beoordeling: de kwaliteit van de zelfstudie. Ook al zijn de zelfstudies of zelfevaluaties een primaire verantwoordelijkheid van de instellingen zelf, ze zijn wel van belang omdat ze een gewichtige bron zijn voor de commissie. Ook op de zelfstudies van w.o.-opleidingen heeft de inspectie over het algemeen weinig tot geen kritiek (Hulshof e.a., 2003). Logisch ook want de visitatiecommissie toetste het rapport al vóór het definitief inleveren ervan, zij het op volledigheid en niet zozeer inhoudelijk. Opnieuw dus een positief punt.

Beoordeling 5: De kwaliteit van de visitatierapporten, eerste ronde

De inspectie had achteraf zelden of nooit aanmerkingen op de kwaliteit van het visitatierapport volgens een onderzoek dat gebaseerd is op veertien visitaties uit de periode 1988-1993. Een visitatierapport is nooit afgekeurd, volgens de onderzoekers (Hulshof e.a., 2003: 30). *'Er wordt vertrouwd op de intersubjectiviteit en de consensus binnen de visitatiecommissie'*, aldus Hulshof e.a. (2003: 30).

De conclusie op basis van het voorgaande luidt: de eerste ronde van visiteren laat over het algemeen een redelijk positief beeld van de kwaliteit van visitatieprocessen in het wetenschappelijk onderwijs zien. De uitkomsten van visiteren worden grosso modo beschouwd als geldig (de goede dingen zijn gemeten), betrouwbaar (de uitkomst is stabiel; bij herhaling van de procedure zou dezelfde uitkomst blijken) en bruikbaar voor verbeteracties. De bevindingen van de visitatiecommissie worden dus herkend ('dit is de goede foto'; 'dit zijn wij') en erkend ('geaccepteerd').

Maar dit visitatieproces is qua zoeken naar geldigheid en betrouwbaarheid niet vergelijkbaar met zuiver-wetenschappelijk onderzoek dat gericht is op het toetsen van hypothesen en op precieze uitkomsten om kennis verder te brengen. Visitaties van opleidingen zijn gericht op feitenverzameling en duiding daarvan; het gaat om - in termen van James Coleman - op besluitvorming gericht onderzoek. Ze kennen daarmee onvermijdelijk ook elementen van intersubjectiviteit (aan de kant van opleiders en commissieleden). Visitaties zijn gericht op beoordeling en besluitvorming over eventuele zwakten en verbeterpunten in opleidingen.

6.4 Nadere analyse van visitatiekwaliteit in de tweede ronde: 1994-1999

Wat toont de tweede ronde, dus de tweede cyclus van visiteren in het wetenschappelijk onderwijs, waarbij alle w.o.-opleidingen dus opnieuw tegen het licht zijn gehouden? De inspectie heeft zich na de eerste ronde opnieuw over de kwaliteit van de visitatierapporten

gebogen. De eerste visitatieronde betrof 1988-1993 en de tweede ronde voltrok zich in 1994-1999.

Beoordeling 6: Draagvlak voor visitatie?

In de eerste ronde is volop draagvlak gebleken voor de visitatie en dat blijkt ook in de tweede ronde. De inspectie stelt ook vast dat er 'een goed draagvlak bestaat voor het visitatiesysteem' (Laman, 2003: 42). Deze steun bleek in feite in alle drie de visitatierondes, te bestaan bij het coördinerend agentschap (VSNU), de instellingen, de inspectie, het ministerie. Weer een positief beeld van een kritische inspectie.

Beoordeling 7: Kwaliteit van rapporten in tweede ronde

De visitatierapporten zijn in de eerste ronde over het algemeen positief uit de bus gekomen. Laman (2003: 41) meldt: 'Met betrekking tot de visitatierapporten die in het laatste jaar van de tweede cyclus (1997) zijn verschenen, is de inspectie van mening dat de kwaliteit daarvan zeer bevredigend is'. Opnieuw een positief beeld, dus.

Beoordeling 8: Benutting van visitatierapporten en 'leren'

Blijkt dat ook te gelden voor een ander criterium: is er door opleidingscoördinatoren en decanen rekening gehouden met conclusies en aanbevelingen uit de visitatie van de opleidingen in de eerste ronde? Op dat aspect is in het voorgaande nog niet ingegaan maar dat moet wel want visiteren was meer dan verantwoorden; visiteren in het w.o. pretendeert ook een bijdrage te leveren aan 'leren' en dus aan verbeteren (Jeliazkova, 2002; Laman, 2003). Wat blijkt dan?

De Algemene Rekenkamer stelt vast dat de verbeterfunctie van het visitatiestelsel 'goed tot haar recht komt' (AR, 2000: 4). En: 'In de onderzochte gevallen bleken de aanbevelingen uit de visitatierapporten aantoonbaar en in voldoende mate door te werken in het onderwijskundig proces' (2000: 4). De onderwijsinspectie stelt in het kielzog van de Algemene Rekenkamer vast dat 'over het geheel genomen de instellingen op adequate wijze zijn omgegaan met de aanbevelingen van de visitatiecommissies' (Laman, 2003: 42).

Wie onderneemt het meest initiatief tot verbetering, wie het minst? Onderwijsinstellingen waar het kwaliteitszorgsysteem intern al behoorlijk ontwikkeld is, blijken zich minder door rapporten van een visitatiecommissie aangesproken te voelen. Dat komt omdat ze op veel fronten een 'voldoende' scoren. En wie voldoende scoort, ervaart een visitatierapport blijkbaar niet als een krachtige prikkel tot verdere verbetering. Wie 'onvoldoendes' scoort, is eerder geneigd tot verbeteracties.

Komen decanen van verschillende universiteiten gezamenlijk tot verbeterinitiatieven of niet? We stoten hier op het 'eigen instelling-eerst-effect'. Gezien vanuit het gehele universitair bestel is elke instelling sterk op zichzelf gericht; ze stelt het eigen onderwijs centraal en *niet* het onderwijs in het hele vakgebied in alle relevante instellingen. Instellingen zoeken over het algemeen zelf oplossingen in termen van verbeteringen. De neiging is gering om te komen tot landelijke gecoördineerde reacties en verbeteracties (Laman, 2003: 42). Logisch eigenlijk? Inderdaad. Waarom zou een instelling vrijwillig zeggenschap uit handen geven en waarom zou een verantwoordelijke decaan zich ondergeschikt maken aan een oordeel of voorstel van de meerderheid van onderwijsdecanen uit zijn vakgebied? Maar decanen overlegden toch al in een door de VSNU ondersteund landelijk disciplineoverlegorgaan? Zeker, maar ervaringen daarin van decanen in de sector bestuurswetenschappen, bedrijfskunde en economie in de periode 1986-2003 wijzen uit dat er wel sprake is geweest van bezinning en uitwisseling maar dat een dergelijk orgaan geen bevoegdheden van onderop krijgt van afzonderlijke instellingen. Disciplineoverlegorganen nemen over het algemeen dus geen besluiten die instellingen

binden. Besluiten worden genomen daar waar de bevoegdheden en de verantwoordelijkheden liggen; bij de universiteiten zelf: bij het college, bij de faculteitsbesturen en decanen. Dat wil niet zeggen dat de overlegorganen betekenisloos waren; dat zeker niet. Er vond informatieverstrekking en -uitwisseling plaats en de leden gingen ook wel persoonlijke 'commitments' aan maar ze namen die mee naar de eigen instelling om daar in te brengen.

Er valt nog meer te zeggen over 'knowledge utilization'. Om daarvoor iets te zeggen, moeten we eerst terug naar de aanbevelingen.

Uit analyse van visitatierapporten blijkt dat er in grote trekken twee soorten aanbevelingen zijn aan te treffen (Westerheijden, 1997; Scheele, Maassen en Westerheijden, 1998). De eerste groep betreft *specifieke problemen* die zich laten plaatsen *binnen* een door de visitatiecommissie en onderwijsstaf van een universiteit gedeelde visie op de onderwijsprogrammadoelstellingen. De tweede groep van aanbevelingen plaatst het onderwijsprogramma in een *breder kader* van een discipline, universiteit en/of samenleving. Dan komen de programmadoelstellingen in het vizier en soms ter discussie, bij voorbeeld omdat een pleidooi wordt gehouden voor een betere afstemming van het programma met preferenties van studenten, voor een opleiding relevante werkgevers of de samenleving als geheel. Bij deze tweede aanbevelingen komen de achterliggende waarden van onderwijsprogramma's ter discussie.

Deze twee typen aanbevelingen zijn niet even makkelijk te 'verwerken', te benutten (Jeliazkova, 2002).

Hoe verloopt de benutting? Er blijkt in termen van Jeliazkova (2002) sprake van 'levels of processing review recommendations'. Een indeling blijkt behulpzaam, namelijk in gebruik of benutting ('utilization') op korte of lange termijn, en op instrumentele wijze of conceptuele wijze (Korsten, 1983; Scheele, Maassen & Westerheijden, 1998; Jeliazkova, 2002).

De eerste groep van aanbevelingen, die betrekking heeft op problemen binnen bestaande onderwijsprogramma's, laat zich over het algemeen direct instrumenteel verwerken, zo blijkt (Jeliazkova, 2012). We spreken in navolging van Duitse bestuurskundigen van beleid en beleidswijziging als '*Problemverarbeitung*'. De centrale vraag is bij de eerste groep steeds: functioneert het onderwijsprogramma zoals bedoeld? Het antwoord komt erop neer dat het onderwijs met een reactie op de eerste groep aanbevelingen min of meer gerepareerd wordt. Het kan gaan om triviale verbeteringen. De oproep van onderwijsstaven om in elkaars keuken te gaan kijken en te profiteren van geijkte praktijken past hierin ook (zie Jeliazkova, 2002).

Bij de tweede groep is instrumenteel en kortetermijngebruik lastiger (cf Jeliazkova, 2002). Hier raken de aanbevelingen de essentie, de kern van een onderwijsprogramma en zal vaker sprake zijn van bezinning op langere termijn, dus niet binnen enkele maanden, en in conceptuele zin (Jeliazkova, 2002). De centrale vraag is bij de tweede groep aanbevelingen: is het onderwijsprogramma vanuit een helicopter gezien eigenlijk wel op 'de goede dingen' gericht? Een meer transcendentiaal perspectief dus, dat kan dwingen om het hele onderwijsconcept en de centrale onderwijsdoelstellingen nog eens tegen het licht te houden door middel van een kritische '*frame-reflective discourse*', dus een debat over het onderliggende perspectief ('frame') en andere perspectieven (Rein & Schön, 1986; Korsten, 1988; Schön & Rein, 1994; Rein, 2006). Daarbij kan naar voren komen dat wat tot dan toe niet als problematisch is gezien of als niet conflicterend met programmadoelstellingen nu ter discussie komt en nieuwe cursussen worden opgenomen, bij voorbeeld omdat ze de studie meer bij de tijd brengen of zaken betreffen, zoals vereiste nieuwe vaardigheden, die in de toekomst meer gaan spelen. Misschien verandert met het naar voren komen van nieuwe doelstellingen de naam van een opleiding ook

wel, bij voorbeeld van 'bedrijfskunde' naar een opleiding 'bedrijfswetenschappen' of van 'bedrijfswetenschappen' naar 'managementwetenschappen'.

Naarmate de denkwijze van de visitatiecommissie ten aanzien van dergelijke aanbevelingen door de onderwijsstaf meer gedeeld wordt, is de kans groter dat het onderwijs ook in deze meer fundamentele zin verandert en daarmee de 'aanbevelingen van de tweede soort' van de commissie dus serieus genomen zijn. Dat is wat bedoeld wordt met geleidelijke en conceptuele benutting van aanbevelingen van visitatiecommissies in het w.o. (Jeliazkova, 2002).

Conceptuele benutting impliceert dat de achter onderwijs liggende *waarden* als gevolg van aanbevelingen van de tweede soort ter discussie komen. Dat blijkt het geval te zijn geweest met de rapportage van een visitatiecommissie op het gebied van milieustudies. Die rapportage leidde tot een bezinning op de rol van milieustudies in interdisciplinaire zin, meer algemeen in de universiteiten en zelfs in de 'moderne samenleving' (Jeliazkova, 2002).

De conclusie van de analyse van benutting van aanbevelingen uit visitatierapporten is dat zeker sprake is geweest van instrumenteel gebruik op de korte termijn, mede omdat het visitatiesysteem dwingt tot verwerking van zwakheden. De inspectie ziet daar op toe. Maar er is ook sprake van meer fundamenteel gebruik: 'The conclusion is that review recommendations, even when found problematic, have a significant influence on the way programs are run, and eventually contribute to the evolution of staff's deeper convictions and beliefs. Long-term changes on system level probably have more far-reaching effects on the quality of programs than the measures undertaken within a clear and undisputed framework. It is on this system level where change in culture in the long run can be observed' (Jeliazkova, 2002: 6). Het wordt door Jeliazkova een misvatting genoemd te veronderstellen dat de veranderingen op concreet en operationeel niveau in het w.o. slechts cosmetisch en oppervlakkig zijn geweest.

Moet de benutting ook niet bekeken worden in een interactieve context? Dat gebeurde ook wel. Zo is vastgesteld dat de reactie op een zeer kritische visitatie van onderwijs op het vlak van 'computer engineering' niet leidde tot een remedie als gevolg van interne conflicten aan de kant van de gevisiteerden (Jeliazkova, 2002: 6).

Kritiek a: Benutting door de inspectie

De benutting van visitatierapporten door opleidingsverantwoordelijken kan gevolgd worden door de vraag wat de inspectie met de rapporten deed.

Het visitatiestelsel is niet primair ontwikkeld om de onderwijsinspectie beter haar werk te laten doen. Dat neemt niet weg dat de inspectie wel visitatieprocedures en ook rapporten beoordeelt en de visitatieprocedures moet goedkeuren. Die taak wordt vanzelfsprekend ook ter hand genomen. Maar op systeemniveau komt er wel een kritiek, die in het begin van dit stuk niet is geformuleerd maar na een aantal jaren visiteren wel verwoording vond. De zelfstudies binnen een vakgebied zijn uiteraard als gevolg van protocollering vergelijkbaar, zo dat de visitatiecommissie ook de kwaliteit van opleidingen kan vergelijken. Maar in de tijd opeenvolgende visitatierapporten zijn *'slecht onderling vergelijkbaar'* - er is nogal eens sprake van dynamiek in opleidingen - en ook *tussen vakgebieden* is vergelijking moeilijk. Daardoor kunnen niet makkelijk op basis van de rapporten uitspraken over het gehele wetenschappelijk onderwijs gedaan worden, klaagt de Algemene Rekenkamer (2000: 4). Een merkwaardig verwijt want daarvoor is het visitatiesysteem zeker niet primair in het leven geroepen.

De inspectie komt er ook in andere zin wat bekaaid vanaf, want 'de wijze waarop de inspectie beslist of een onderwijsinstelling in aanmerking komt voor toezichtactiviteiten vormt een zwakke schakel'. De inspectie beoordeelt de reacties van onderwijsinstellingen op

visitatierapporten en wat blijkt? Als de reactie niet toereikend is, dan krijgt de onderwijsinstelling het predicaat 'zorgelijk'. Hoe de inspectie en instelling hiertoe komen, is de Algemene Rekenkamer niet geheel duidelijk.

Wat te doen? De Algemene Rekenkamer neigt naar verdere protocollering. De Rekenkamer schrijft namelijk: 'Om de stimulerende werking van het kwaliteitszorgsysteem voor de 'betere' opleidingen te vergroten zou er een *meer uitgewerkt beoordelingssysteem* moeten komen. Daarmee zou ook de onderlinge vergelijkbaarheid van de opleidingen gediend zijn. Het visitatieprotocol 2000-2005 voor het wetenschappelijk onderwijs en het basisbeoordelingskader voor visitaties in het hbo komen aan beide punten reeds tegemoet: deze bevatten meer uitgewerkte concrete normen met een meerpuntsschaal om het oordeel op te scoren' (AR, 2000: 5).

Kritiek b: daagt het stelsel het onderwijs uit?

Het voorgaande laat zien dat niet alleen het hoofdbeeld van de kwaliteit van opleidingen, maar ook de afzonderlijke analyses in de tweede cyclus over het algemeen een positief beeld opleveren. Maar er is ook in de tweede cyclus wel enige kritiek mogelijk. Een van die kritieken betreft *de stelselkritiek*. In dat verband is de blik van de inspecteur-generaal van het onderwijs relevant. Die heeft immers de positie die erop gericht is de verworvenheden én beperkingen bloot te leggen. Welke kritiek is er? Inspecteur-generaal Ferdinand Mertens stelt dat de neiging bestaat in het buitenland de loftrumpet te steken op het Nederlandse stelsel van visitaties in het gehele hoger onderwijs (w.o. en hbo samen).

Die waardering bleek waarneembaar op een internationale conferentie over kwaliteitszorg in het wetenschappelijk onderwijs die rond 1999 in Lissabon plaatsvond. In Nederland werden rond 1999 opleidingen disciplinegebonden gevisiteerd, bijvoorbeeld alle opleidingen rechtsgeleerdheid. In Nederland is dat doenlijk omdat het nog om een overzienbaar geheel van opleidingen gaat van een tiental universiteiten, maar Duitsland kent al veel en veel meer universiteiten en ook meer opleidingen rechtsgeleerdheid. Een land als Turkije kent nog meer universiteiten dan Duitsland. In landen met veel universiteiten bestaat de neiging om een ander aggregatieniveau te zoeken dan dat van opleidingen in een vakgebied, namelijk de visitatie van een faculteit of vergelijkbare faculteiten of van een universiteit. Met name Franse deelnemers aan die conferentie uitten zich in die zin. Visiteren moet overzichtelijk en uitvoerbaar zijn, zo luidde de redenering van onder andere Franse zijde. In Nederland werd een dergelijk aggregatieniveau aanvankelijk genegeerd, maar in het accreditatiesysteem is hier wel oog voor ontstaan.

Inspecteur-generaal Mertens komt hier nog niet mee maar met andere kritiek dan het beoordelingsniveau, een kritiek die wel het pad naar accreditatie zou effenen. Hij erkent de kwaliteit van visiteren en meent zelfs dat het systeem bepaalde problemen oplost, maar hij signaleert ook een beperking op systeemniveau. Mertens vindt dat de visitatiecommissies veel te veel binnen bestaande patronen opereren. We noemen die kritiek het '*begrensd uitdagingsverschijnsel*'. Visitaties zijn in zijn optiek te gesloten. Mertens gaat om zijn stelling uit te werken terug naar de politieke ontstaansgeschiedenis, naar de HOAK-nota.

Het Nederlandse kwaliteitszorgsysteem (van disciplinegebonden externe beoordeling, daarmee verbonden zelfevaluaties en visitatiecommissies, de rol van de inspectie en de overheid als toezichthouder) is in feite in zijn optiek het product van een *politieke afruil* tussen te vergaande statelijke regulering en instellingsautonomie (Mertens, 1999). In die afruil is het visitatiestelsel als een symbolische *trait d'union* (tussen instelling en inspectie) ontstaan in het concept van sturen op afstand. Daardoor blijven bepaalde problemen in het hoger onderwijs die er eerder ook al waren, volgens hem bestaan. Visiteren verhelpt bepaalde problemen in het

wetenschappelijk onderwijs niet. Hij stelt dat het stelsel nog onvoldoende in staat is om de veronderstellingen waarop het Nederlands hoger onderwijs is gebaseerd '*uit te dagen*' (Laman, 2003: 43). Waar gaat het dan om? Dat voert ons naar een derde kritiekpunt.

Kritiek c: de lage studielast wordt niet gecorrigeerd

Een van de problemen in het hoger onderwijs is wat pathetisch '*the unbeatable lightness of academia*' is genoemd. Sommige studies zouden te licht zijn en 'studenten zouden in de watten worden gelegd' (Bruinsma, 2000; Leeuw, 2003: 19). Daarmee komt de hier en daar te geringe studielast in beeld. Een kritiek is: visiteren leidt *niet* tot het aantrekken van de studiebelasting voor studenten. Waarom?

Visitatiecommissies gaan uit van de stand van zaken van opleidingen, op een bepaald moment. Ze vertrekken vanuit bestaand (opleidings)beleid. Dan wordt er dus geen injectie gegeven om zwakheden in dat bestaand beleid bloot te leggen. Doen instellingsbesturen en decanen wel de goede dingen in plaats van een analyse van of ze de dingen goed doen? Die vraag komt nauwelijks in visitaties aan bod. Ergo, visitatiecommissies *blijven binnen de bestaande denkkaders*. Vrijmoedige of gedurfde bespiegelingen over de toekomstige ontwikkeling van onderwijsdisciplines zal men niet aantreffen in visitatierapporten. Geen 'dynamiek' in visitatierapporten dus. De toekomst van een opleiding komt niet erg in beeld en voor zover dat wel het geval is, is doorgaans slechts sprake van een correctie op het bestaande, van wat in lijn met het denken van Charles Lindblom een incrementele verbetering binnen het bestaande systeem in de vorm van suggesties en eventueel aanbevelingen genoemd kan worden. Visiteren leidt dus tot een beleid van kleine stappen. Visitatierapporten neigen in handen van onderwijsbesluitvormers tot voorzichtige stappen voorwaarts en - in onze woorden - tot verlicht conservatisme.

Begrijpelijk? Het geven van vrijmoedige bespiegelingen over de gehele discipline was ook niet de taak van visitatiecommissies. Er worden in visitatierapporten ook *geen vanzelfsprekendheden in vakgebieden ter discussie gesteld*. Logisch want de gereputeerde vakgenoten die de visitatiecommissies bevolken zijn onderdeel van een '*community of assumptions*' (Mertens, 1999; Laman, 2003: 44). De visitaties zijn daarom wel als *gesloten* beschouwd. Die geslotenheid wordt versterkt doordat van een internationale vergelijking binnen de discipline (die aan visitatie wordt onderworpen) geen sprake is.

Ons commentaar is dat de kritiek van Mertens wel en niet terecht is. Wie innovaties in het kader van externe kwaliteitszorg wenst, moet eenvoudig andere instrumenten inzetten dan visitaties of visitatiecommissies een andere opdracht geven. Visiteren moet van de andere kant niet 'overvraagd' worden. De protocollen voor visitatiecommissies lieten tot 2000 helemaal geen vrijmoedige of uitdagende beschouwingen toe. Dat was niet voorzien, het was de opdracht niet. Dus slaat de kritiek in feite niet op die visitaties.

Beoordeling 9: De kosten in verhouding tot de baten

Eerder is aangegeven dat ook aandacht verdient of de baten de kosten van visiteren in het wetenschappelijk onderwijs wel of niet overstijgen. Daarbij is de vraag wiens kosten of baten het betreft. Een eerste perspectief is dat van de opleidingen zelf. Sommige sceptici merken op dat het voorbereidende werk aan de zijde van de opleidingen wel veel inspanningen vergt in termen van personele inzet en dat aan die kant de kosten de baten overschreden. Onderzoeksgegevens ontbreken daarover echter. Ontegenzeggelijk vergt het verzamelen en ordenen van gegevens en het schrijven van een zelfstudie het nodige werk maar in de tweede ronde is dat al makkelijker dan in de eerste.

Het tweede perspectief is dat van de instellingen. De onderwijsinstellingen dragen het merendeel van de visitatiestelselkosten. Die kosten zijn rond 1999 op bijna f 20 miljoen getaxeerd, pakweg 0.4 procent van de totale onderwijsuitgaven voor het hbo en w.o.; exclusief

verbeterkosten. Dat cijfer van 0.4 zou in vergelijking met de V.S. laag zijn omdat daar ongeveer 10 procent van het budget aan kwaliteitszorg besteed zou worden (Dill, 1997; Algemene Rekenkamer, 2000: 9).

Het derde perspectief is dat van de minister. De Algemene Rekenkamer stelt dat de visitatie een 'toereikend instrument is voor de minister om tegen relatief geringe kosten te bewaken dat de instellingen de hen toebedeelde verantwoordelijkheid voor de kwaliteit van het onderwijs waarmaken' (AR, 2000: 4). De Rekenkamer is de mening toegedaan dat het 'stelsel de minister in staat stelt om binnen de wettelijke en bestuurlijke kaders van het hoger onderwijs zijn verantwoordelijkheid te dragen voor de deugdelijkheid van het onderwijs en voor de besteding van publieke middelen die het hoger onderwijs ontvangt'. Dat betekent dat de kosten-batenbalans voor de minister op het moment van meting - in de tweede ronde - positief was.

Conclusie van beoordelingen in de tweede ronde: positief maar minder euforie

De conclusie over de tweede ronde wijkt beperkt af van de eerste ronde. De inspectie is zich de kritiek op visiteren bewust geweest. In de lof op de tweede visitatieronde (1994-1999) is de inspectie dan ook minder euforisch dan over de eerste ronde. Maar de inspectie is ook over het functioneren van het visitatiestelsel in het hoger onderwijs in de tweede ronde 'over het geheel genomen tevreden' (Laman, 2003: 44). De kritiek van de inspectie betreft de zelfstudies en verbeterinitiatieven. Er groeit een groep personen op die wel heel vaak zelfstudies opstellen en daar handigheid in krijgen. En het valt de inspectie op dat vooral in het w.o.-onderwijs instellingen zo weinig samen initiatieven nemen om het onderwijs te verbeteren. Als verbetermechanisme kent visiteren blijkbaar grenzen van instellingsautonomie.

6.5 Nadere analyse van visitatiekwaliteit in de derde ronde: 2000-2003

Na de eerste twee visitatiecycli (1988-1993; 1994-1999) volgde een derde ronde van visitaties in het w.o. (2000 en later). Men kan het de derde generatie visitaties in het wetenschappelijk onderwijs noemen. Het einde van deze periode valt rond 2003 als de accreditatie van start gaat.

In de derde ronde veranderde er weer iets. Nu kregen visitatiecommissies ook de mogelijkheid om naast een openbaar rapport een *management letter* uit te brengen aan de instelling. Met dit instrument kan een visitatiecommissie desgevraagd of op eigen initiatief nader concreet advies geven. Hoe is die ontwikkeling beoordeeld? De inspectie heeft zich hier neutraal tegenover opgesteld (Laman, 2003: 42). Ze kan immers geen invloed uitoefenen op de inhoud van 'management letters'. Wel valt te erkennen dat het werken met 'management letters' de transparantie van het systeem niet bevordert omdat andere instellingen niet kennis kunnen nemen van deze management letter voor een specifieke instelling.

Beoordeling 10: algemeen oordeel over derde ronde

Hoe is de waardering voor deze derde ronde? Het systeem van kwaliteitszorg in het Nederlandse hoger onderwijs wordt *nog steeds als 'werkbaar'* gezien (Laman, 2003: 45). De inspectie beschikt ook nu weer over meta-evaluaties: 15 uit het hbo en 15 uit het w.o.-onderwijs. Geen enkel van de 30 visitatierapporten wordt door de inspectie afgekeurd maar er zijn er ook weinig om 'door een ringetje te halen' (Laman, 2003: 45). Er komen wel enkele kritische kanttekeningen over de samenstelling en de werkwijze van de commissies naar voren maar die gelden meer voor het hbo dan het w.o.

Kritiek d: proeven van bekwaamheid

In deze derde periode van visiteren komt ook enige kritiek naar voren. We noemden al de drie kritiekpunten a, b en c en voegen hier een punt aan toe.

Na 1999 brengt de inspectie geen aparte rapportages meer uit over de kwaliteitszorg (Laman, 2003: 44). In het Onderwijsverslag 2000 van de inspectie wordt hierop wel ingegaan. De inspectie meldt dat visitatierapporten over het algemeen niet erg diepgaand ingaan op toetsing en examens, ook niet op de studielast (een punt van Mertens) en de meting daarvan, en de rendementen van opleidingen. Laman stelt: 'Thema's als de authenticiteit van proeven van bekwaamheid onder andere ten gevolge van nieuwe media, de spanning tussen individueel werk en groepswork bij toetsresultaten, de gevolgen van flexibilisering en de nog beperkte mate van operationalisering van nieuwe toetsmethoden zouden hoog op de visitatieagenda moeten staan, maar dat is nog niet het geval' (Laman, 2003: 45).

Conclusie van beoordelingen in de derde ronde: nog steeds 'werkbaar'

Tot welke algemene bevindingen leidt onze analyse in de derde ronde? Er worden in de derde ronde geen visitatierapporten afgekeurd maar de inspectie neemt een kritischer toon aan. Er komt stelselkritiek (Laman, 2003: 46). Zoals bij elk kwaliteitszorgsysteem openbaart zich 'na verloop van tijd een geheel van onderliggende assumpties, overtuigingen en probleempercepties waarvoor op productieve wijze een uitweg gevonden moet worden' aldus Huitema, Jeliaskova en Westerheijden (2003: 33). De vooravond van een evolutie naar accreditatie is aangebroken.

Douma voegt toe dat de eerder aangeduide visie van Mertens op het visitatiestelsel, hoewel in onze ogen grotendeels onterecht, minister Hermans van onderwijs de legitimatie verschaft voor de doorontwikkeling van visitatie naar accreditatie, die de Tweede Kamer accepteert. Maar er is meer. De invoering van bachelor-mastersysteem in het kader van de Bologna-verklaring impliceert dat een internationalisering in het onderwijs voorzien wordt. Studenten worden mobieler en moeten voor hun keuzen kunnen beschikken over een transparant systeem van beoordelingen. Ook docenten worden mobieler. Dan is het nodig om dat bacheloropleidingen en masteropleidingen beoordeeld zijn en de bekostiging verzekerd is. Ze moeten een accreditatie hebben ondergaan.

7 Terug naar de positieve en kritische beelden van visiteren

In deze studie is het visiteren in het wetenschappelijk onderwijs gedurende de periode 1985-2003, toen van accreditatie nog geen sprake was, op kwaliteit beoordeeld. Het visiteren van opleidingen in zelfde vakgebieden vond in drie cycli plaats. Het visitatiestelsel was alles dekkend; alle w.o.-opleidingen vielen er onder.

We begonnen deze analyse met te stellen dat er positieve en kritische oordelen waren over het visiteren in het wetenschappelijk onderwijs. Wat is nu de oogst?

Uit de analyse van het visitatiestelsel in het wetenschappelijk onderwijs in de periode 1982-2003 komt een positief hoofdbeeld naar voren. Ook andere auteurs die eerder de balans opmaakten, concluderen dat zoals Frederiks (1996), Van Berkel (2000), de Algemene Rekenkamer (2000), Huisman en Theisens (2001) en auteurs van inspectiezijde zoals Hulshof e.a. (2003) en Laman (2003). Dat hoofdbeeld wordt gedeeld door instellingsbesturen, de VSNU, de inspectie en de opeenvolgende ministers van onderwijs.

Enkele indicaties hiervoor zijn de volgende. Visitaties leiden bij de bij opleidingen betrokkenen tot een herkenbaar beeld (geldigheid: 'dit zijn wij') en erkend beeld (betrouwbaarheid: 'de foto is goed'). Zwakke plekken zijn in het onderwijs wel uit de weg geruimd (Laman, 2003: 47). Begrijpelijk want de inspectie voor het hoger onderwijs zag daarop toe. Visiteren in het w.o. leidde ook tot verbeteracties, vooral bij opleidingen die gedwongen waren tot verbeteren. Huisman en Theisens (2001: 171) schrijven: 'De intentie van de kwaliteitszorg was uitdrukkelijk gericht op het verbeteren van het onderwijs binnen de opleiding. Onderzoek toont dat deze

functie ook is waargemaakt'. Door middel van visitaties kon ook bijgedragen worden aan verantwoording van opleidingen door universiteiten aan de samenleving. Op de realisatie van dit doel bestaat geen kritiek.

Over het algemeen komt een positief beeld naar voren over de rol van de VSNU, de visitatiecommissies, de processtappen, de zelfstudies, de rapportages. De kosten-batenbalans is dus positief. In het verband van kwaliteitszorg en -verbetering moet ook melding gemaakt worden van specifiek overheidsbeleid dat in de jaren negentig gericht was op 'kwaliteit en studeerbaarheid'. In de jaren negentig is meer dan 700 miljoen gulden geïnvesteerd in ongeveer drieduizend projecten voor kwaliteitsverbetering. Die projectbudgetten werden niet slechts 'ge Gund' als ze een antwoord waren op kritiek in visitatierapporten maar dat er geen pogingen gedaan werden om in het wetenschappelijk onderwijs te innoveren kan op grond hiervan niet volgehouden worden (Jenniskens, 1997; Huisman en Theisens, 2001: 171).

De meeste eerder geuite bezwaren tegen verantwoording door visiteren zijn, als ze toetsbaar waren, dan ook naar het rijk der fabelen te verwijzen.

Het positieve hoofdbeeld kan in zoverre genuanceerd worden dat de inspectie in de derde visitatieronde minder euforisch was over visitatie in het wetenschappelijk onderwijs. Na een aantal jaren werden ook enkele zwakten duidelijk. Leeuw (2003), tot mei 2003 werkzaam als Hoofdinspecteur Hoger onderwijs, stelt dat het niet bij voorbaat zo is dat altijd met een visitatie ook alle stenen boven water worden 'gevisiteerd'. Andere kritiek is: een zekere versnippering in de beoordeling (checklist-denken), het niet uitdiepen door commissies van examenvraagstukken en het hier en daar niet vergroten van de te beperkte studielast. De inspectie heeft het (naderhand gewenste) neveneffect om uitspraken te kunnen doen over de kwaliteit van het hoger onderwijs niet op een makkelijke wijze via visitaties kunnen realiseren.

Nogmaals de beginselen van behoorlijk visiteren

De positieve waardering is grotendeels veroorzaakt doordat de visitatie gebaseerd is op een aantal beginselen van goed visiteren (Korsten, 2004a). De volgende beginselen waren van toepassing op het 'succes' van visiteren van opleidingen in het wetenschappelijk onderwijs.

1. Er is sprake van een initiatiefnemer tot visiteren. In geval van een dekkend visitatiestelsel kunnen immers niet alle visitaties gelijktijdig starten maar is er een clustering en volgorde.
2. Het coördinerend agentschap (VSNU) voert regie en treedt dus op als regisseur.
3. Het visiteren geschiedt op basis van een bepaalde opzet: het protocol. Daarin is voorzien in een beoordelingskader.
4. De procedure voorziet in de aanwezigheid van een visitatiecommissie met een opdracht (gegeven in het initiatief en het protocol).
5. De visitatiecommissie bestaat uit onder meer uit vakgenoten en is onafhankelijk.
6. De dataverzameling (in het kader van een zelfstudie) vindt plaats op basis van het protocol (in casu het beoordelingskader).
7. Er is sprake van betrokkenheid van de gevisiteerde bij de dataverzameling.
8. De procedure voorziet in interactie met betrokkenen (docenten, studenten, etc.) door bezoek ter plaatse.
9. Er is sprake van contact met externen (zoals een werkveldcommissie).
10. Het proces resulteert in een openbaar rapport, dat in conceptvorm wordt voorgelegd aan betrokkenen voor beoordeling en eventuele correctie van ontbrekende of niet juiste feiten.
11. Er vindt enigerlei bezinning plaats bij de gevisiteerde op de uitkomsten en zeker 'toezicht' daarop.

Van deze lijst behoeft geen punt te worden afgevoerd. En er hoeft geen punt te worden opgevoerd.

Op grond van deze in 2004 gepubliceerde lijst (Korsten, 2004a: 320) is een *congruentieverwachting* te formuleren: Als in een andere sector dan wetenschappelijk onderwijs een visitatie - op grond van de elf genoemde punten - veel gelijkenis vertoont met de gang van zaken in het wetenschappelijk onderwijs zal de kans groot zijn op erkenning van de procedure en herkenning van de uitkomsten. De check van visiteren in het hbo bevestigt deze regels ons inziens. Deze principes vonden we alle terug bij hbo-visitaties (zie ook rapport commissie-Korsten).

Toch is er ook wel kritiek geleverd op het stelsel van kwaliteitszorg die aan het visiteren *niet* geweten kan worden. Wie bijvoorbeeld wil dat het wetenschappelijk onderwijs vernieuwt, moet niet van visitaties verwachten dat ze daartoe bijdragen tenzij men in plaats van retrospectieve visitaties prospectieve visitaties vraagt of andere instrumenten inzet. Want de in de periode 1985-2003 uitgevoerde visitaties in het w.o. geven uitkomsten in termen van 'voldoendes' of 'onvoldoendes' of kwalificaties van vergelijkbare strekking. De begrijpelijke reactie is dan om in hetzelfde stramien te blijven werken door juist te trachten om zwakten weg te werken.

Eerder was ook sprake van kritiek. Tot de serieuzere kritiek behoren de volgende kritiekpunten.

1. *De checklistbenadering.* Bij visitaties wordt gepoogd een *integraal beeld* te krijgen van opleidingen zonder alles te onderzoeken. Hoe dat te doen? Protocollering met een expliciet kader wijst de weg. Het gevaar van een *checklistbenadering* dreigt. Er mag vooral niets vergeten worden. Dus dient zich tellen, turven en afvinken aan. Als daardoor een gefragmenteerd of versnipperd beeld van een opleiding ontstaat, kan de samenhang zoek raken. Dit bezwaar is beperkt van betekenis. Het is opgemerkt, maar het is geen kritiek die veel gehoord is. Immers, de integraliteit en de protocollering is te voren in de sectoren met betrokken decanen besproken. Men kan er wel kritiek op hebben maar betrokkenen hebben zich gebonden (motto: 'zo gaan we het doen'). Bovendien kunnen onderwijsdecanen de samenhang aangeven en uitwerken.
2. *Het protocolleringsfetisjisme.* Een tweede kritiek houdt in dat het gevaar van protocolleringsfetisjisme op de loer licht: de neiging bij critici om het integrale beeld van een organisatie, organisatiedeel of opleiding zelfs verder te willen *verbreden* (het 'governance'-beeld te verruimen), *te verfijnen* (niet slechts in termen van voldoende of onvoldoende oordelen maar met oog voor meer oordeelsmogelijkheden) of *te dynamiseren* (niet slechts een momentopname maken maar meerjarencijfers gebruiken). Jeliaskova en Westerheijden (1999) noemen een van deze drie punten. Zij analyseerden acht opleidingen en stellen vast dat er behoefte bestaat aan meer discriminerende beoordelingen die het verschil verduidelijken tussen goede en beste opleidingen. Dat is een pleidooi voor een technische (benchmarkachtige) verbetering, die later al in gang is gezet (zie Algemene Rekenkamer, 2000). Dit bezwaar was in het w.o. beperkt. Wat de verbreding en verfijning betreft constateerden leden van de commissie-Korsten die de hbo-opleidingen op het gebied van bestuurskunde/overheidsmanagement tegen het licht hielden, dat sprake was van een uitgebreider normenkader dan bij het belichten van w.o.-opleidingen, zodanig dat de leden vooraf aan de visitatie eerst getraind moesten worden om ermee te kunnen omgaan.

3. *Het bezwaar van de selectieve verbetering.* In het meergenoemde rapport van de Algemene Rekenkamer uit 2000 wordt gepleit voor meer genuanceerde beoordelingen dan in termen van 'voldoende' of 'onvoldoende' want w.o.-opleidingen die 'voldoende' scoren ervaren geen stimulans tot verder leren in de vorm van verbeteringsinitiatieven. De Rekenkamer meldt dus *onderbenutting* (weinig aandacht voor 'voldoendes') en *selectieve benutting* (wel aandacht voor 'onvoldoendes'). Gevisiteerde instellingen berusten in hun 'voldoende', alsof het luie leerlingen van hun eigen instelling zijn. Het daadwerkelijk verbeteren van opleidingen richt zich dus voornamelijk op zwakke plekken, niet of nauwelijks op opleidingselementen die juist voldoende zijn, meent ook Laman (2003). De aard van de benutting is verder te kenschetsen als: directe en instrumentele benutting, niet zozeer als indirecte en conceptuele benutting (Weiss, 1979; Korsten, 1983; Bekkers e.a., 2004; Johnson e.a., 2009). Zou een verdere differentiatie in oordelen dan wel meer verbeteringsdruk opleveren of uitdagen tot direct en instrumenteel verbeteren? Dat een differentiatie tussen 'goed' en 'nog beter' zou leiden tot druk voor de goede opleidingen om nog verder te verbeteren is een onbewezen stelling. Verbeteringsdruk ontstaat vooral als gevolg van het toezicht van de inspectie, vermoedelijk minder of niet van differentiaties in oordelen. De inspectie komt langs om te controleren wat instellingen deden met de rapporten en heeft dan vooral oog voor de zwakke punten, dus voor het wegwerken van de 'onvoldoendes'.

4. *Het bezwaar van de systeemimmanente beperking.* De kritiek van inspecteur Ferdinand Mertens was dat visiteren het hoger onderwijs onvoldoende uitdaagt. Verbeteraars blijven doorgaans in het bestaande opleidingsstramien werken, zegt Mertens. Dat is begrijpelijk omdat het ontwikkelen of herzien van een opleidingsprogramma niet alleen een cognitieve zaak is van uitdenken maar ook een relationele van overleggen en eventueel onderhandelen. Als vakgroepen een aandeel in de uitvoering van een onderwijsprogramma hebben, willen ze dat om drie redenen behouden: a) omdat de expertise van de vakgroep staat voor beroepstrots en erkenning van het belang van een vakgebied; b) het deelnemen aan een opleidingsprogramma een soort werkgelegenheidsgarantie is en indirect is het een mogelijkheid om studenten naar het eigen vakgebied te zuigen, wat ook weer werk verschaft; en c) zo ook de expertise van de individuele medewerkers benut wordt. Van radicale veranderingen in onderwijsconcepten of onderdelen is vaak als *gevolg* van visitaties, let wel, *geen* sprake. Dat is een verwijt dat begrijpelijk is, maar niet spoort met afspraken. Men kan het de visitatiecommissies niet verwijten. Alleen vanuit een ander visitatieconcept zou er een kans bestaan op meer uitdaging en discussie over of opleidingen wel de 'goede dingen' doen. Werkte die kritiek van Mertens nog door? Leeuw trad in zijn voetsporen. Hij spreekt van *systeemimmanente beperkingen* (ook Laman, 2003: 46). De toenmalige hoofdinspecteur hoger onderwijs Frans Leeuw (2001) merkt op dat begrijpelijk en verdedigbaar is dat activiteiten worden gestart die beogen een stimulans te zijn tot hogere kwaliteit en doelmatigheid in het hoger onderwijs: de nota *'Keur aan kwaliteit'* (2000). Maar verandert er dan weinig of niets aan universitaire onderwijsprogramma's? Onze indruk, die gebaseerd is op kennis van onderwijs in de politicologie, bestuurskunde, bedrijfskunde en economie uit de periode 1985-2010 aan de Radboud Universiteit, de Universiteit van Maastricht en de Open Universiteit, is dat er elk jaar wel degelijk wat verandert aan universitaire onderwijsprogramma's; soms meer, soms minder. Maar dan niet alleen of uitsluitend als gevolg van kwaliteitszorg maar a. als gevolg van bezuinigingen (die dwingen tot maatregelen, waaronder vakken schrappen of vakinhoud laten opgaan in een nieuw vak), b. als gevolg van het ontstaan van een kleinere vakgroep door bezuinigingen of mobiliteit van docenten (een vak kan niet meer gegeven worden als een vertrokken specialist niet vervangen kan worden), of c. als gevolg van hergroepering van vakgebieden (facultaire politiek). Bezuinigingen

kunnen leiden tot structuurwijzigingen en die weer tot wijzigingen in onderwijsprogramma's. Soms is dat een nadrukkelijk bedoelde maar niet steeds transparante operatie. Een structuurwijziging kan *op zich* al, zonder dat sprake is van bezuinigingen leiden tot een veranderd vakkenpalet hoewel iets dergelijks niet expliciet gezegd wordt. Dat neemt niet weg dat als gevolg van deze bezuinigingen en structuurwijzigingen doorgaans geen sprake is van paradigmaveranderingen in onderwijsprogramma's. Een radicale marktuitval zal wel een grote wijziging kunnen oproepen, zoals het opheffen van volledige studierichtingen en de creatie van een nieuwe richting (een vorm van 'policy succession').

5. Is het glas met benutting van bevindingen van visitatiecommissies half vol en half leeg? Volgens Leeuw (2001) is de 'impact' of benutting van visitaties in het huidige bestel 'gering' omdat alleen kleine stappen worden gezet; het glas is bij wijze van spreken half leeg of nog iets leger. Daarmee kiest Leeuw een wat andere toon dan de Algemene Rekenkamer, de inspectie en de minister. Hij spreekt van perverse leereffecten omdat de betrokkenen intussen geleerd hebben 'hoe om te gaan met visitaties'. Hij verwijst naar onderzoek over monitoring en auditing waarin dergelijke onbedoelde neveneffecten ook volop zijn aangetoond (o.a. Smith, 2000). Voorbeelden? We noemen al strategieën van strategisch gedrag om kritiek te omzeilen (zoals het evaluatiestroomeffect) en voegen daar nu drie nieuwe aan toe maar bedenk dat een visitatiecommissie lang niet altijd ruimte hiervoor geeft. Een vorm van strategisch gedrag is het aanbrenge van *faseringen* bij opleidingen die toevalligerwijze samenvallen met bezoeken van visitatiecommissies. Daardoor kan beargumenteerd worden dat net voor de komst van de visitatiecommissie 'alles fundamenteel gewijzigd is'. Maar ook weer niet zoveel dat de commissie zegt dat ze zich van een oordeel gaat onthouden. De commissie kan daarop besluiten een opleiding alsnog buiten het rapport te houden. Dan heeft de opleidingsverantwoordelijke een probleem. Tweede voorbeeld: de *orchestratie* van een commissiebezoek door studenten slim te selecteren of ze 'in te fluisteren' of 'enkele suggesties mee te geven'. Derde voorbeeld: door verslaggeving in een zelfstudie zo te presenteren dat uitspraken *niet goed toetsbaar zijn*. Dat kan door verleden, heden en toekomst door elkaar te laten lopen. Er zijn intussen tal van strategische zetten in het visitatiespel doordacht maar bewijs voor het systematisch optreden van specifieke strategische handelwijzen in met naam en toenaam genoemde gevisiteerde w.o.-opleidingen – anders dan het aanbieden van een diner en ontvangst 'met bloemen' - zijn niet geleverd en niet door onderzoekers waargenomen. Veel kan de commissie voorzien omdat leden op de hoogte zijn met strategisch gedrag. Dus, Leeuw heeft door te wijzen op strategisch gedrag wel een relevant aspect van 'doorlichten' benoemd maar daarmee is nog geen substantiële doorwerking van strategisch gedrag aangetoond. Op grond hiervan is de verwachting dat als in een andere sector een visitatie veel gelijkenis vertoont met de gang van zaken in het wetenschappelijk onderwijs hetzelfde effect optreedt: strategisch gedrag wordt vermoed maar is daarmee nog niet aangetoond.
6. *Het inburgeringseffect*. We komen bij nog een kritisch punt dat in lijn is met het vorige. Als een opleiding voor de tweede of derde keer gevisiteerd wordt, dreigt een 'inburgeringseffect'. De spelregels zijn intussen bekend, het opstellen van een zelfstudie wordt 'een klusje', de betrokkenen kennen het klappen van de zweep bij de ontvangst van de commissie en men weet bij benadering wat men mag verwachten. Er ontstaat een groep opstellers van zelfstudies die dit werk vaker hebben gedaan en de formuleringen langzamerhand kent. Op basis hiervan is de verwachting dat als in een andere sector een visitatie veel gelijkenis vertoont met de gang van zaken in het wetenschappelijk onderwijs hetzelfde effect optreedt: bij herhaalde visitatie is sprake

van een inburgeringseffect. Visitatie in het w.o.-onderwijs zou zo gezien een instrument zijn dat (lichte?) slijtage vertoont. Betrokkenen bij onderwijs zijn zich zeker na eerste visitaties bewust van beperkte mogelijkheden om hier en daar strategisch gedrag bij de opstelling van een zelfstudie aan de dag te leggen.

Onverwachte effecten

Zijn er ook onverwachte effecten van visiteren opgetreden? Media hebben uitkomsten van de visitaties gebruikt om opleidingen onderling te 'ranken'. Daardoor is de functie van het visitatiestelsel licht gewijzigd, stellen Huisman en Theisens (2001: 171). De uitkomsten zijn in de loop der jaren steeds meer gebruikt om de verschillende opleidingen binnen een zelfde discipline onderling te positioneren. Deze functiewijziging is mede beïnvloed door landelijke vergelijkingen tussen studieplaatsen en opleidingen, zoals die op grond van studentoordelen en oordelen van hoogleraren zijn uitgevoerd en verschenen in bladen als Elsevier, menen deze auteurs. Instellingen nemen zelf geen afstand hiervan want ze gebruiken uitkomsten van visitaties vervolgens in de strijd om studenten in reclames en advertenties. Universiteiten doen zelf in de jaren negentig ook mee aan die 'ranking' (Huisman en Theisens, 2001: 171). Zo wordt in de public relations gesproken van de beste universiteit van de klassieke universiteiten, van de beste van de nieuwe universiteiten of de beste van de grote universiteiten en van de kleine. Wie zo doorgaat stelt vast dat elke universiteit wel ergens de beste in is. Een verband tussen de kwaliteit van de opleiding en toenemende interesse van studenten voor een opleiding is overigens niet sterk, maar beelden van geboden onderwijs doen er wel toe (Huisman en Theisens, 2001: 171). De studiekeuzen blijken niet erg beïnvloed te worden door kwaliteitsverschillen. Reisafstand speelde in de onderzoeksperiode nog altijd een belangrijke rol.

8 Enkele algemene conclusies

Met visiteren in het visitatiestelsel van het wetenschappelijk onderwijs zijn pro- en contra-argumenten verbonden. Anders gesteld: het stelsel laat zich positief verdedigen, maar er is ook kritiek op te uiten. We noemden een aantal positieve en kritische punten in het begin van dit stuk. Zijn hierover nu uitspraken te doen en zo ja, wat zijn die argumenten bij nader inzien waard?

Dat visiteren in het wetenschappelijk onderwijs positief wordt beoordeeld, impliceert vermoedelijk dat de positieve argumenten bevestigd worden. We noemden in het begin van dit stuk een aantal argumenten pro positieve visitatie:

- a) het argument van de brede steun voor opzet en werkwijze;
- b) het argument van de kwaliteit en betrouwbaarheid van de momentopname;
- c) het argument van de benuttingspotentie;
- d) het argument van de behoorlijkheid van de procedure.

Deze argumenten blijven na onze analyse van onderzoek naar externe kwaliteitszorg in het wetenschappelijk onderwijs overeind.

De eerste twee argumenten worden op empirische gronden ondersteund: het visiteren genoot in de periode van onderzoek brede steun in de universitaire wereld van bestuurders en andere verantwoordelijken voor opleidingen, alsmede bij de inspectie, het departement van onderwijs en de Tweede en Eerste Kamer. En de kwaliteit van de visitatie was weliswaar intersubjectief als gevolg van de oordelen van de visitatiecommissie maar solide: opleidingsdecanen leverden een zelfstudie aan, waarop de commissie zich mede baseerde, en de bevindingen van de commissie werden door opleidingsdecanen over het algemeen herkend en erkend. Dat visitatiecommissies alles dat aangeleverd werd onkritisch voor 'zoete koek' slikten, is niet

aangetoond en evenmin dat de decanen de commissieleden masseerden om te komen tot welgevallige bevindingen over de eigen opleiding.

Het derde proargument slaat op de doelstelling van 'visiteren als leren'. Deze doelstelling lijkt gerealiseerd. Kwaliteitszorg vroeg en vraagt volgens de officiële leer de deelname van veel betrokkenen, die elk in een nieuwe ronde lering kunnen trekken uit een vorige ronde. De leereffecten leiden vervolgens tot interne dynamiek van kwaliteitszorgsystemen. Huitema e.a. (2003) nemen daar een decennium later met betrekking tot leren in het wetenschappelijk onderwijs geen afstand van. De inspectie zag hier op toe en door ons geraadpleegd onderzoek geeft geen indicaties dat de benutting afwezig was. Dat die benutting van bevindingen zich concentreerde op 'onvoldoendes' is een inperking, een begrijpelijke inperking omdat de natuurlijke neiging bestaat om het goede te behouden en het slechte te corrigeren. De benutting van de bevindingen van visitatiecommissies is niet maximaal geweest voor wie nastreeft dat het goede te verbeteren is naar 'het beste' (in spreektaal: meer benutting had gekund).

Het vierde proargument betreft de beginselen van behoorlijk visiteren in het wetenschappelijk onderwijs. In het begin van dit stuk is gesteld dat aan positieve visitatieprocessen in het w.o. een aantal beginselen ten grondslag ligt (Korsten, 2004a). We herhalen ze kort: 1. de aanwezigheid van een initiatiefnemer die een paar uitgangspunten voor visiteren aangeeft, zoals de volgorde in het visiteren opleidingen (niet alles tegelijk); 2. de aanwezigheid van een coördinerend agentschap dat regie voert, zoals het starten van een visitatiepilot en de formulering van de aard van de visitatie (retrospectief en/of prospectief?); 3. de aanwezigheid van een protocol over de opzet van het visiteren (fasering), een rol voor een zelfstudie of andere dataverzameling, het bestaan van een commissie en een beoordelingskader; 4. de aanwezigheid van een visitatiecommissie; 5. de samenstelling van de commissie (bestaande uit vakgenoten) en de onafhankelijkheid daarvan; 6. de dataverzameling op basis van een protocol; 7. betrokkenheid van de gevisiteerde bij de dataverzameling (i.c. een zelfstudie); 8. interactie met betrokkenen door bezoek ter plaatse ('site visit' genoemd); 9. contact met externen over het te visiteren object (eventueel een werkveldcommissie); 10. een openbaar visitatierapport, dat in conceptvorm wordt voorgelegd aan betrokkenen bij elke opleiding voor commentaar op juistheid en volledigheid van de feiten; 11. bezinning op de uitkomsten van de visitatie als onderdeel van de procedure en toezicht daarop van de inspectie.

Beginselen

De positieve betekenis van deze beginselen wordt in het voorgaande bevestigd. We stelden het volgende vast. Dat het visiteren ergens een start moet vinden, is door niemand bestreden. De colleges van bestuur van de universiteiten hebben het initiatief genomen, in VSNU-verband. De procescoördinator VSNU is (o.a. door de inspectie) geprezen vanwege de organisatirol. Met visiteren is bovendien niet over een nacht ijs gegaan. Het visiteren kwam immers na een 'pilot' van de grond. En dat er sprake was van een protocollaire opzetnota van de zijde van de VSNU is eveneens positief geweest. De universitaire bestuurders werden daarbij volop betrokken. Visitatiecommissies werden samengesteld en erkend als goed werkende, onafhankelijke commissies van vakgenoten ('peer reviewers'). De zelfstudies waren geschikt omdat ze impliceren dat degenen die het onderwijs kennen en uitvoeren ook de gegevens moeten kunnen verstrekken op grond waarvan commissies oordelen. Het proces van visiteren was voorts volop iteratief: de voor de opleiding verantwoordelijken kregen de gelegenheid suggesties te doen over de samenstelling van de visitatiecommissie, de relevante feiten over de opleiding in de vorm van de zelfstudie aan te leveren en later in het proces commentaar te geven op een conceptrapport van de visitatiecommissie. De ontvangers van de visitatierapporten, dus de betrokkenen bij de opleidingen, hebben de procedures geaccepteerd en zich in de bevindingen herkend en die ook erkend. Er zijn voor zover bekend geen

visitatierapporten afgekeurd. Daarmee kunnen we vaststellen dat de beginselen van behoorlijk visiteren die bij de start van het proces waren gehanteerd na drie rondes van visiteren in het wetenschappelijk onderwijs herbevestigd worden. Er zijn geen alternatieven bedacht voor bepaalde beginselen. De suggestie is *niet* gedaan om opleidingen in plaats van door visitatiecommissies te laten doorlichten door een adviesbureau. Er is evenmin gesteld dat zelfstudies beter afwezig konden blijven omdat de commissies wel zelf hun inventariserend werk konden doen. Commissies konden de zelfstudies niet missen.

Was het toezicht van de inspectie op de (kwaliteit van de) visitatierapporten eigenlijk wel nodig? Ook op dit punt heeft weinig tot geen kritiek geklonken omdat de universitaire wereld zich niet wilde mengen in het bestaan en de taakstelling van de onderwijsinspectie.

Conclusie: Het stelsel van visitaties in het wetenschappelijk onderwijs is in de periode 1985-2003 in de kern solide gebleken. De Inspectie van het Onderwijs heeft over het algemeen de gang van zaken rond visiteren erkend als adequaat. De minister van onderwijs stelt in 2000 dat het visitatiestelsel 'internationaal in hoog aanzien staat'.

De doelen

Visiteren dient in de w.o.- sector twee doelen: verantwoorden en verbeteren. Vanuit beide perspectieven gezien, is visiteren in beginselen evident 'boven de streep'. De verantwoording is geslaagd als de rapportages zijn voltooid en de inspectie de procedure en uitkomsten erkent. Dat was het geval. De inspectie heeft er ook op toegezien of colleges van bestuur en de opleidingsdecanen de bevindingen van de visitatiecommissies voldoende verwerkt hebben. Dat bleek in het algemeen het geval.

Welke verbeteringen zijn opgetreden door visitatie? Twee effecten zijn gebleken: het *anticipatie-effect* en het *selectief concentratie-effect* om 'zwakheden weg te werken'. Het anticipatie-effect impliceert dat betrokkenen bij onderwijs, zoals decanen, er alleen al vanwege de noodzaak tot opstelling van een zelfstudie toe neigen om voorafgaand aan de visitatieprocedure te reflecteren op wat sterk en zwak is en na te denken over interne beleidsaanpassingen. Dus er ontstaat hier en daar al een doorwerking van de visitatie voorafgaand aan het uitbrengen van de zelfstudie, dus nog voor de visitatiecommissie ter plaatse komt kijken. Het selectief concentratie-effect impliceert dat instellingen ertoe neigen om bij de score 'voldoende' geen verbetertraject te starten, maar zich te richten op het wegwerken van de 'onvoldoendes', zoals die in een visitatierapport naar voren komen. Op zich is dat logisch. Wie de dakgoten heeft bekeken, gaat niet vanzelfsprekend alle dakgoten vervangen maar probeert te volstaan met het dichten van de lekken of het wegwerken van andere mankementen. Dakgoten worden doorgaans eerst geheel en al vervangen als de algehele staat beneden peil is.

De gang van zaken in het wetenschappelijk onderwijs, gebaseerd op de genoemde beginselen, was in het algemeen zo geschikt dat een toepassing bij visiteren in het hoger beroepsonderwijs, dat mede voortkwam uit de HOAK-nota, vergelijkbare vormen aannam. Ook daar zien we de genoemde beginselen van behoorlijk visiteren allemaal terugkeren. Er is dus grond voor ondersteuning van *de congruentieverwachting*: visiteren in het hbo geschiedt grosso modo vanuit dezelfde beginselen als het w.o.. Ergo, visitatiestelsels die overeenkomen met de visitatie in het wetenschappelijk onderwijs en hbo kunnen als deugdelijk worden beschouwd. De onderwijsinspectie is ook een dergelijke mening toegedaan.

Het glazen huis en de transparantieparadox

Zijn er in de reeks van beginselen van behoorlijk visiteren nog onderdelen die bij nader inzien misschien overbodig of minder belangrijk zijn?

We hebben even getwijfeld of openbaarheid (en transparantie) van prestaties even belangrijk is als de andere beginselen. Wie immers wantrouwend en machiavellistische kijkt naar openbaar gemaakte informatie over prestaties van een organisatie kan *een transparantieparadox* veronderstellen:

hoe meer de prestatiegegevens naar buiten worden geplaatst, hoe minder druk er is van politici om zich ermee in te laten' (De Beus, 2006).

Burgers, docenten, managers, opinielidderen en anderen kunnen zich immers op de uitkomsten van de publiek gemaakte meetresultaten storten en het debat er over aangaan. Daardoor kunnen politieke representanten of andere verantwoordelijken in het openbaar bestuur achterover leunen: 'zij zijn niet aan zet'; 'van hen hoeft het initiatief om iets met de uitkomsten te doen, even niet te komen'. Van uitkomsten van metingen gaat bovendien nogal eens geen 'handelingsdwang' uit, menen Swierstra en Tonkens (2008). Dat versterkt de politieke luiheid, of eufemistischer gesteld de politieke terughoudendheid. Dat bezwaar geldt echter meer voor vaak onduidelijke of niet goed interpreteerbare tabellen uit benchmarkingrapporten dan voor visitatierapporten, waar een onderwijsinspectie naar kijkt.

Ko Scheele (2003: 75), in 2003 als inspecteur hoger onderwijs verbonden aan de Inspectie van het Onderwijs, stelt 'een van de lastigste dilemma's van hoger onderwijskwaliteitssystemen is de mate van openheid en openbaarheid' (ook Karsten e.a., 2004). De Europese markt van w.o.-opleidingen was toen hij dit schreef 'nog verre van transparant'. Belangrijker is volgens hem dat een van de bezwaren tegen openbare rapportage is dat het wordt misbruikt om te 'ranken' en vervolgens public relations te gaan bedrijven. Het gevaar bestaat dat tijdens het proces van het maken van rangordes allerlei voor studenten en anderen relevante informatie buiten de berichtgeving en overwegingen valt (Scheele, 2003: 75).

Dat we toch vasthouden aan het belang van openbaarheid is eenvoudig.

Het eerste argument luidt: geheim regulier onderwijs op het vlak van wiskunde of diergeneeskunde hoort niet te bestaan. De staatsveiligheid is doorgaans als gevolg van openbaarheid niet in gevaar. Er is dus geen noodzaak om de openbaarheid ter discussie te stellen.

Tweede argument: visitatierapportages waarvan hoogleraren en andere docenten van andere universiteiten die op eenzelfde vakgebied werkzaam zijn of burgers niet mogen weten wat de inhoud ervan is, zijn een absurditeit. Visitatierapporten horen openbaar te zijn.

Derde argument: Als de visitatierapporten niet openbaar zouden zijn, kan iemand die werkzaam is op het gebied van pakweg sociologie niet kennis nemen van een visitatierapport over rechtsgeleerdheid, waarin ook de rechtssociologie aan bod komt, of van een visitatierapport over de psychologie, waarin de sociale psychologie aan de orde is. Dat is merkwaardig omdat dan zo iemand niet betrokken zou kunnen zijn bij de bezinning op kwaliteitsverbetering.

Vierde argument: bevindingen over de kwaliteit in het onderwijs, zoals kwaliteitskaarten van voortgezet onderwijs, worden al lang en breed gepubliceerd. 'Rankings' bestaan al. Dus ligt het in de rede om ook de kwaliteit van w.o.-opleidingen openbaar te maken.

Vijfde argument: er hoort forumwerking en daarmee debat mogelijk te zijn over de externe kwaliteitszorg. Dat debat is alleen mogelijk als rapporten verkrijgbaar zijn.

Zesde: er zijn geen negatieve effecten van openbaarheid van visitatierapporten. Dus handhaven we het beginsel dat visitatierapporten openbaar behoren te zijn.

De openbaarheid heeft overigens wel tot enige schichtigheid aanleiding gegeven vanwege de vrees dat dagbladjournalisten kritische punten uit visitatierapporten zouden kunnen uitvergrooten tot incidenten. Dit speelde niet zozeer in het onderwijs dat in grote delen van de samenleving toch als abracadabra wordt gezien, zoals wiskunde- of scheikundeonderwijs. Decanen, vertegenwoordigd in enkele disciplineoverlegorganen, hebben hier en daar wel

geprobeerd zich in te laten met het persbericht over het visitatierapport omdat een (niet evenwichtig geformuleerd) persbericht een verkeerde toon kan zetten en de uitkomst van de visitatie in hun ogen scheef kan trekken. In die zin zijn gedurende de verschillende visitatierondes ervaringen opgedaan.

Een andere specifieke ervaring was, dat in een eerste ronde de opleidingen bestuurskunde en politicologie samen zijn gevisiteerd. In een volgende visitatieronde is een splitsing aangebracht: twee afzonderlijke visitaties in plaats van een. De decanen op het gebied van bestuurskundeopleidingen hadden weinig neiging om een gezamenlijk rapport te laten beïnvloeden door zwakheden in politicologieopleidingen. Beelden van opleidingen en vrees voor het inzakken van aantallen studenten kunnen dus hier en daar ook van invloed zijn op de opstelling aan de vooravond van de start van een visitatie van (zelfde) opleidingen.

Terug naar de kritiek

We begonnen deze beschouwing met de argumenten die voor en tegen visiteren pleit(t)en. Wat zijn de contra-argumenten, de argumenten tegen een visitatiepraktijk, waard gebleken? Daarop moeten we nog ingaan, zeker om het verwijt van een positieve geneigdheid ('bias') te hebben, te ontlopen. We noemden de volgende kritische argumenten:

- a) het argument van de beoordelingsziekte: er is sprake van teveel beoordelingen in de Alma Mater, van te veel verborgen kosten van al die opeenvolgende beoordelingen, van een afname van arbeidsmotivatie bij medewerkers als gevolg van de 'beoordeelziekte' en bij critici is sprake van weinig geloof in de voordelen van prestatiemeting; het onderwijs is in het algemeen al deugdelijk;
- b) het argument van de uitlokking van pervers gedrag: universitaire medewerkers gaan zich als gevolg van de visitaties belangengericht gedragen, blijkend uit het entertainmenteffect in het onderwijs om zo goed in de evaluaties van colleges te voorschijn te komen en uit het co-auteurseffect van samen elkaar aan meer publicaties helpen;
- c) het argument van de disciplinerende: universitaire medewerkers worden 'geknecht' in een universitaire setting van steeds meer bedrijfsmatigheid; de knechten moeten dingen doen die ze niet willen of die niet primair zijn;
- d) het (macro-)argument van overdreven luxe: blijkbaar kan de Nederlandse staat zich in financieel opzicht een hele 'visitatie-industrie' veroorloven die in tijden van grotere schaarste niet te betalen is (zou in de 'universiteit' van Verwegghistan ook visitatie voorkomen?);
- e) het argument van de zwakke legitimatie van de starter: wel andere organisaties de maat nemen maar dat bij jezelf niet doen? (laat het departement zich beoordelen?; laten de colleges van bestuur zich de maat nemen?; laat de onderwijsinspectie zichzelf ook weer beoordelen?);
- f) het argument dat visiteren politiek speelgoed is: het gaat niet uitsluitend om beter onderwijs maar voor politici om schone handen te houden als incidenten optreden en mediakritiek op het onderwijs volgt (de branche heeft immers zgn. zelfreiniging door visitatie);
- g) het argument van de zwakte van het integraal beeld: de essentie van onderwijs blijft ongemeten; sommige aspecten of elementen van onderwijs blijven in visitaties buiten schot, zoals oog voor de samenhang in het onderwijs?;
- h) het argument van de matige toegevoegde waarde van visitatierapporten;
- i) het argument van de weinige leereffecten die van rapporten uitgaan;
- j) de adresseringskwestie: wie moet zich de bevindingen van de commissie eigenlijk aantrekken?; iedereen dus niemand?;
- k) de negatieve kosten-batenbalans: het visiteren kost meer dan het opbrengt (zie ook a, h en i);
- l) het tijdelijkheidsargument: wat je ook verbetert, het onderwijs zal vooral veranderen door externe omstandigheden zoals crises (daarom maar van visiteren af zien?).

Deze argumenten ogen deugdelijk maar snijden bij nader inzien niet veel hout. Daarvoor hebben we de volgende argumenten. a) De kritiek slaat meer op onderzoeksvisitaties dan op

opleidingsvisitaties en heeft betrekking op de cumulatie van beoordelingen. b) Deze kritische argumenten kunnen weliswaar ingebracht worden in de discussie maar ze hebben in de periode 1985-2003 nog geen grote rol gespeeld in het afsprakenstelsel in de sector van het wetenschappelijk onderwijs. Het departement van onderwijs, het parlement, de onderwijsinspectie, de colleges van bestuur, de opleidingsdecanen in de disciplineoverlegorganen, ze stemden allemaal in met het toenmalige visitatiestelsel. c) Externe kwaliteitszorg was met de aanvaarding van de HOAK-nota van het ministerie een onvermijdelijkheid geworden. d) In de praktijk had het visitatiestelsel genoeg toegevoegde waarde en genoeg leereffecten om gehandhaafd te blijven, zo laat onderzoek zien. De baten overstegen de kosten. De contra-argumenten hebben het uiteindelijk dus af moeten leggen tegen de positieve argumenten.

Was het onderzoek naar visitaties misschien toch eenzijdig waardoor er geen specifieke uitdieping van de contra-argumenten plaatsvond? Veel onderzoek was inderdaad wel gericht op (in termen van Hoogerwerf) de bestaande beleidstheorie achter de kwaliteit van visiteren en de rol van visiteren in het hoger onderwijs, dus op het geheel van veronderstellingen over doelen, middelen en doel-middelrelaties (zie o.a. Leeuw en Van der Graaf, 1988).

Er is volgens ons inderdaad nog te weinig onderzoek gedaan vanuit een ander perspectief. Het feit dat een artikel van Frey en Osterloh (2008) in vertaling verscheen in de NRC en er kritische beschouwingen verschenen in andere bladen zegt wel wat. Blijkbaar hebben Frey en Osterloh en ook Boomkens en Lock de vinger wel gelegd op bezwaren maar die analyses verschenen eerst rond 2008, vijf jaar na afloop van drie rondes visitaties van wetenschappelijke opleidingen. De analyses van de filosoof Lock kregen ook wel aandacht in de dagbladen maar landden vooral bij een belangengroep, BON, die kritisch is op ontwikkelingen in het voortgezet- en middelbaar beroepsonderwijs (zoals het competentiegericht onderwijs), minder op het wetenschappelijk onderwijs. De analyses van Lock en Boomkens kwamen minder aan in de universitaire arena's of die van de onderwijsinspectie (verg. Hulshof e.a., 2003).

Was er dan helemaal geen kritiek die hout snijdt? Wel degelijk. De kritiek is te verdelen in twee categorieën: terechte en onterechte kritiek. Een belangrijke onterechte kritiek was de systeemkritiek, die eerst rond 1999-2000 opkwam: visiteren zou het onderwijs te weinig uitdagen. Dat is op zich wel juist maar een onterecht verwijt. Het is kritiek die eigenlijk niet aan het visiteren geweten kan worden. Wie het onderwijs echt wil uitdagen moet *andere* instrumenten inzetten of had expliciet visitaties daarop moeten *inrichten* en dat ook als een doel moeten meegeven aan commissies.

Een tweede kritiekpunt betrof de dreigende fragmentatie in de oordeelsvorming door visitatiecommissies. Deze kritiek dreigde inderdaad, maar het is een wolk geweest waaruit uiteindelijk weinig regen is gevallen.

Dan de opgave aan een commissie om zich een integraal beleid te vormen van elke gevisiteerde opleiding. Door het checklistdenken (fragmentatie in oordeelsvorming) zou de samenhang van vakken onvoldoende in beeld komen. Ook deze kritiek heeft veel commissies niet geraakt want opleidingsdecanen konden zelf de samenhang aangeven als dreigde dat de commissie deze samenhang zou vergeten. Een storm in een vingerhoed, bij nader inzien.

En hoe zat het dan met het strategisch gedrag vanuit de opleidingen? Dat dergelijk gedrag denkbaar is, is zeker (Scheele, 2003; Willemsen en Leeuw, 2007; De Bruijn, 2008). Dat dergelijk gedrag voorgekomen zal zijn, ligt zelfs voor de hand. Maar dat dergelijk gedrag substantieel zou zijn geweest en tot veel vertekening van bevindingen van commissies zou hebben geleid, is niet aangetoond. Een belangrijke uitkomst want Schillemans stelt:

'De belangrijkste opgave voor visitaties is om niet te detoneren door strategisch gedrag. Wanneer visitaties verantwoordingsarrangementen zijn, dus een rol spelen voor

minister(ie)s en buitenstaanders, bestaat het risico op strategisch gedrag dat de beoogde leerwaarde van visiteren ondermijnt' (Schillemans, 2009: 163).

Een 'lock-in'-effect, dat universitaire medewerkers wel meededen en meedoen aan al de onderwijsvisitaties omdat het nu eenmaal moet en weinig enthousiast zijn, lijkt reëel, is dus aannemelijk maar is niet nader onderzocht.

Is er nog een 'controlitis'-effect in termen van demotivatie van het personeel? De toegenomen meetcultuur ondervindt in het algemeen onder professionals nogal wat kritiek. Swierstra & Tonkens (2008: 137) stellen: 'Zo zegt 64% van de professionals dat de toenemende nadruk op controle-instrumenten, productiecijfers, targets, tijdschrijven, registratie en verantwoording een negatieve invloed heeft op het werkplezier. Voor 41% is het reden (geweest) om ander werk te zoeken'. Het onderzoek waarop deze auteurs terugvallen, had betrekking op docenten uit het middelbaar onderwijs en zorgverleners die lid zijn van een vakbond. Daarmee kan dit onderzoek gerelativeerd worden.

Hoe zit het met een dalende motivatie onder wetenschappelijk personeel als gevolg van verantwoording in het algemeen en visitaties in het bijzonder? Aangenomen mag worden dat universitair personeel in het algemeen iets meer bewegingsvrijheid heeft dan ander onderwijspersoneel. Wat zegt onderzoek over (de)motivatie van het w.p.? Het GoodWork Project[®] van Gardner c.s. heeft bij uitstek aandacht voor motivatie van professionals, waaronder universitair personeel. Uit vijftien jaar studie komt naar voren dat 'goed werk' onder meer van de grond komt als wenselijk gedrag verinnerlijkt wordt en de werknemers niet tegengewerkt worden door externe krachten, zoals marktwerking of politieke of sociale krachten (Gardner, Csikszentmihalyi & Damon, 2009: 48; Smid, 2012). Een geheel van visitaties op het vlak van onderwijs en onderzoek is te beschouwen als een externe kracht. Een terugloop in motivatie van onderwijspersoneel is daarom niet helemaal uit te sluiten. Maar stel dat die motivatie terugliep dan nog is niet zeker of er in de periode 1985-2003 een verband met opleidingsvisitaties is geweest. Een verband is niet waargenomen (Steijn, 2005, 2006; Kluijtmans, 2010).

Is er misschien verantwoordingsdruk opgetreden en vastgesteld? Onderzoek naar verantwoordingsdruk richt zich op personeel in de zorg, huisartsen, leraren en politiemensen en niet direct op universitair personeel. Noordegraaf en Sterrenburg (2009: 249) stellen vast: 'Professionals lijken minder moeite te hebben met het afleggen van verantwoording dan door opiniemakers wordt gesuggereerd'. We nemen daarom vooralsnog aan dat ook de verantwoordingsdruk onder universitair onderwijspersoneel dat betrokken is bij visitaties meevalt.

De conclusie kan niet anders zijn dan dat de kritiek op opleidingsvisitaties in het w.o. in de periode 1985-2003 overtrokken was en grotendeels onterecht. Of de kritiek wel opgaat voor de periode na 2003 is niet door ons nagegaan. We handhaven daarom de stelling dat de opleidingsvisitaties in de periode 1985-2003 over het algemeen succesvol waren.

9 Enkele verwachtingen voor verder onderzoek naar visitatiekwaliteit

De analyse van visiteren in het wetenschappelijk onderwijs heeft niet alleen bevindingen opgeleverd over de kwaliteit van visiteren maar ook een rijke oogst aan bevindingen of conclusies die te formuleren zijn als hypothesen voor verder onderzoek. We noemen deze hypothesen in lijn met Haverland & Yanow verwachtingen, dus verwachte uitkomsten voor verder onderzoek. Het oogmerk van formulering is om verder onderzoek naar visiteren op andere gebieden, met name op dat van gemeentebesturen, van focus te voorzien. Voorgaande

beschouwing heeft een rijke oogst opgeleverd, met als kern de congruentieverwachting: naarmate een visitatiestelsel meer gelijkenis vertoont met het stelsel in het wetenschappelijk onderwijs - af te meten aan de beginselen van behoorlijk visiteren - is de kans groter dat de kwaliteit van die andere visitatie als positief is te beoordelen. De opsomming aan verwachtingen is overigens niet limitatief.

Er zouden om te beginnen verwachtingen zijn te formuleren over de vraag waarom organisaties tot visitatie overgaan en waarom organisaties 'meten'. Wouter van Dooren formuleert in zijn dissertatie *'Performance measurement in the Flemish Sector- A supply and demand approach'* (2006) al zes hypothesen over de vraag naar de startomstandigheden. Maar dit vraagstukken laten we terzijde.

Overzicht van verwachte uitkomsten van verder onderzoek naar visiteren in andere sectoren dan wetenschappelijk onderwijs

1. De congruentieverwachting
2. De verwachting van contextafhankelijkheid
3. De doelmatigheidsverwachting
4. De verwachting van visitatiekenmerken als randvoorwaarde voor succes
5. De verwachting van de concluderende kracht als succesindicator
6. De verwachting van de uitblijvende aanbevelingen
7. De verwachting van de grote proceskracht van het protocol
8. De verwachting van het immer te beperkte integraal kader voor beoordeling
9. De verwachting van de visitatie-illusie
10. De verwachting van strategisch gedrag bij visiteren
11. De verwachting van het inburgeringseffect
12. De verwachting van de benuttingsimpuls
13. De verwachting van gezwellenheid
14. De verwachting van niet collectief verbeteren ('het op zichzelf'- effect)
15. De verwachting van slijtage in de externe kwaliteitszorg en de roep om evolutie
16. De verwachting van de begrensde uitdaging van visitatieobjecten
17. De verwachting van de begrensde verantwoording.

We geven in het navolgende steeds eerst de inhoud van de verwachting en daarna een verduidelijkende toelichting. Er wordt hier niet ingegaan op operationaliseringsvraagstukken en hoe de verwachting verder te onderzoeken.

1: De congruentieverwachting

Het visitatiebestel in het wetenschappelijk onderwijs was succesvol ('satisficing'). Daaraan heeft de aanwezigheid van de 'beginselen van behoorlijk visiteren' krachtig bijgedragen. De congruentieverwachting houdt in dat de kans groot is dat een visitatiebestel uit een sector dat ook gebaseerd is op de 'beginselen van behoorlijk visiteren' eveneens succesvol is. In vergelijking met de tweede verwachting noemen we dit de *maximale* congruentieverwachting.

In dat verband dient zich het grotestedenbeleid als object van vergelijking aan, want daar opereerden ook visitatiecommissies in een meerlagig bestuurlijk systeem.

De toelichting op deze verwachting luidt als volgt.

In deze analyse is de kwaliteit van visiteren in het wetenschappelijk onderwijs, zoals toegepast in de periode 1985-2003, beoordeeld. Het hoofdbeeld van visiteren in het wetenschappelijk onderwijs is positief ('satisficing') en overheerst de kritische bevindingen. Deze conclusie is niet alleen door ons getrokken maar op tal van momenten ook door betrokken actoren in de

sector van het wetenschappelijk onderwijs zelf. Daartoe heeft het toepassen van de beginselen voor behoorlijk visiteren bijgedragen. Elf beginselen zijn genoemd. Die vinden we niet alleen terug in het w.o. maar ook in het hbo. Ze hebben de tand des tijds doorstaan nadat ze eerder zijn geformuleerd.

Dit inzicht, dus de betekenis van de beginselen voor de positieve kwaliteit van visiteren in zowel het w.o. als het hbo, leidt tot de verwachting dat het toepassen van een visitatiesysteem in een andere sector, die gelijkenis heeft met visitatie in het wetenschappelijk onderwijs, ook als kwalitatief positief uit de bus zal komen. Dit noemen we daarom de (positieve) congruentieverwachting.

Dit laat onverlet dat eerder genoemde effecten ook daar kunnen voorkomen. Zoals het filteringseffect: het oordeel over een opleiding wordt gefilterd door de aangeleverde zelfstudie die is opgesteld door de instelling en door de kijk van de visitatiecommissie. En het hoera-effect: betrokkenen bij de opleiding die gevisiteerd wordt, willen bij voorkeur een positieve uitkomst uit de visitatie. En het evaluatiestroomeffect: om tot een positieve uitkomst te komen, zullen verantwoordelijke betrokkenen bij een opleiding graag veel gegevens verstrekken.

Is voor de congruentieverwachting nog aanvullend bewijs te vinden? Het hoger beroepsonderwijs heeft heel veel weg gehad van de visitaties van opleidingen in het wetenschappelijk onderwijs. Dat is ervaren bij een visitatie op het vlak van bestuurskunde/overheidsmanagement en door de blik te richten op andere hbo-visitaties (commissie-Korsten). Ook daar is, zoals bijvoorbeeld het visitatierapport *'De kunst van het balanceren'* uit 2003 laat zien, gewerkt met procescoördinatie (HBO-Raad); protocollen; onafhankelijke visitatiecommissies; zelfstudies per opleiding; site visits; mondelinge terugkoppeling na de bezoeken; een conceptrapport per instelling waarop een reactie mogelijk was; een openbaar overallrapport van alle gevisiteerde opleidingen. De protocollen verschilden wel enigszins van die in het wetenschappelijk onderwijs maar na zekere tijd trad enige harmonisatie op (Laman, 2003).

We bekeken in dat verband het rapport van de commissie-Korsten (2003), die de opleidingen bestuurskunde/overheidsmanagement heeft doorgelicht, om zo voeling te krijgen met visitaties in het hbo. Ook over de visitaties van hbo-opleidingen worden over het algemeen positieve bevindingen gemeld. Ook hier is sprake van geldige en betrouwbare uitkomsten uit door de inspectie geanalyseerde visitaties van uiteenlopende opleidingen, zoals op het gebied van hotelmanagement, milieu (mer), journalistiek, werktuigbouwkunde, chemische technologie, bouwkunde (Hulshof e.a., 2003; Laman, 2003). Logisch eigenlijk, want de processtappen bij visiteren in disciplinekader hebben een grote mate van vanzelfsprekendheid, aldus Hulshof e.a. (2003). Het vergelijken van zelfde vakgebieden tussen verschillende instellingen wordt als meer voor de hand liggend gezien dan het vergelijken van 'vreemde' disciplines of opleidingen per zelfde instelling.

Samenvattend: de congruentieverwachting kon in het hbo al bevestigd worden. Gegevens die van inspectie en andere zijden gerapporteerd zijn wijzen dit uit. Vertonen visitaties elders hier overeenkomst mee of niet?

2: De verwachting van contextafhankelijkheid

Visiteren van w.o.-opleidingen vond en vindt plaats in een bepaalde sturingscontext, zo is in het begin van dit stuk gesteld. Visiteren kwam in het w.o. op toen een overgang werd gemaakt van sterk hiërarchische sturing door de centrale overheid van onderwijs naar meer autonomie voor instellingen en zelfregulering op het vlak van onderwijs (sturen op afstand; HOAK-nota). Het ging dus om visiteren in een bepaalde interbestuurlijke setting: meer autonomie voor

instellingen, een visitatiestelsel waarin iedereen deelneemt, verplichte deelname, meta-evaluatie door de inspectie, openbare uitkomsten.

De verwachting van contextafhankelijkheid impliceert het aan de kant schuiven van de maximale congruentieverwachting. De tweede verwachting houdt in dat de sturingscontext, de bestelafspraken en het object van visitatie met zich kan brengen dat meerdere beginselen afwezig zijn in een bestel dat afwijkt van het bestel in het wetenschappelijk onderwijs maar de visitatie toch succesvol is in termen van vier kwaliteitsclusters: ontwerp kwaliteit, kaderkwaliteit, proceskwaliteit en concluderende (en bruikbaarheids-)kwaliteit. Vanuit de contextafhankelijkheid gezien is de tweede verwachting de verwachting van het *minimale stramien* te noemen. Maximale congruentie (eerste verwachting) versus minimaal stramien (tweede verwachting) dus. Deze tweede verwachting worden op het einde van dit stuk toegelicht.

Toelichting: Deze tweede verwachting neemt afstand van de eerste verwachting. Dat komt omdat visitatie in het wetenschappelijk onderwijs wel typisch is door de aanwezigheid van de startomstandigheden (colleges van bestuur maken samen afspraken, in overeenstemming met het ministerie), de samenwerking in *VSNU-kader* bij de uitvoering, de verplichte deelname van instellingen door zelfbinding, de aanwezigheid van een krachtige gelegitimeerde regisseur (VSNU) en vooral door de aanwezigheid van een op de visitatie toezichthoudende onderwijsinspectie. Vooral verdient dus vermelding dat visitatie in het w.o. verbonden is met toezicht. Niet uit te sluiten lijkt dat er meerdere wegen naar Rome leiden en dus de afwezigheid van een of meer beginselen tot gevolg heeft dat de visitatie toch succesvol ('satisficing') kan zijn; dus dat een minimaler stramien tot succes leidt vanwege de bijzondere context. Zo is bij voorbeeld de visitatie van grote steden in het kader van het grotestedenbeleid weliswaar verantwoordingsgericht maar komt toezicht hier in een heel ander licht te staan.

De verwachting van solide (dus succesvolle) visitaties in eerstgenoemde verwachting houdt in dat structureel visiteren in een visitatiebestel een succes is (dus 'satisficing' is) als die visitaties steeds gebaseerd zijn op (nagenoeg) *alle* beginselen van behoorlijk visiteren die achter visiteren in de sector wetenschappelijk onderwijs schuil gingen. Als aan die beginselen vastgehouden worden is de kans groot dat visitaties een succes worden. Dat betekent dat onder meer sprake is van een regisseur, een onafhankelijke visitatiecommissie, een zelfstudie, een bezoek ter plaatse, een conceptrapport waarop gereageerd mag worden en een openbaar eindrapport.

Helemaal irreëel is die stelling niet. Ook Schillemans stelt: 'Visitaties verlopen over het algemeen volgens hetzelfde stramien' (Schillemans, 2008: 165). Hij noemt het visitatiecollege, dat een bezoek aflegt om zich een oordeel te vormen over een bepaald aspect van een organisatie. Die doet dat aan de hand van een beoordelingskader, een zelfevaluatie en gesprekken. Op basis van deze bronnen volgt een visitatieverslag met bevindingen en eventueel aanbevelingen. Dat wordt meestal openbaar gemaakt. Daardoor kan het verslag ook door verschillende actoren worden benut, zoals een minister, belanghebbenden, journalisten. 'Daarna is het aan de gevisiteerde organisatie om de aanbevelingen te verwerken' (2008: 165). Hij stelt nadrukkelijk dat gebruik van bevindingen in zoverre mogelijk is omdat de visitatie 'feedbackinformatie genereert vanuit alle belanghebbenden'. Schillemans ziet voor zich dat rond visitaties een aantal inrichtingsbeslissingen genomen moeten worden. Hij noemt er zeven.

Volgens deze eerste congruentie- en stramienverwachting is het mogelijk dat alle succescriteria contextonafhankelijk zijn. Anders gesteld: wil visiteren van grotestadsbesturen, die met het rijk een bondgenootschap sloten om het grotestedenbeleid krachtig uit te voeren,

een succes zijn, dan zullen alle elf beginselen achter deze visitaties schuil moeten gaan. Hetzelfde geldt voor de visitatie van politiekorpsen of woningcorporaties. Maar niet uit te sluiten is (volgens de tweede verwachting) dat een of meer van deze beginselen niet zijn aan te treffen en de visitatie toch succesvol in termen van 'satisficing' is. Dat is de inhoud van verwachting twee.

De vraag is zal het praktijksucces een smal basisstramien zijn met veel afwijking (verwachting twee) of robuust gelijk stramien met nauwelijks afwijkingen (verwachting een).

Sommigen menen dat die contextafhankelijkheid juist van groot belang is. Elke sector zijn eigen inrichting. Congruentie leidt veel te veel tot standaardisering. Het is eerder nuttig dat de visitaties pluriform zijn dan uniform, menen Swierstra en Tonkens (2008) in de voetsporen van Scott's analyse in *'Seeing like a State'* (Scott, 1996).

Behn (2003) houdt ons in *'Why measure performance?'* meer in het algemeen voor dat het zelfs zeer gewenst is om een meting afhankelijk te laten zijn van doel en context. De ondertitel bij het artikel luidt: *'different purposes require different measures'*. Maar hij accentueert meer de doelstellingen dan de context als zodanig.

3: De doelmatigheidsverwachting

De doelmatigheidsverwachting houdt in dat visitaties gebaseerd op de essentie van de visitatieaanpak in het wetenschappelijk onderwijs niet onaanvaardbaar ondoelmatig zijn. Tot de essentiële elementen van visiteren behoren: de aanwezigheid van een regisseur, de aanwezigheid van een protocol, een zelfstudie evenals een site visit, spreken met externen, het voorleggen van het conceptrapport aan opleidingsdecanen. Deze verwachting spoort met de verwachting dat beginselen van behoorlijk visiteren belangrijk zijn om van visiteren een succes te maken. Als die beginselen nodig zijn, inclusief bijvoorbeeld een zelfstudie, dan is de kans klein dat de procedure ondoelmatig is. Dan bevat het visitatiebestel geen overbodigheden.

Toelichting:

Sinds pakweg 1990 is de aandacht voor bureaucratiesering in het onderwijs toegenomen (Onderwijsraad, 2004; Swierstra en Tonkens, 2008). Raakt bureaucratiesering ook het visiteren? Op visiteren is wel als kritiek gekomen dat de kosten-batenbalans scheef was. De baten zouden betrekkelijk laag zijn en de kosten hoog. Voor kosten kan dan gedacht worden aan a) de voorbereidingskosten binnen de universiteit zelf van de visitaties door de opstelling van een zelfstudie (in de periode 1985-2003); b) de gang van zaken tijdens het bezoek van de visitatiecommissie: veel betrokkenen moeten voor de commissie verschijnen of zich beschikbaar houden; c) de noodzaak te moeten reageren op een conceptrapport van de commissie; en d) de nasleep van het verschijnen van een rapport, zoals het opstellen van een persbericht. Is van ondoelmatigheid bij - in de periode 1985-2003 uitgevoerde - visitaties feitelijk sprake geweest? Uit rapportages van de onderwijsinspectie en uit specifieke studies naar aspecten van visiteren, zoals in het voorgaande genoemd, komt op dat vlak geen overtuigend bewijs naar voren. De noodzaak van een zelfstudie is alom erkend.

Wel is het zeker zo geweest dat de opstelling van eerste zelfstudies intern binnen universiteiten nogal wat werk gekost heeft. Dat was deels een gevolg van het feit dat de interne kwaliteitszorg in universiteiten nog beperkt was en van 'onwennigheid' met het opstellen van zelfstudies sprake was. Ook op de gang van zaken tijdens de 'site visit' - te kenschetsen als een heel intensieve operatie voor de commissie - was weinig aan te merken. Van overbodigheden in de procedure bleek over het algemeen geen sprake.

Conclusie: dat de gang van zaken rond de visitaties bij de universiteiten veel werk met zich bracht, zeker in de eerste ronde, is ontegenzeggelijk juist maar weinigen hebben een alternatief daarvoor gegeven of gesuggereerd dat zelfstudies niet nodig waren.

4: De verwachting van visitatiekenmerken als randvoorwaarde voor succes

De eerste twee verwachtingen gingen over het begrijpen van het succes van visiteren. Ligt dat aan de aanwezigheid van elf beginselen van behoorlijk visiteren (verwachting één) of kan het goed gaan met een beginsel of enkele beginselen minder.

Maar bepaalde visitaties zullen mogelijk ook helemaal niet of minder succesvol zijn. Welke? De verwachting van visitatiekenmerken als randvoorwaarde voor succes houdt in dat de basiskennmerken van invloed zijn op of misschien zelfs bepalen en omgrenzen wat visitatie kan bieden.

Dat er sprake is van een visitatiebestel in het wetenschappelijk onderwijs impliceert niet vanzelfsprekend dat een visitatiebestel ten aanzien van het functioneren van grotestadsbesturen in het verband van het grotestedenbeleid daarmee gelijk staat. Het is als met vissen. Wie het vissen van vissers analyseert moet er rekening mee houden dat vissen op zee vanaf een boot wat anders is dan vissen op een 'stekkie' bij een rivier of meertje. De kenmerken van een visitatie maken uit voor de kans op het optreden van bepaalde gevolgen. Bij voorbeeld, als een commissie meedenkt over verbetermogelijkheden voor een concrete opleiding stijgt - ceteris paribus - de kans dat een onderwijsdecaan ook aan noodzakelijke verbeteringen gaat werken.

De toelichting luidt als volgt. Visiteren kan verschillende vormen aannemen.

1. Zo kan een visitatiecommissie evaluerend terugkijken op de prestaties van de organisatie ('retrospectieve visitatie') en/of vooruitkijken op toekomstige handelingsperspectieven ('prospectieve visitatie').
2. Visiteren kan geschieden door een geheel onafhankelijke commissie van vakgenoten maar ook door een anders samengestelde commissie. Welke heeft legitimiteit in het visitatiebestel?
3. Een derde onderscheid is dat in besloten of openbare visitatie. Visiteren kan in een verschillende openbaarheidscontexten plaatsvinden. Visiteren kan tot een openbaar rapport leiden of in volledige (semi-)beslotenheid plaatsvinden, zoals bij ziekenhuisvisitaties. Is het beginsel van behoorlijk visiteren 'openbaarheid' essentieel of niet?
4. Verder kan visiteren gericht zijn op verschillende doelen: een organisatie verantwoordt zich over prestaties en een organisatie poogt wellicht ook te leren van zwakheden en verbeterpunten.
5. Visitaties kunnen naast conclusies aanbevelingen bevatten (grote concluderende kracht en grote aanbevelende kracht?).

Doorgaans zijn combinaties van deze variabelen aan de orde. Visitaties in het wetenschappelijk onderwijs waren terugkijkend en beperkt prospectief. De bijeenkomsten van en met de commissie waren besloten maar het eindrapport was openbaar.

Veel te visiteren of gevisiteerde organisaties geven aan zich te willen verantwoorden en te willen leren. Daartoe wordt de 'past performance' gezien. De mate waarin sprake is van prospectie verschilt echter aanzienlijk. Visitaties in het w.o. en hbo waren primair gericht op een beeld van de actualiteit en op terugkijken. Eerst meer recent kwam in het w.o. de mogelijkheid van het opstellen van een management letter.

Een *uitsluitend* retrospectieve visitatie zal weinig zeggen over de toekomstkansen van het gevisiteerde thema of object. Een prospectieve visitatie zal juist ingaan op de toekomst van een thema of object (eventueel condities, weerstanden, scenario's) maar weinig zeggen over het verleden.

Hoe prospectiever, hoe meer kans op benutting van een visitatierapport. Anders gesteld, als een visitatie geen prospectieve elementen kent, zal een rapportage vermoedelijk toch niet steeds benut worden.

5: De verwachting van de concluderende kracht als succesindicator

Een indicatie voor een succesvolle visitatie is of een visitatierapport verschijnt met duidelijke conclusies in het licht van het door de visitatiecommissie gehanteerde beoordelingskader.

Toelichting: Een onderzoek kent een probleemstelling. Een van daaruit uitgevoerd - op besluitvorming gericht - onderzoek zal bij voltooiing bevindingen bevatten en daarmee conclusies kennen. Zo is het ook met visitatieprocessen. Ook visitatieprocessen zijn te zien als een vorm van onderzoek.

Een visitatie die een rapport oplevert, zonder conclusies, is in onze ogen een fiasco. Visitaties in het wetenschappelijk onderwijs bevatten overigens steeds voldoende concluderende kracht. De inspectie moest toezien op de kwaliteit van visitatierapporten en die heeft over het algemeen visitatierapporten geaccepteerd; ze waren voldoende concluderend in termen van het uitspreken van oordelen als 'voldoende' of 'onvoldoende'. Een rapport zonder duidelijke conclusies zou afgewezen zijn.

6: De verwachting van de uitblijvende aanbevelingen

Te verwachten zou zijn dat een visitatierapport niet alleen concluderende kracht moet hebben maar ook aanbevelende kracht om de visitatie tot een succes te maken. Want duidelijke conclusies en duidelijke aanbevelingen vergroten de bruikbaarheid van een rapport en daarmee zal de kans op benutting van een rapport door een opleidingsdecaan en betrokken docenten toenemen. Ergo, zou daarmee het lerend vermogen kunnen stijgen. Maar deze veronderstelling is te eenvoudig. Een visitatiecommissie is geen adviesbureau dat zich richt op een specifieke problemdiagnose van een onderdeel van een groter geheel en in het onderwijs moeten vakgenoten zelf zoeken naar wat hoe beter moet. We kiezen daarom op basis van visitaties in het wetenschappelijk onderwijs voor een *andere* verwachting, namelijk de verwachting van de uitblijvende aanbevelingen.

Deze verwachting van de uitblijvende aanbevelingen houdt in dat het *niet* opnemen van aanbevelingen in een visitatierapport niets afdoet aan de positieve kwaliteit van een visitatierapport. Het niet opnemen van aanbevelingen in een visitatierapport hoeft de benutting van een rapport niet in de weg te staan als de conclusies duidelijk zijn en aangeven waar een zwakke plek zit, en als betrokkenen zelf moeten zoeken naar verbetering.

Toelichting:

Visitatierapporten in het wetenschappelijk onderwijs bevatten, zeker in de eerste ronde, over het algemeen wel conclusies maar géén aanbevelingen; dat later, in de derde ronde, hier en daar door een visitatiecommissie 'management letters' werden uitgebracht is een ander verhaal. Omdat oordelen van visitatiecommissies over een opleiding uit het wetenschappelijk onderwijs op zich vaak al een richtingwijzer bevatten voor verbetering maar niet concreter kunnen zijn, zijn geen aanbevelingen in rapporten opgenomen. Stel immers dat het rendement van een opleiding kritisch wordt beoordeeld omdat het te laag zou zijn, dan kan een visitatiecommissie niet eenvoudig aangeven hoe dat te verhogen. Daarvoor is nader onderzoek en overleg tussen betrokken docenten nodig. En het ontbreekt een commissie eenvoudig aan

tijd om dit werk zelf te doen. Tweede voorbeeld: stel dat examinatoren ertoe neigen te hoge cijfers voor afstudeerscripties te geven, welke aanbeveling dan te geven? Visitatiecommissies laten het aan betrokkenen over om in dit geval naar oplossingen te zoeken.

Er is overigens nog een andere redenering mogelijk om aanbevelingen terzijde te laten. Die redenering grijpt aan bij de basale en integrale beoordeling van een opleiding.

De visitatie van een organisatie of onderdeel ervan is te vergelijken met een basaal oordeel van een schoolarts over de gezondheid van een kind afgemeten aan tien kleine testjes, zoals onder meer de tongtest (die indicatief is voor bepaalde aspecten van de menselijke gezondheid), de test van de longfunctie (diep ademen), de test van de motoriek (kniebuigingen). Daarmee krijgt de arts, mede gebaseerd op de stand van onderzoek naar gezondheid, een integraal, maar ook globaal beeld van de gezondheid van het kind. Dat beeld is niet het beeld van een specialist, dus niet de kijk van een kinderarts, orthopeed, gynaecoloog, neuroloog of uroloog. De arts maakt slechts een integrale scan die tot de bevinding kan leiden dat een nader onderzoek nodig is naar een aspect of onderdeel van het menselijk lichaam. Dan zal een verwijzing via een huisarts naar een specialist plaatsvinden. Visitaties hebben daar wat van weg; er is sprake van een integrale kijk en niet van één specialistische beoordeling of een beoordeling door twee of drie specialisten. Een visitatiecommissie die een opleiding bekijkt, zal een iets degelijker beeld dan de schoolarts verkrijgen, mede door a) de samenstelling van de commissie (niet slechts een vakgenoot), b) door de beschikbaarheid van een zelfstudie, en c) een 'site visit' waarbij met tal van groepen personen gesproken wordt.

7: De verwachting van de grote proceskracht van het protocol

Het protocolbeginsel maakt deel uit van de reeks 'beginselen van behoorlijk visiteren'. Het van start gaan van een visitatie op basis van een door de te visiteren organisatie aanvaard protocol bevordert het succes van de visitatie.

De verwachting van de grote proceskracht van het protocol in een visitatiebestel houdt in dat een vooraf vastgesteld en gedragen koersdocument afspraken bevat over werkwijzen en over een beoordelingskader, die gedurende het gehele traject van visiteren een houvast vormen, dus een ankerpunt zijn voor de visitatiecommissie, de gevisiteerde organisatie en anderen. Een dergelijk ankerpunt voorkomt discussies en afwijkingen.

Toelichting: Er was in de periode 1985-2003 bij visitaties in het wetenschappelijk onderwijs sprake van een '*bij de les blijven-effect*'. Dat effect heeft betrekking op het feit dat de visitatiecommissies de handen vol hadden aan de uitvoering van hun taak. Wat de taak was, was gegeven in het (basis)protocol dat vooraf in VSNU-verband besproken en vastgesteld was. Het protocol bevatte de fasering van het visitatieproces, de werkwijze van de commissie en het kader voor beoordeling. In het kader van verwachtingenmanagement is dat belangrijk.

De commissies traden over het algemeen dan ook niet buiten het protocol. Toen dat een keer rond de visitatie van w.o.-opleidingen bedrijfskunde (in de eerste visitatieronde) wel gebeurde, door de opleidingen ook te beoordelen in het licht van de instellingsambities en de strategische koers van de instellingen, was de inspectie kritisch en afhoudend. Dat viel namelijk niet binnen het protocol. De visitatiecommissie had zich teveel vrijmoedigheid gepermitteerd volgens de inspectie die toezicht hield op visitaties. Ergo, de visitatiecommissie werd gepusht om binnen het protocol te blijven en geen 'uitstapjes' in de vorm van additionele analyses of vergezichten te plegen.

We nemen aan dat deze de verwachting van de grote kracht van een protocol, dat een verdere procedure kadert, zich ook buiten het wetenschappelijk onderwijs kan voordoen als meerdere

organisaties aan afzonderlijke visitaties deelnemen maar dat moet worden onderzocht. Overigens kent protocollering ook een schaduwkant: de visitatie-illusie. Zie verderop.

8: De verwachting van het immer te beperkte integraal kader voor beoordeling

Bij visitatie in het wetenschappelijk onderwijs deed zich het verschijnsel voor dat de wereld van het onderwijs hier en daar misschien iets mooier wordt afgeschilderd dan die in werkelijkheid is, aldus Leeuw (2003: 19). Opleidingsdecanen neigen ertoe een positiever beeld te hebben en te scheppen van de eigen opleiding dan strikt genomen gerechtvaardigd is. Dit verschijnsel kwam ook elders voor. Van alle 90 directeuren van naftakrakers waren er in het begin van deze eeuw 89 van mening dat ze in de (positieve) top tien zouden staan, stelt Leeuw. Dat is natuurlijk onmogelijk. Daarom zijn goede beoordelings- of toetsingskaders ook nodig, meent Leeuw (2003).

Het is wel mogelijk over een integraal beoordelingskader te beschikken dat weinig kans laat op 'positievere inkleuring dan strikt genomen gerechtvaardigd is' maar kan dat integraal kader de kwaliteit van het onderwijs geheel vangen of moet er steeds weer een normering en meting aan worden toegevoegd? De verwachting van het steeds te beperkte integraal kader houdt in dat het vooraf bepalen dat de organisatie of een deel (bijv. een opleiding) integraal doorgelicht wordt niet alleen leidt tot de keuze van één werkwijze en beoordelingskader, maar na zekere tijd ook tot een neiging bij critici om meer in dat beoordelingskader te willen stoppen (verbreding, verdieping, enz.). De neiging viel in het wetenschappelijk onderwijs duidelijk te onderkennen. Een soort *'rupsje-nooit-genoeg'-effect* waardoor het integraal kader dreigt te groeien en groeien. Omdat er echter afspraken gemaakt zijn over de stabiliteit van de werkwijze en het uniforme beoordelingskader wordt gaande de doorlichting van opeenvolgende organisaties echter geen wijziging aangebracht. De kritiek blijft daardoor bestaan.

Toelichting: Wie kijkt naar de externe kwaliteitszorg waarin verantwoording een doel is, in het w.o. of elders, ziet dat altijd een kader nodig is voor beoordeling van een reeks opleidingen op eenzelfde vakgebied en de prestaties daarvan. Dat kader is ook in het wetenschappelijk onderwijs ontstaan en na een pilot in het protocol gevat. Over dat kader was sprake van veel overeenstemming in VSNU-verband en in de disciplineoverlegorganen.

Bij het ontwikkelen van een kader voor de 'evaluatie' van een opleiding ligt het voor de hand om oog te hebben voor onder meer de doelstellingen waarop de opleiding mikt en de realisatie daarvan door middel van de inhoudelijke programmering met vakken en per vak, de proeven en examens, de studieorganisatie, de doorstroming, de uitval en het rendement, de arbeidsmarktperspectieven en de kwaliteit van docenten. Bij het kader waarmee een opleiding in het wetenschappelijk onderwijs bezien wordt, in feite de focus van de visitatiecommissie, zijn echter altijd wel enige kanttekeningen te plaatsen. Die blijken vaak na de eerste of tweede visitatieronde. Dan wordt gesproken over de wenselijkheid van het uitdiepen van een onderdeel, of over het voorkomen van afvinkgedrag en het in beeld krijgen van de gehele opleiding en de samenhang tussen onderdelen, of over meer genuanceerde beoordelingen van commissies die verder gaat dan een indeling in 'voldoende' en 'onvoldoende'. Wat we hier willen benadrukken, is dat het altijd te beperkt geachte kwaliteitskader voor beoordeling een soort wetmatigheid lijkt omdat elke focus op kwaliteit van een te beoordelen object andere foci uitsluit. Dus ontstaat er de voortdurende neiging tot het verbreden van het integraal kader. Zo gezien, is visiteren *'immer eine Unvollendete'*. Het is alsof de patiënt in het ziekenhuis komt en steeds maar weer naar meer en verdergaand onderzoek vraagt. Zo gezien, genereert visiteren altijd kritiek op de integraliteit.

De vraag is of we dit ook bij andere visitaties in andere sectoren terugzien. Blijken zal van wel. Zo vond bij voorbeeld in de tweede ronde van de visitatie van gemeentebesturen van in Limburg gelegen gemeenten een financieel verdiepingsonderzoek plaats. Deze analyse, die niet werd uitgevoerd door de visitatiecommissie-Rutten was een aanvulling op de eigenlijke bestuurskrachtmeting van de visitatiecommissie (Schutgens, Maessen en Korsten, 2009).

9: De verwachting van de visitatie-illusie

De verwachting van de visitatie-illusie impliceert dat visiteren in het wetenschappelijk onderwijs altijd in een dwangbuis van beperkte tijd plaatsvindt waardoor de kans stijgt dat een gefragmenteerd beeld van een opleiding ontstaat als gevolg van 'checklist-denken' en geen (voldoende) samenhangend beeld naar voren komt. Dat geeft het risico dat het visitatiestelsel niet de essentie van het onderwijs raakt (Van Berkel, 2000; Scheele, 2003: 71; Leeuw, 2003: 19). Voor visitatiecommissies dus een opdracht om aan dit gevaar van werken met af te vinken checklists te ontsnappen.

Toelichting: De term *visitatie-illusie* is gemunt door Van Berkel (2000) die zelf lid was van een aantal visitatiecommissies. De *visitatie-illusie* impliceert het bestaan van een *fragmentatie-effect*. Daarmee wordt bedoeld dat visitaties in het w.o. zich vooral richtten op onderdelen van onderwijs, zoals bij voorbeeld op de onderwijsprogrammering, de vakinhoud, de doorstroming, het rendement van de opleiding. Door de visitaties ontstaat mogelijk niet een geheel integraal maar een gefragmenteerd beeld. Hoe bevalt de opleiding als geheel? Is de opleiding studeerbaar? Dat blijft mogelijk buiten beeld. De interesse wordt volgens de *visitatie-illusie* als het ware naar veel afzonderlijke punten getrokken, naar details gezogen, naar punten uit een checklist. Als het de opleiding lukt om de commissie op gevraagde onderdelen tevreden te stellen kan het gevolg zijn dat betrokkenen menen dat het onderwijs als geheel ook op orde is. Dat is niet vanzelfsprekend het geval. De integrale op samenhang gerichte blik is nodig omdat er altijd, in alles wat gevisiteerd wordt, wel zwakheden blijken. Wie als tuinman elk boompje of elke heester goed verzorgt, heeft nog niet vanzelfsprekend zicht op de kwaliteit van de tuin; het is ook nodig te letten op de bouw van de tuin en de tuinidentiteit ('tuinarchitectuur'). Anders gesteld, wie de kwaliteit van een tuin wil beoordelen, moet niet alleen kijken naar de kwaliteit van elke boom of heester. Het is de *visitatie-illusie* dat dit geheel niet in beeld komt.

Van Berkel richt zijn pijl dus op checklists, op verdingelijking, dus – anders gesteld - op fragmentatie. Hij meent dat checklists ertoe kunnen leiden dat commissieleden op verschillende plaatsen zelfde vragen stellen en dat ook de informatie zo volledig mogelijk is om tot een goed rapport te komen maar van de andere kant kunnen checklists geen greep geven op en inzicht verschaffen in ingewikkelde onderwijsvraagstukken, in de integratieve kwaliteiten van een opleiding.

Er is ook nog een ander aspect. Door het hanteren van checklists kan de aandacht ook nog eens verschuiven naar processen ('het is niet op orde, maar we werken eraan') in plaats van op inhouden. En de checklists vangen het totaal en de samenhang mogelijk niet.

Het gaat hier om een dilemma, de spanning tussen geldigheid en betrouwbaarheid. Wie geldigheid wil bevorderen zal instructies geven om uiteenlopende interpretatie van wat je wilt waarnemen te voorkomen (standaardisatie, interpretatie-uitleg) maar een strakke taakomschrijving voor beoordelaars in de vorm van 'bullits' beperkt de beoordelingsvrijheid en verhindert het stellen van bepaalde vragen, van open vragen. Voor Van Berkel is dan ook een schrikbeeld om de kwaliteitseisen verder uit te werken, een advies dat de Algemene Rekenkamer (2000, 2008a,b), welhaast de uitvinder van ver(der)gaande protocollering, gaf.

De vraag is of visitaties in andere branches, zoals ziekenhuizen of woningcorporaties of politiekorpsen, bij beoordelingen van de organisatie het geheel en de samenhang vangen of vervallen in fragmentatieve oordelen over afzonderlijke zaken en details en checklist-denken.

10: De verwachting van strategisch gedrag bij visiteren

Strategisch gedrag van gevisiteerden komt voor, hoewel het vaak niet of moeilijk te bewijzen is en visitatiecommissies strategisch gedrag willen voorkomen of dempen. In het algemeen zijn strategische manoeuvres om gunstiger uit een visitatie te komen dan de waarheid toelaat of rechtvaardigt ongewenst omdat het de kwaliteit van een visitatierapport en de geloofwaardigheid kan ondermijnen (Schillemans, 2009: 163).

Eerder is aangegeven dat protocollering nuttig is omdat deze lijn brengt in het visitatietraject. Maar protocollering kan niet in alles voorzien. Er wordt namelijk wel gesteld dat het 'de oorzaken van strategisch gedrag niet wegneemt' (Ehren, 2007: 201). Logisch eigenlijk want opleidingsdecanen willen goed uit visitaties komen. Een visitatierapport met negatieve oordelen zorgt voor slechte 'public relations' en demotivatie van docenten.

Ondanks de protocollering moeten we dus rekening blijven houden met 'moves' aan de kant van de te visiteren opleiding om goed uit de visitatie te komen. De verwachting van strategisch gedrag bij visiteren (in het wetenschappelijk onderwijs) houdt in dat betrokkenen aan de kant van de gevisiteerde 'partij' strategieën kunnen en veelal zullen hanteren om 'goed' uit de visitatie te komen; dus om een goed rapport te krijgen (Leeuw, 2003). Strategisch gerichte pogingen kunnen tot uiting komen in de zelfstudie, in aantal en type te verstrekken documenten (hoeveelheid, wat wordt achtergehouden), op wat tijdens een 'site visit' gezegd kan worden of vermeden, en op inbreng bij een conceptrapportage (aanvechten van bepaalde passages, enz.). Kansen op voor de geldigheid en betrouwbaarheid van de visitatie negatief strategisch gedrag kunnen over het algemeen worden beperkt door protocollering, een alerte commissie en openbaarheid van het eindrapport.

De redenering om tot deze verwachting te komen, luidt als volgt. Op visitaties zijn meerdere perspectieven mogelijk. De kwaliteit is te beoordelen vanuit het beeld dat visiteren een vorm van onderzoek is. Dat perspectief leidt tot een vraag of het onderzoek geldig en betrouwbaar is. Een ander, daarmee samenhangend beeld houdt in dat men visiteren kan zien in termen van strijd en waarheidsbeïnvloeding. Gaat men uit van spel, strijd en waarheidsbeïnvloeding, dan dient zich het fenomeen van strategisch gedrag aan (Laman, 2003; De Bruijn, 2008). Dat houdt in dat een gevisiteerde organisatie kan proberen het procesverloop, de informatievoorziening, de gespreksonderwerpen en de conclusies van een visitatiecommissie te beïnvloeden om een ander beeld in het visitatierapport te schetsen dan het geval zou zijn bij het wegvallen van dat gedrag. In dat verband is het *hoera-effect* aan de orde. Daaronder wordt verstaan het streven om - door middel van handigheden of quasi-onhandigheden - gunstig uit de visitatie te komen.

Protocollering verhindert strategisch gedrag? Grotendeels omdat omschreven is wat in een zelfstudie moet worden behandeld. Maar vastgesteld is ook dat de 'site visits' aan universiteiten zich meer leenden voor strategische manoeuvres dan de zelfstudies. Ook hier is wel sprake van een strakke agenda maar gesprekspartners kunnen hun eigen woorden kiezen.

Hoe komt men in de hoerasfeer van een positieve visitatie terecht? Door die docenten te selecteren die een cursus of proef degelijk kunnen toelichten en door de zwakkere (onhandige enz.) buiten beeld te houden. Het *evaluatiestroomeffect* is een ander serieus voorbeeld van strategisch gedrag: overdonder de commissie met informatie zodat je als decaan de eigen interpretatie van een feit of reeks feiten ingang kunt doen vinden. Weer een andere strategie is

het *lokale omstandighedeneffect*: benadruk dat deze opleiding anders is en daarom heel bijzonder; commissie: hou daar rekening mee! Een volgende optionele aanpak betreft het *strategisch beleidseffect*: benadruk dat de opleiding gaat veranderen vanuit een nieuw strategisch beleid; de toekomst wordt anders. Als de commissie het daarover bij het bezoek ter plaatse langdurig heeft, kan ze geen vervelende vragen stellen over de 'past performance' en kan de opleiding benadrukken dat de toekomst het volgende heden is. Een volgende strategie die een hoera-effect beoogt, is het *meebuigeffect*. Denk als decaan met de commissie mee vanuit de verwachting dat de commissie altijd iets moet zoeken en vinden. Er is bovendien altijd kritiek mogelijk. Benadruk dus zwakheden, maar kies daarvoor zwakheden of risico's 'waarvoor anderen dan de decaan verantwoordelijk zijn' ('het is misschien wel juist, maar u moet daarvoor bij het college van bestuur zijn, niet bij mij'). Een volgende strategie is de retorische: kies 'zachte woorden voor harde feiten' als de kans zich voordoet. Ga geen 'gezwel wegsnijden'. 'Framing' doet ertoe (De Bruijn, 2011). Vermijd medische metaforen over ongezond gedrag. Denk eerder aan het volgende: een dreigende blunder is 'een risico op kwetsbaarheid'. En denk in termen van risico's die je klein wilt houden door proceduremaatregelen, zoals het regelmatig houden van een op preventie gerichte quick scan (de opleiding kent een web aan 'voorzorgsmaatregelen'). Tenslotte: denk aan het *orchestratie-effect*: het influisteren van de personen die bij de commissie hun opwachting maken.

Kunnen we specificeren wat matig presterende organisaties aan gedrag vertonen bij 'public reporting' over prestaties of is dat niet mogelijk? We komen daarmee op het terrein van *'The Blame Game'* (Hood, 2011). Gebleken is dat *'lower-performing public agencies'* meer hun best doen om het beleid in zelfevaluaties te rechtvaardigen: om zaken goed te praten en om 'vuiltjes weg te werken'. Er treedt dus meer gedrag op om niet voor schut gezet te worden. Er is - anders gesteld - sprake van ontwijkgedrag, van pogingen om zich vrij te pleiten, of van inzet van talige voorbehoedmiddelen (*het 'blame avoidance'-effect*). Goed presterende organisaties voeren juist minder rechtvaardigingen ('justifications') voor hun handelen op. Ze hebben het als het ware het niet nodig. De feiten spreken hier voor zich (Charbonneau & Bellavance, 2012).

Voor 'blame-avoidance' kunnen de volgende strategieën ingezet worden:

- a. presentatiestrategieën: 'spin your way out of trouble';
- b. de strategie om een zwart schaap te vinden; een ander 'van ver weg' krijgt de schuld ('zoek een externe crisis');
- c. de strategie om geen oordelen te geven waardoor verliezers aangewezen kunnen worden; dus het juist vermijden dat iemand de schuld krijgt ('het raakt ons allen').

Presentatiestrategieën lijken een onvermijdelijk verschijnsel als een 'self report' moet worden gemaakt. De volgende strategieën komen voor:

- a. de keuze van een laag profiel (dus: 'hou je koest');
- b. verspring van onderwerp: schenk niet veel aandacht aan wat zwak is en juist meer aan het neutrale of positieve ('we klampen ons vast aan wat goed is');
- c. kies een krachtige argumentatie: bied excuses aan en toon dus erkenning van een zwakte ('kwetsbaarheid tonen werkt...') en zeg dat het anders gaat worden (goede bedoelingenpolitiek);
- d. trek een grens: voorbij dat punt accepteren wij iets zelf niet (zie Hood, 2007b: 201; Hood, 2011: 18).

Maar het is in een spel zo dat er minimaal een en meestal meerdere partijen zijn die de strategische manoeuvres van de ene partij trachten te verhinderen, ongedaan te maken of te neutraliseren.

Is er bewijs gevonden voor het voorkomen van strategieën en zo ja voor welke strategieën? Echt bewijs van substantieel vertekende visitierapportages in het wetenschappelijk onderwijs is in de literatuur niet te vinden. Maar uit gesprekken met betrokken decanen en hoogleraren uit opleidingen economie, milieukunde, bedrijfs- en bestuurskunde is wel bekend dat degenen die tijdens een visitatie bij een 'site visit' aantreden wel bij een of meer oefensessies betrokken waren. Soms krijgen studenten ook tips aangereikt. En tal van visitatiecommissies erkennen dat als ze vragen om afstudeerscripties en ze zeggen niet dat ze een spreiding in de punten willen, ze vooral scripties onder ogen krijgen van studenten die een cijfer zeven, acht of meer daarvoor kregen. Een evaluatiestroomeffect is evident waarneembaar: de neiging om spontaan dan wel gevraagd tal van evaluatiedocumenten te overleggen over bijna alles wat rond opleidingen zich als evalueerbaar aandient. Daar gaat de hoop achter schuil dat de commissie gaat denken dat dit of dat weliswaar 'niet sterk is' maar men houdt in elk geval de vinger aan de pols. Opleidingsdecanen gokken dan op ontvankelijkheid bij de commissie voor proceskwaliteit. Een commissie die oog heeft voor strategisch gedrag zal echter vragen naar zwakkere en sterkere scripties en kijken naar de relevantie en kwaliteit van de evaluaties.

Wie niet alleen visitatie in het wetenschappelijk onderwijs beziet maar ook bijvoorbeeld visiteren van gemeentebesturen kan vanuit de congruentiegedachte - verderop aan de orde - nagaan of gemeentebesturen ook strategisch gedrag vertonen om een gunstige uitkomst te beïnvloeden. Ze kunnen zich voordoen, maar ze zijn ook lastig aantoonbaar omdat niemand belang heeft bij vermelding.

11: De verwachting van het inburgeringseffect

De verwachting van het inburgeringseffect houdt in dat als organisaties of opleidingen aan een tweede of volgende visitatie deelnemen, de visitatie routinematig gaat worden en verwordt tot een 'af te handelen zaak'. De druk neemt daardoor af om zich als bestuurders en ambtenaren te prepareren en naderhand ook werk te maken van verbeteracties. Visitaties dragen dan niet meer bij aan een cultuur van kwaliteitszorg en/of innovatie in het onderwijs, als ze dat eerder al deden.

Toelichting: Als een opleiding voor de tweede of derde keer gevisiteerd wordt, dreigt een 'inburgeringseffect'. De spelregels zijn bekend, het opstellen van een zelfstudie wordt 'een klusje' dat met veel gevoel voor taalnuances tegemoet getreden moet worden en de betrokkenen kennen het klappen van de zweep bij de ontvangst van de commissie ('bloemen op tafel' en 'goed in het pak'). Er ontstaat een groep opstellers van zelfstudies die dit werk vaker hebben gedaan en de formuleringen langzamerhand kennen.

Op grond hiervan is de verwachting dat als in een andere sector een tweede of derde (opeenvolgende) visitatie veel gelijkenis vertoont met de gang van zaken in het wetenschappelijk onderwijs hetzelfde effect optreedt: routinematige gedragspatronen om - als het even kan - gunstig uit een visitatie te komen. Maar ook kan zich voordoen dat de kans op slordigheden tijdens het visitatieproces toeneemt.

Deze verwachting heeft een verbinding met de verwachting dat na zekere tijd sleet op het visiteren kan komen. Visitatie in het onderwijs zou zo gezien een instrument zijn dat slijtage vertoont. Betrokkenen bij onderwijs zijn zich zeker na eerste visitaties bewust van mogelijkheden om hier en daar strategisch gedrag bij de opstelling van een zelfstudie aan de dag te leggen.

12: De verwachting van de benuttingsimpuls

De verwachting van de benuttingsimpuls houdt in dat belangrijke bevindingen van visitatierapporten benut worden ('performance information use') door de doelgroep van gevisiteerde organisaties, met name door opleidingsdecanen.

Op de bevindingen uit verplichte visitaties volgt eigenlijk altijd bezinning en actie, terwijl vrijwillige visitaties 'soms geen gevolg krijgen' (Schillemans, 2009: 167; Schillemans, 2007: 230).

De benutting zal bij verplichte visitaties, zoals in het w.o., van het type zijn: direct, instrumenteel en op korte termijn (binnen een jaar na het uitkomen van het visitatierapport). Dus niet: indirect, conceptueel en op de langere termijn. Gevisiteerde organisaties zullen als gevolg van die benutting zelfreinigend vermogen tonen en daarmee 'lerende organisaties' zijn.

De geformuleerde verwachting past in dit perspectief, dat krachtig aanwezig is in de 'performance'-literatuur, aldus Swierstra en Tonkens (2008: 121-122), Van Dooren en Van de Walle (2008) en Moynihan (2008).

Een belangrijk tweede argument pro reflectie op benutting ontleen we aan de ervaringen met visitaties in het wetenschappelijk onderwijs omdat de benuttingscondities daar, gezien vanuit de kennis die de '*utilization research*' heeft opgeleverd over benutting of gebruik, in de periode van onderzoek (tamelijk) gunstig waren. De visitatierapporten waren over het algemeen leesbaar, relevant, concluderend, overtuigend, en de ontvangers waren ontvankelijk voor de inhoud; de externe druk van een meekijkende onderwijsinspectie zorgde voor de noodzaak om te verbeteren en om met name de 'onvoldoendes' uit het visitatierapport weg te werken. Visitatierapporten over opleidingen uit het w.o. spreken zich uit over wat 'voldoende' of 'onvoldoende' is en geven duidelijk richting aan waar de verbetering gezocht moet worden.

Er is eerder de verwachting uitgesproken dat sprake is van concluderend gehalte van visitatierapporten (verwachting nr. 5) maar ook van uitblijvende aanbevelingen (verwachting nr. 6). Desondanks en ondanks het inburgeringseffect (nr. 11) zijn we van mening dat de voorwaarden voor benutting volop aanwezig zijn. Wat we stelden over benutting van visitatierapporten in het w.o. ondersteunt dus deze verwachting. Van de andere kant is deze verwachting niet eenvoudig te analyseren. In par. 10 gaan we er daarom nader op in.

13: De verwachting van gezwollenheid

Er is ook een verwachting mogelijk die enigszins contrair is aan de verwachting van de benuttingsimpuls. De verwachting van gezwollenheid impliceert dat achter visiteren meer ambities schuilgaan dan waargemaakt worden: een organisatie zegt zich, door aan een visitatie deel te nemen, te willen verantwoorden en te willen leren maar het organisationeel leren is zwak of niet institutioneel in de eigen organisatie verankerd doordat een dagelijks bestuur zich niet verplicht weet om verbeterinitiatieven te ontplooiën (over leren: Argyris & Schön, 1978; Swieringa en Wierdsma, 1990; Casey, 1993; Argyris, 1996; Huysman en Van der Vlist, 1998; Meurs e.a., 2006; Caniëls, 2011). Achter de façade van geproclameerde ambitie tot deelname aan de visitatie gaat geen echte verbeterpraktijk schuil; men is tevreden als men 'goed' uit de visitatie komt en verder niet in actie hoeft te komen (Scheele, 2003: 71). Daardoor is een - al of niet openbaar - rapport of verslag over wat na de visitatie is verbeterd voor onderzoekers doorgaans onvindbaar en een verbetering die toe te schrijven is aan een visitatierapport niet door centrale actoren te benoemen. Het gaat hier om een *plofkipverwachting*: de (plof)kip is minder goed dan die lijkt.

De *prestatieparadox* sluit hierop aan: organisaties kunnen wel willen meewerken aan prestatie meting maar zijn daarmee nog niet onmiddellijk ook effectieve organisaties (Van Thiel en Leeuw, 2003).

Toelichting: Deze verwachting heeft te maken met een combinatie van a) een formele ambitie (op papier altijd de lat wat hoger leggen dan de praktijk) én b) een zekere neiging tot zelfgenoegzaamheid bij bestuurders en anderen ('wat er ook gezegd wordt: we doen het voldoende tot goed') en c) de tolerantie voor enige suboptimaliteit ('och, er is altijd wel iets dat beter kan').

De gezwollenheid bestaat hierin dat de leerpraktijken niet volgen op de (opgeblazen) leerambities. Als er geen evidente 'onvoldoendes' gerapporteerd zijn en een commissie betrekkelijk vaag is in de beoordeling verdwijnt het visitatierapport in de kast of volgt slechts een rituele agendering (wellicht wordt een rapport aangeboden 'ter kennisname').

In politieke contexten oefent een politieke representatief algemeen bestuur zelden druk uit om te leren. Overheidsbesturen munten niet uit in zelfreiniging. Als een dagelijks bestuur geen consequenties formuleert bij een toegezonden visitatierapport nemen fracties, bij voorbeeld uit een gemeenteraad, in meerderheid een rapport voor kennisgeving aan. Gemeenteraden munten niet uit in zelfreinigend vermogen (Korsten en Schoenmaker, 2011; Schoenmaker, 2011; Abma, 2012).

14: De verwachting van niet collectief verbeteren ('het op zichzelf-effect')

Deze verwachting van niet collectief verbeteren maar op zichzelf betrokken blijven, handelt ook over benutting van visitatierapporten. De verwachting houdt in dat als er verbeterd moet worden, verantwoordelijke decanen deze verantwoordelijkheid zelf nemen en er niet toe neigen om in het vakgebied in disciplineoverlegorganen samenwerking met decanen van andere universiteiten te zoeken. Dit noemen we de 'verwachting van het niet collectief verbeteren' of het 'op zichzelf-effect'.

Toelichting: In het wetenschappelijk onderwijs zijn door instellingen in de periode 1985-2003 – voor zover bekend - weinig pogingen ondernomen om op basis van een visitatierapport collectief tot een verbetering op een vakgebied te komen. Vergaderverslagen van discipline-overlegorganen op het gebied van bestuurswetenschappen en bedrijfskunde en raadpleging van enkele voormalige leden die participeerden vanaf 1985-2003 (de hoogleraren Korsten, Heemstra, Van den Bosch) laten dat zien. Blijkbaar waren de omstandigheden in de eigen instelling zodanig specifiek dat een decaan op zichzelf bleef werken en waren de 'problemen' in de opleiding ook met in de eigen instelling aanwezige expertise weg te werken. Bovendien willen decanen hun verantwoordelijkheid niet laten afhangen van of beïnvloeden door andere decanen.

Vanuit de congruentieverwachting is aannemelijk dat bij visitaties in een andere sector dan het w.o. ook sprake zal zijn van een 'op zichzelf-effect'. Zo zullen gevisiteerde gemeenten voor het corrigeren van imperfecties of zwakten mogelijk niet direct de samenwerking zoeken met buurgemeenten of andere gemeenten, en zullen politiekorpsen dat niet doen met andere politiekorpsen.

15: De verwachting van slijtage in de externe kwaliteitszorg en de roep om evolutie

We komen bij de verwachting van slijtage in de externe kwaliteitszorg ('het inzakken') en de roep om evolutie. Na enkele rondes van visiteren, kunnen betrokkenen met zekere waarschijnlijkheid bovendien voorspellen wat sterk en zwak is in de eigen opleiding en hoe de commissie ernaar zal kijken. Dan nog kan visitatie nut hebben omdat de decaan en anderen zelf niet goed de feiten onder ogen durven zien en zwakten te 'zacht' inschatten. Dit fenomeen zal zich ook elders voordoen. Daarom heet het een verwachting, lees verwachte bevinding.

Toelichting: Beleid slijt en ook bij systemen van externe kwaliteitszorg blijkt na zekere tijd sleet; dan komen zwakten naar voren. In een visitatiesysteem waarin zelfde opleidingen van verschillende universiteiten vanuit een uniform beoordelingskader doorgelicht worden, bestaat weinig tot geen oog voor de eigenstandige kwaliteiten van elk van de opleidingen en evenmin voor de strategische koers van het college van bestuur en/of het faculteitsbestuur. Elke opleiding wordt door een dezelfde mal geperst (het uniforme protocol met uniform beoordelingskader).

Een voorbeeld. Wie vanuit het onderwijskundig concept van competentiegericht leren veel accent legt op individuele keuzen die leerlingen of studenten in het onderwijs moeten maken, krijgt na een aanvankelijk beloftevolle periode, waarin ongunstige effecten van het concept nog onbekend zijn, te maken met meer serieuze kritiek: van docenten die zich ongemakkelijk voelen omdat hun vakmatige expertise niet goed meer wordt benut en van cognitieve hiaten bij studenten die gecorrigeerd moeten worden via een klassikale aanpak van kennisoverdracht. Heeft de interne kwaliteitszorg dit tijdig gesignaleerd? Vaak in het begin niet.

In de uniforme aanpak van visiteren wordt geen aandacht besteed aan pluriformiteit en wordt geen voldoende distantie genomen tot tekortkomingen die elk onderwijskundig concept in zich draagt. De kwaliteitszorg is gericht op optimalisering binnen onderwijskundige systemen en *niet* op detectie van grondfouten of zwakten van onderwijskundige aanpakken. Dat achter een opleiding een eigen strategisch beleid schuilgaat, wordt in een te verrichten min of meer uniforme vakmatige visitatie niet goed belicht omdat er 'geen plaats voor is'. Dat is te ondervangen door een instelling zelf de keuze te laten om een eigen externe visitatie te laten verrichten waarbij wel naar eigen doelstellingen gekeken wordt.

En wat ook niet meer passend te krijgen is, is de puzzel dat als een land veel instellingen en opleidingen kent, het wel heel ver gaat om alle zelfde studierichtingen gelijktijdig te visiteren en alle opleidingen eens per vier-vijf jaar te beoordelen. Dat wordt dan immers een heel arbeidsintensieve en dure werkwijze. Dan gaat een hoger aggregatieniveau gezocht worden zoals de faculteit of instelling, met de optie van een instellingsvisitatie. Dan evolueert het visitatieconcept en de praktijk van visitatie.

Doet zich dit alleen voor bij visiteren in het wetenschappelijk onderwijs? Neen. Bij elk kwaliteitszorgsysteem is sprake van generaties die elkaar opvolgen. Er openbaart zich 'na verloop van tijd een geheel van onderliggende assumpties, overtuigingen en probleempercepties waarvoor op productieve wijze een uitweg gevonden moet worden' (Huitema, Jeliaskova en Westerheijden, 2003: 33). Dat betekent vanuit het congruentiedenken redenerend dat als een visitatiesysteem uit het wetenschappelijk onderwijs gekopieerd wordt naar een andere sector ook daar na zekere tijd beperkingen en slijtage blijken. Ergo, visiteren in de gemeentelijke sfeer zal na zekere tijd ook een evolutie doormaken of misschien zelfs verdwijnen.

16: De verwachting van de begrensde uitdaging van visitatieobjecten

De verwachting van de begrensde uitdaging houdt in dat elk visitatiebestel na zekere tijd op systeemimmanente beperkingen stoot. Dan komen nieuwe vragen op (zoals over de kwaliteit van het onderwijsconcept of de basisaanpak in de ziekenzorg of de keuzen in de politiezorg) die het bestaande externe kwaliteitszorgsysteem niet kan beantwoorden. Vermoedelijk geldt die verwachting ook in andere sectoren dan het wetenschappelijk onderwijs.

Toelichting: In het wetenschappelijk onderwijs is vastgesteld dat visiteren van opleidingen leidt tot kwaliteitsoordelen en aanbevelingen binnen 'bestaande kaders', binnen een bestaand paradigma. Visiteren zou opleidingscoördinatoren en andere betrokkenen in het

wetenschappelijk onderwijs niet echt uitdagen. Visiteren zou te gesloten zijn. Het gaat hier om stelselkritiek, die is verwoord door inspecteur-generaal Mertens. We hebben die kritiek gebracht als de kritiek van de 'begrensde uitdaging'. De kern van de kritiek is: doen instellingsbesturen en decanen wel de goede dingen in plaats van een analyse van of ze de dingen goed doen? Die vraag komt nauwelijks in visitaties in het wetenschappelijk onderwijs aan bod. Wordt dit opgemerkt dan zit visiteren aan de vooravond van een evolutie naar een verandering.

In het wetenschappelijk onderwijs trad die verandering ook op: de overgang naar accreditatie (Douma, 2004). Als gevolg daarvan blijft visiteren wel betekenis houden, maar in een nieuwe setting, beleidsopvolging te noemen.

Uitgaande van de congruentieverwachting mogen we aannemen dat als in een andere sector ook visitatie plaatsvindt, bij voorbeeld van gemeentebesturen of de uitwerking van het protestedenbeleid door stadsbesturen, na verloop van tijd gezegd gaat worden: bij visiteren van gemeentebesturen worden deze gemeentebesturen te weinig uitgedaagd. Er wordt dan slechts nagegaan of de gemeentebesturen hun beleid goed uitvoeren; niet nagegaan wordt of gemeentebesturen de goede dingen doen.

17: De verwachting van de begrensde verantwoording

Een van de beginselen van behoorlijk visiteren is: openbaarheid van de bevindingen van de visitatiecommissie. Dit beginsel komt uit het w.o. In het w.o. was sprake van een openbare verantwoording, niet slechts naar collega-instellingen of de onderwijsinspectie. Maar hoe ver gaat die openbaarheid? In hoeverre verantwoorden organisaties die meedoen aan een visitatie zich actief of passief? Deze vraag voert, op grond van onderzoek van de kwaliteit van de externe kwaliteitszorg in het wetenschappelijk onderwijs, naar de verwachting van de begrensde verantwoording. Deze verwachting houdt in dat de verantwoording naar een visitatiecommissie volledig is conform protocol maar dat de samenleving in bepaalde opzichten het nakijken heeft. Zelfstudies zijn niet openbaar, en de terugkoppeling van bevindingen door een visitatiecommissie na een 'site visit' gaat niet vergezeld van persberichten, de aanwezigheid van journalisten of persconferenties.

Toelichting: Zelfstudies over opleidingen uit het w.o. zijn - voor zover bekend - vrij zelden door onderwijsinstellingen actief openbaar gemaakt. Decanen neigen er (uit concurrentieoogpunt) toe die zelfstudies slechts in handen te stellen van een visitatiecommissie maar de rapporten waren na enige aandrang meestal voor collega-decanen uit een disciplineoverlegorgaan wel te raadplegen ('uitruil'). Van actieve openbaarheid naar de samenleving was dus in dit opzicht geen sprake. Dat was ook *niet* de afspraak. Visitatierapporten van w.o.-opleidingen waren volop openbaar; de VSNU publiceerde ze allemaal, maar zelfstudies niet. Hoe is dat elders?

De visitatierapporten van ziekenhuizen zijn niet voor burgers openbaar en die van politiekorpsen, enkele uitzonderingen daargelaten (omdat een rapport op internet verscheen), ook niet. Veel visitatierapporten van woningcorporaties zijn daarentegen wel openbaar (Van Overmeeren, 2007).

Actieve fora die visitatierapporten kritisch belichten zijn zeldzaam. De Consumentenbond besteedt er geen aandacht aan. In dat opzicht is de openbaarheid ook begrensd. In de onderwijsarena is het de onderwijsinspectie die toeziet op de kwaliteit van rapportages.

Reflectie

Deze verwachtingen lenen zich voor verder onderzoek. Welke meegenomen worden in verder onderzoek valt nader te bezien. Een verwachting als die over strategisch gedrag voor en tijdens

visitatieprocessen is slechts te analyseren door grote betrokkenheid tijdens een visitatieproces. Omdat observatoren niet makkelijk toegelaten worden, is dit aspect in het w.o. ook nauwelijks onderzocht.

10 Benutting van onderzoeksbevindingen nader bezien

Benuttingsvraagstukken geven aanleiding 'tot veel hoofdbreken' (Korsten, 1983; Bekkers, 2007: 318; Hoppe, 2008: 16). Drie redenen zijn daarvoor te geven.

- 1) Benutting is niet eenvoudig zichtbaar ('low visibility'; Korsten, 1983: 28).
- 2) De zeer hoge verwachtingen over benutting dragen bij aan het beeld dat 'litte utilization takes place' (Korsten, 1983:28).
- 3) Er is wel eens gezegd dat benutting van onderzoeksbevindingen te omschrijven is als een kwestie van 'stukjes en beetjes'. Of anders gesteld: '*ambiguous, amorphous, incremental and meandering*' (Webber, 1991; Bekkers, 2007: 318).

We lopen aan tegen:

- de begripsomschrijving,
- wat te verwachten is aan benutting (normatieve modellen),
- de differentiatie naar typen benutting (o.a. de indeling in direct en indirect gebruik),
- hoe benutting te onderzoeken (we geven een aantal manieren, zoals de 'utilization research'),
- hoe veel of weinig benutting te verklaren (drie clusters verklaringen).

Begrippen

Een van de lastige onderwerpen is de terminologie. We gebruiken hier de termen benutting en gebruik door elkaar. We spreken niet:

van '*rendement*' (wel bij Van de Vall en Van Dijkum-De Jong, 1979);

van '*doorwerking*' (wel bij Bekkers e.a., 2004; Hupe en Van der Meer, 2006);

van '*cruciale koppeling*' (Hupe en Van der Meer, 2006);

en niet van '*invloed*' (Mark & Henry, 2004).

Rendement achten we een eenzijdig economische term. Doorwerking verwijst naar gebruik van in procedures ingebedde rapportages als milieu- en andere effectrapportages, wier bestaan niet ontkend mag worden. En doorwerking verwijst teveel naar het zonder meer overnemen van resultaten, wat ongewenst kan zijn omdat er in een politieke democratie altijd een politiek-bestuurlijke afweging nodig is en slecht onderzoek niet moet worden overgenomen. Ook de term '*cruciale koppeling*' of *kennistoepassing* negeren we omdat die in de literatuur geen gemeengoed zijn. Kennistoepassing is een te deterministisch concept dat wellicht meer past bij natuurwetenschappelijk onderzoek.

We komen in de literatuur ook het begrip '*kennistransfer*' tegen (Korsten, 1983: 28). We gebruiken dit begrip verder niet omdat het ons te doen is om de betekenis van visitatierapporten.

'Benutting' en 'gebruik' verwijzen ons inziens vooral naar bezinningstrajecten en laten toe te denken in *fasen* en *vormen* van benutting (Van de Vall, 1980; Korsten, 1983; Pröpper en De Vries, 1995; Dal Santo e.a., 2002; Landry e.a., 2003).

Is de benutting of het gebruik van onderzoeksbevindingen gering dan spreken we van *onderbenutting* of *ondergebruik*. Dat betekent dat we ervan uitgaan dat *maximalisering* van benutting of gebruik de norm is, althans zeker bij in opdracht verricht evaluatieonderzoek (cf Carter, 1994; Pröpper en De Vries, 1995). Daar zijn natuurlijk enkele kanttekeningen bij te plaatsen. Habermas (1987) heeft uitgelegd dat er drie vormen omgaan met onderzoek zijn: het

decisionistisch model, het technocratisch model en het pragmatisch dialoogmodel. In het technocratisch denken domineert het onderzoek de besluitvorming, in het decisionistische model domineren de besluitvormers en zijn onderzoekers en andere technocraten ondergeschikt terwijl in het dialoogmodel wisselwerking bestaat tussen wetenschap en beleid. In een politieke democratie dient een pragmatisch aanpak gericht op dialoog tussen onderzoekers en beleidsvoerders voorop te staan. Zo gezien, mag niet verondersteld worden dat onderzoek domineert en benutting van onderzoeksbevindingen op voorhand maximaal zou moeten zijn. Immers, dan is men een adept van het decisionistische denken.

Heeft het zin om het begrip 'invloed' te gebruiken in plaats van gebruik of benutting? Wie de specifieke literatuur over evaluaties beziet, komt tot de conclusie dat het gaat om benutting van het (evaluatie)resultaat. Sommigen stellen evenwel een heroverweging van het begrip 'benutting' voor (Kirkhart, 2000; Henry & Mark, 2003). Karin Kirkhart stelt:

'The term influence (the capacity or power of persons or things to produce effects on others by intangible or indirect means) is broader than use, creating a framework with which to examine effects that are multidirectional, incremental, unintentional, and instrumental (Kirkhart, 2000: 7).

Kirkhart meent dat het zoeken naar 'benutting' ('utilization' of 'use') te beperkt is. Henry & Mark (2003) treden in haar voetsporen. Het analyseren van de evaluatie-*invloed* zou de mogelijkheid bieden om te begrijpen waarom, wanneer en hoe het evaluatieresultaat benut wordt. Er moet volgens deze auteurs meer zicht ontstaan op processen en ketens die het wel of niet reflecteren op het evaluatieresultaat kanaliseren. 'Benutting' acht Kirkhart slechts een onderdeel van invloed als breder kader. Een evaluatierapport kan bij voorbeeld ook invloed hebben op opvattingen en houdingen van ambtenaren en/of bestuurders. Dat kun je dan niet echt tot 'benutting' rekenen maar dit proces heeft wel betekenis, zegt zij. Duidelijk is dat deze auteurs ook zien dat benutting te makkelijk leidt tot de waarneming van onderbenutting en daarmee leidt tot frustratie over de betekenis van beleidsevaluaties. Soms is er wel degelijk invloed en het is zaak daarop ook zicht te krijgen. De auteurs stellen dat benutting in het algemeen onvoldoende aandacht schenkt aan (intra)persoonlijke, inter-persoonlijke en maatschappelijke veranderingsprocessen (Henry & Mark, 2003: 297; Mark & Henry, 2004). Het is goed dit te beseffen maar we houden toch vast aan de termen 'benutting' en 'gebruik'.

Benutting van kennis: welke?

Waar moeten we aan denken bij benutting van onderzoek? We noemen een aantal mogelijkheden (zie ook Dunn, 2008; Derksen, 2011: 115-116).

1. Datagebruik: Onderzoek levert cijfers die een bestuurder kan gebruiken om beleid te onderbouwen. Bij voorbeeld cijfers over de 'leeftijd' van rioolbuizen in het stadscentrum en daarbuiten? Cijfers over de reinheid en mate van verontreiniging van oppervlaktewater.
2. Operationele kennis: Onder geeft de operationalisering van beleid. Welke mogelijkheden heeft de burgemeester om de vestiging van drie pedofielen in een bepaalde woonwijk tegen te gaan? Als de omvang van de hennepcultuur in de gemeente wordt vastgesteld, is het nodig om
3. Probleemanalyses: waar gaat het na agendering van een item en van melding van incidenten eigenlijk precies om? Wat is voor wie hoe en waarom een probleem? Is het probleem minuscule of groot?
4. Vergelijking van alternatieven: Als alternatief c gevolgd wordt heeft dat meer positieve of negatieve consequenties voor luchtverontreiniging dan in geval van alternatief b of a.
5. Kosten-batenanalyse van alternatieven: Hoe groot is de financiële meerwaarde als een weg wordt aangelegd via dbfm (PPS) in vergelijking met db?
6. Prognoses van groei of krimp: De toename van files....in jaar o in vergelijking met jaar k en l. De toename van drugsgebruik in een gemeente onderscheiden naar soorten drugs.

7. Het aanreiken van een aansprekend idee of beleidssuggestie: Bij voorbeeld over duurzaamheid of innovatie. 'Als Nederland nou eens een windmolenpark in zee zou bouwen.....'.
8. Evaluatieve (resultaat)kennis (o.a. monitoring en 'performance'): Wat is het normenkader voor de beoordeling van de kwaliteit van het gemeentebestuur? Wat is het draagvlak voor introductie van beleid a versus b? Is het doelmatiger om huisvuil zelf op te halen of huisvuilophaal uit te besteden?
9. Effectkennis: Wat is er terecht gekomen van beleidsprogramma a en wat is de oorzaak van het niet (in een bepaalde mate) bereiken van doelstellingen? Wat zijn de (neven)effecten van politiek k? Wat zijn de gevolgen van de introductie van buurtcoaches op? Heeft de introductie van maatregelen op het gebied van verkeersveiligheid geleid tot minder ongevallen van het type a, b en c?
10. Uitvoeringskennis: Hoe werken organisaties aan de reïntegratie van werklozen in het arbeidsproces? Is dat conform de verwachtingen?
11. Risico-analyses: Welke risico's lopen burgers in geval van het uitbreken van ziekte a en b? Of in geval van een overstroming?
12. Consequentionele kennis. Wat is de 'impact' van het op grote schaal niet achterhalen van de schuldigen van woninginbraken voor het oordeel over de politie? Wat zijn de ingeschatte gevolgen van de introductie van een wietpas op illegale drugshandel op straat? Wat leert simulatie van de introductie van de mestboekhouding op?
13. Optionele kennis: Wat zou de implicatie kunnen zijn van het uitrusten van alle handhavers met een iPad met gegevens over? Welke mogelijkheden biedt ICT voor gemeente-ambtenaren? Wat impliceert e-government voor de mogelijkheden op het vlak van dienstverlening?
14. Vergelijkende kennis: hoe waarderen burgers in verband met de dienstverlening de openingstijden van gemeenten? En wat vinden ze van de loketbehandeling? Wat leren benchmarkings?

Normatieve denkmodellen: wat mogen we verwachten aan 'benutting'?

Alvorens in te gaan op hoe je deze of andere benutting of ander gebruik van onderzoek of visitatierapporten zou kunnen onderzoeken, behandelen we kort enkele normatieve modellen over de relatie tussen onderzoek en beleid. Het gaat hier om hoe onderzoek benut moet of kan worden. Wat mag je verwachten aan benutting?

Die behandeling biedt de mogelijkheid om voor eenzijdigheden te waarschuwen. We noemen de volgende perspectieven of concepten:

het push-model; het demand-pull model, het verspreidingsmodel, het interactiemodel, het enlightenment-model, politiek-argumentatieve model (o.a. Weiss, 1977, 1979, 190; Weiss & Bucuvalas, 1980; Korsten, 1983; Bekkers, 2007; Dunn, 2008).

Allereerst is er het '*push-model*' (Korsten, 1983; Bekkers, 2007: 318). Volgens deze denkwijze, dus dit perspectief, is de kwaliteit van nieuwe ideeën en inzichten die uit onderzoek voortkomen de belangrijkste determinant voor de toepassing van kennis door beleidvoerders. Een evaluatierapport leidt een-op-een tot besluitvorming over beleid in de vorm van aanpassing, herontwerpen of stoppen van beleid. Dit is een hoofdzakelijk *lineair model*. Als onderzoekskwaliteit bepalend is, dan is de vraag wat de geldigheid en betrouwbaarheid van kennis is. Deze kwaliteit is dus van groot belang voor het begrijpen van de benutting in dit model: hoe meer kwaliteit een onderzoek heeft, hoe meer of betere benutting. Dit rationele en lineaire model is helder maar eenzijdig omdat er bij besluitvorming veel meer actoren, met verschillende macht en belangen, betrokken zijn. Die doen mee aan een 'spel' van trekken en duwen, waarbij sommigen het onderzoek naar zich toetrekken en anderen het afstoten of negeren. Beleid is niet alleen een zaak van cognitie (lees: uitdenken) maar ook van

communicatie over en weer (lees: uitvechten). Daarom zal dit lineaire model vermoedelijk de werkelijkheid van benutting niet goed representeren. Tenzij het een bijzonder onderzoek is dat gerelateerd is aan een crisis. Zoals bij voorbeeld een onderzoek naar het falen van een raket uit een maanprogramma omdat de oorzaak bekend moet zijn alvorens het lanceringsprogramma te vervolgen. Of een onderzoek naar uitputting van een belangrijke energiebron (Korsten, 1983).

Een tweede conceptueel model is probleemoplossingsmodel (Korsten, 1983), dat ook wel het *'demand-pull-model'* wordt genoemd (Bekkers, 2007: 318). Dit is een veel gehanteerd model. Het accent wordt hier *niet* gelegd bij het onderzoek of de onderzoekers maar *wel* bij de opdrachtgevers, vragers en gebruikers van kennis. De redenering houdt in dat als een beleidsevaluatie of ander onderzoeksverslag goed aansluit op 'beleid' en beleidsvraagstukken het rapport naar alle waarschijnlijkheid ook benut zal worden. Hoe meer aansluiting, hoe meer benutting. Opdrachtgevers en gebruikers moeten dan wel goede wensen of eisen formuleren ten aanzien van het onderzoek, en de onderzoeker moet die als het ware verwerken in de opzet en uitwerking van het onderzoek en de reflectie op uitkomsten respectievelijk de formulering van aanbevelingen. In deze optiek gaat veel aandacht uit naar een gemeenschappelijk referentiekader, naar overeenstemming over de opdrachtformulering en het begeleiden van het onderzoek en de onderzoeker(s).

Ook dit model is evenwel eenzijdig: het is veel gehanteerd maar slechts op een beperkt aantal gevallen van toepassing (Korsten, 1983). We zitten ook hier nog in de sfeer van *lineariteit en (eenzijdige) rationaliteit* (Bekkers, 2007: 318). Dit model houdt te weinig rekening met het feit dat een overheid zich nogal eens bevindt in een beleidsnetwerk en dat beleid maken, vaststellen en uitvoeren vaak een zaak is van zowel uitdenken als uitvechten (chiquer gesteld: van cognitie en sociale interactie). Bestuurders hebben medewerking en draagvlak nodig. In een netwerk is zelfs sprake van zodanige afhankelijkheden dat niet een actor beschikt over alle taken, bevoegdheden en middelen om alleen zijn wil door te zetten. Dat impliceert dat het netwerk productief moet worden door een gemeenschappelijk perspectief, doelvervlochten en middelenballing. Overigens geldt niet voor alle beleidsvraagstukken dat sprake is van een netwerkcontext.

Het derde conceptueel model ten aanzien van benutting van onderzoeksbevindingen is het *verspreidingsmodel* ('dissemination model'). In deze optiek, die we onder meer tegen komen op medisch terrein, wordt geaccentueerd dat wat van benutting wordt vooral een zaak is van een goede communicatiestrategie kiezen, waardoor de bevindingen ook goed voor het voetlicht komen en aandacht trekken (Korsten, 1983; Pröpper en De Vries, 1995; Hanney e.a., 2003; Bekkers, 2007: 318). Deze wijze van denken zoekt de oorzaak voor onderbenutting vooral in een inadequate presentatie en verspreiding van kennis en de reden voor benutting in een goede presentatie en disseminatie. Daar is wel iets voor te zeggen. Deze variabelen zijn belangrijk. Onderzoekers wie werk niet of beperkt benut wordt, neigen ertoe om met één rapport te volstaan en als ze nog iets publiceren dat dan te doen in een wetenschappelijk tijdschrift dat niet gericht is op de doelgroep.

Ook dit perspectief kan zekere eenzijdigheid niet ontzegd worden. Het perspectief neigt naar monocausaliteit: alsof alles afhangt van slechts één factor, de communicatie. Ook in dit perspectief wordt te lineair gedacht. De kwaliteit van het onderzoek kan er ook toe doen, de kenmerken van de onderzoekers en hun kijk op het te onderzoeken vraagstuk, en kenmerken van het ontvangend systeem. In feite is dit model complementair aan de genoemde.

Het vierde perspectief op benutting is het *interactiemodel*. In deze conceptualisering is het lineaire (monocausale) denken nagenoeg verlaten. De centrale gedachte uit het

interactiedenken is dat sprake is van *gescheiden werelden* (onderzoekers versus beleidvoerders en actoren uit het onderzochte veld) die verbonden moeten worden. De stelling is: naarmate kennisproducent, opdrachtgever(s) en gebruiker(s) langduriger en intensiever communiceren, zal de kans op benutting van onderzoeksbevindingen sterk toenemen. Dat betekent dat er sprake moet zijn van overleg tussen onderzoekers, opdrachtgever, gebruikers en wellicht ook 'partijen' uit het betreffende beleidsveld tijdens de opdrachtverlening, de uitvoering van het onderzoek, de publicatiefase en het verdere beleidstraject. Daarmee is wel duidelijk dat dit 'model' elementen uit het 'research push-model' en het 'demand-pull'-denken en het verspreidingsperspectief incorporeert (Korsten, 1983; Bekkers, 2007: 319). De interactie lijkt er niet onder als de onderzoekers proberen geldig en betrouwbaar onderzoek te verrichten, ze de gebruikers tijdig bij het onderzoek betrekken en rekening houden met hun kenmerken en percepties, en als ze proberen de resultaten op meervormige en adequate wijze te verspreiden. In de interactie komt juist een aantal van deze onderwerpen aan de orde. In feite is het interactiemodel dan ook een uitwisseling- en afstemmingsaanpak. Dat is positief.

In dit model wordt niet benadrukt dat in de overlegvormen de onderzoeker gemanipuleerd moet of mag worden. Het gaat eerder om het kweken van begrip voor de context waar de onderzoeksopdracht uit voortkomt, het stellen van de goede vragen in het onderzoek, het op het juiste moment uitbrengen van bevindingen. Dit *sociale* perspectief lijkt dus al beter dan de voorgaande maar de vraag is: is de politieke bestuurlijke component, dus van timing en macht en belangen er voldoende in opgenomen? Voldoende als de overleggende ambtenaren deze componenten aanvoelen en bespreekbaar maken en als de onderzoekers gevoel hebben voor de politieke barometer en met de politiek-bestuurlijke omstandigheden weten rekening te houden.

We komen bij een vijfde perspectief, het '*enlightenment*'-model (Korsten, 1983; Bekkers, 2007). Onderzoeksbevindingen kunnen een *agenderende functie* vervullen. Maar hoe gaat de agendering? In het openbaar bestuur kunnen ideeën, wensen, eisen en voorkeuren van onderop naar boven komen via kanalisatie door journalistenwerk, via de opiniepagina's, via politieke partijen en belangen- of actiegroepen. Ook komt het voor dat agendering van bovenaf geschiedt, doordat politieke bestuurders proefballonnen oplaten of expliciet in de samenleving bij doelgroepen steun aftasten door middel van een interview of discussienota. Ook kan agendering vanuit een tussenorganisatie geschieden, zoals een vakbond. Onderzoek kan daarin een rol vervullen.

Wie uitgaat van agendavorming (Kingdon, 2003) betreedt de arena van probleempercepties, probleemverkenningen, beleids(her)ontwerp, besluitvorming, implementatie. Een wereld die volop bezaait is met politieke voorkeuren. Wie zo kijkt, plaatst benutting van onderzoek in een *politiek-bestuurlijk perspectief*. Leidt onderzoek tot andere prioriteiten, tot erkenning van een vergrote ernst van de problematiek, tot een verschuiving in een controversie en nieuwe conceptualisering, tot een andere probleemformulering of tot herontwerp van een beleid? Wie zo kijkt onderkent dat onderzoek kan bijdragen aan een ander conceptualisering van beleid en 'reframing' (Weiss, 1977; Korsten, 1983; Bekkers, 2007: 320).

In het zesde perspectief is sprake van een '*politiek-argumentatief model*' (Bekkers, 2007: 320; ook Korsten, 1983: 18). Kennis is macht. Wie over kennis beschikt die anderen missen heeft een voorsprong. Onderzoek is in dit perspectief vooral ammunitie voor partijen met eigen belangen die eigen standpunten ontwikkelen, innemen en verdedigen. Ook politieke bestuurders en politieke fracties betrekken stelling en trekken onderzoek naar zich toe of stoten het af. Het verschijnsel 'tactisch gebruik' en 'strategisch gebruik' doet zijn intrede. Onderzoeksbevindingen zullen in dit perspectief dus benut kunnen worden door de bevindingen te omarmen, te kritiseren, af te stoten, of te negeren. Benutting is niet uitsluitend een zaak van kritiekloos aanbevelingen overnemen. Het is in dit perspectief dat ook aandacht

ontstaat voor selectieve benutting: het shoppen in de onderzoekswinkel om te vinden wat uitkomt.

Tot zover zes perspectieven op benutting die in de literatuur zijn aan te treffen. Deze perspectieven geven zicht op hoe met onderzoeksbevindingen omgegaan kan worden en wordt perspectieven genereren bovendien tal van factoren die van invloed zijn op benutting van onderzoeksbevindingen. De perspectieven zijn complementair.

Zijn enkele perspectieven meer belangrijk dan andere? De lineaire conceptualisering is schraal omdat de omvattendheid beperkt is. Niet alle onderbenutting is terug te voeren tot het onderzoek, of tot gebrekkige opdrachtarticulering of tot gebrekkige presentatie. Uit de literatuur komt wel naar voren dat het interactie-, het enlightenment- perspectief en het politiek-argumentatief perspectief in combinatie veel realiteitswaarde hebben. Het politiek bestuur heeft veel van doen met agendering, strijd en argumentatie. Maar nu al kan gezegd worden dat wie de specifieke benutting van visitatierapporten analyseert, wel met push & pull en interactie rekening mag houden want dergelijke rapporten zijn bedoeld om te gebruiken vanwege de leerfunctie en het op gebruik van visitatierapporten bestaande toezicht.

Wat laat het onderzoek naar benutting zien?

Het onderzoek van 'benutting' van kennis: vier richtingen

Visitaties in het wetenschappelijk onderwijs en andere visitaties zijn te beschouwen als een vorm van evaluatieonderzoek. Daarom is onderzoek naar benutting (lees: gebruik) van evaluatieonderzoek ('research utilization') hier van belang. Hoe zouden we greep moeten krijgen op benutting of gebrekkige benutting van visitatierapporten buiten het w.o.? Door aansluiting te zoeken bij het denken over benutting van onderzoek en onderzoeksrapportages. Daarin zijn meerdere denkrichtingen te onderkennen:

- a. de klassieke 'utilization research'-benadering gericht op het begrijpen van de benutting van toegepast (lees: gebruikgericht) sociaal-wetenschappelijk onderzoek (Caplan 1979; Weiss, 1979, 1980, b; Weiss & Bucuvalas, 1980; Van de Vall, 1980; Korsten, 1983; Patton, 1997; Bekkers c.s., 2004; Johnson c.s., 2009);
- b. de grenzenwerk-aanpak (Hoppe, 2002, 2003, 2007, 2008, 2010, 2012; Hoppe en Halfman, 2004; Hoppe en Huijs, 2003; In 't Veld, 2010; Derksen, 2011);
- c. de benadering van de lerende organisatie (Argyris & Schön, 1978; 't Hart, Hufen en Van Duin, 1988; Casey, 1993; Huysman en Van der Vlist, 1998; Greve, 2003; Caniëls, 2011);
- d. de pragmatische, meer bedrijfseconomisch gerichte aanpak (Van Helden e.a., 2005).

We geven hier een kort commentaar op en maken een keuze.

De klassieke 'utilization research'- benadering volgen we hier omdat we daarmee het meest vertrouwd zijn, deze nog steeds actualiteitswaarde heeft en we onvermijdelijk moeten uitgaan van visitatierapporten (Weiss & Bucuvalas, 1980; Van de Vall, 1980; Korsten, 1983; Landry e.a., 2003; Bekkers e.a., 2004; Van Dooren en Van de Walle, 2008). Hebben de andere benaderingen nog betekenis?

De grenzenwerk-aanpak ('boundary work') is primair door de Twentse hoogleraar Rob Hoppe ontwikkeld om zicht te krijgen op de relatie tussen planbureaus en de Nederlandse rijksoverheid en daarom hier minder relevant. Hoppe heeft deze aanpak ontwikkeld als basis voor een onderzoeksprogramma naar expertadvies. Deze auteur meent dat de 'utilization research' voor zijn doel *te instrumentalistisch* is en daarmee 'doorgeslagen' is; een verschraalde

aanpak die bovendien *teveel* uitgaat van eenrichtingsverkeer tussen onderzoek en advieswerk enerzijds en de beleidsorganisatie anderzijds. Dat eenrichtingsverkeer wordt wel heel duidelijk verwoord door de term 'doorwerking' (Bekkers e.a., 2004) die sommigen hanteren in plaats van 'benutting' of 'gebruik'; 'doorwerking' is een eenzijdige term, meent Hoppe (2007, 2008, 2010, 2012). Hoppe spreekt van te consumentistisch denken, alsof beleidvoerders met onderzoek een ijsje kopen in een winkel. Een 'ijsje' is hier de metafoor voor 'een onderzoek bestellen' waarbij de beleidsvoerders 'de volgers' zijn. Het *toeschrijven* van beleidswijzigingen of beleidsbevestiging aan een onderzoek, zou uitlopen op een expertocratie. Dat zou niet alleen een onwenselijkheid zijn, zo gaat die benutting van onderzoeksbevindingen (zoals benchmarkings, visitaties van bestuurskrachtmetingen) in beleid ook niet in de werkelijkheid. Het *interactiegericht denken* vat de verhouding tussen wetenschap en beleid daarentegen veel beter, aldus Hoppe. Over opdrachten en opdrachtformuleringen wordt immers overlegd of zelfs onderhandeld tussen opdrachtverleners en opdrachtnemers; in de beleidsanalyse worden dit de entreehandelingen genoemd. En beleidsvoerder en onderzoekers hebben gaande een onderzoek vaak ook contact, al of niet via een begeleidingscommissie. Dat kan leiden tot een uitbreiding van het onderzoek of een aanbeveling nog vóór het onderzoek afgerond is. Soms wordt een onderzoek daarmee al gebruikt voor het onderzoeksverslag verschenen is. In de 'utilization'-traditie moet dan gezocht worden naar 'taal' om dit effect te verwoorden. Overigens trekt In 't Veld de grenzenwerk-aanpak nog een stuk breder. Of onderzoek benut wordt heeft ook van doen met de heersende probleemduiding (routinematig probleem, wicked problem?, etc.) en het heersende discours (In 't Veld, 2010: 83-100). Discursief onderzoek dat handelt over een andere wijze van kijken naar een vraagstuk zal vaak genegeerd worden tot het moment dat het dominante discours sneuvelt of getransformeerd wordt in een nieuw overstijgend perspectief.

We zullen verderop zien dat de kritiek wel onderkend is dat daarom in de 'performance'-literatuur een meer interactiegerichte benadering wel degelijk terug te vinden is (zie Moynihan, 2008). En auteurs als Weiss (1991) en Weiss & Bucuvalas (1980) hebben gewezen op *het percolator-effect*: onderzoeksinformatie raakt vermengd met andere informatie.

We komen bij het als derde genoemde perspectief op gebruik of benutting van toegepaste research die enige diensten kan bewijzen bij het begrijpen van gebruik van visitatierapporten. Het perspectief van de lerende organisatie gaat ons inziens te weinig uit van de visitatierapporten en sterk van 'leren' als afhankelijke variabele. Maar wel komen beschouwingen over 'organisational learning' en 'the uses of performance information' wel tegen in studies over gebruik van prestatiemetingen (Van Dooren en Van de Walle, 2008; Askim, 2007).

In het pragmatische, meer bedrijfseconomisch gerichte perspectief worden vragen gesteld over de waardering van een meetinstrument in de organisatie, wordt nagegaan in hoeverre de uitkomsten ervan worden geaccepteerd en welke verbeteringsacties naar aanleiding daarvan worden ondernomen. Van Helden e.a. (2005) deden bij voorbeeld onderzoek vanuit dit perspectief naar benchmarking van bedrijfsvoering bij waterschapsorganisaties (ook Tillema en Van Helden, 2003; Van Helden en Tillema, 2004, 2005). Deze sterk economisch gerichte aanpak is voor zover we zien vooral op benchmarking gericht maar is ook te combineren met de 'utilization research'. In deze economische theorie worden hypothesen van de volgende soort geformuleerd (ontleend aan Van Helden en Tillema, 2004):

H1: De gemiddelde prestatie van organisaties zal verbeteren met achtereenvolgens uitgevoerde benchmarkingoperaties;

H2: De prestatieverschillen tussen organisaties zullen verminderen met achtereenvolgens uitgevoerde benchmarkingoperaties;

H3: Slecht presterende organisaties zullen door benchmarking sterker worden geprikkeld ter verbetering van hun prestaties dan goed presterende organisaties;

H4: Naarmate binnen een overheidsorganisatie meer de indruk bestaat dat door deelname aan (onderdelen van) een benchmarkproject of door prestatieverbetering, het bestaansrecht van de organisatie kan worden vergroot, wordt het waarschijnlijker dat de organisatie kiest voor deelname of prestatieverbetering;

H5: Naarmate de omgeving van een overheidsorganisatie hechtere relaties onderhoudt met andere organisaties binnen het organisatieveld, wordt het waarschijnlijker dat de organisatie kiest voor deelname aan (onderdelen van) een benchmarkingproject of prestatieverbetering. Zo kan men doorgaan. Voor benchmarking is mogelijk ook te lezen 'visitatietraject waarbij meerdere organisaties betrokken zijn'.

De klassieke benadering van 'the uses of performance information': twee thema's

De klassieke 'utilization'- benadering gericht op de benutting van toegepast sociaal-wetenschappelijk onderzoek kent twee centrale themata van onderzoek die van belang zijn voor het begrijpen van aspecten van benutting van bevindingen (dus uitkomsten) van visitaties of audits.

1. Het eerste thema betreft de vraag naar *de mate* van benutting en *de vorm* van benutting van bevindingen uit toegepast onderzoek (lees visitatie of audits).
2. En het tweede thema betreft *de verklaring* van matige of juist heel krachtige benutting.

We lichten beide thema's toe.

Mate en vorm van benutting in de klassieke benadering- eerste thema

Het eerste thema betreft *de mate* van benutting of gebruik van rapportages en bevindingen daaruit (ook wel intensiteit van benutting genoemd), en *de vorm van benutting*. De mate en vorm van benutting zijn op grofweg drie manieren te onderzoeken en ook onderzocht:

- a. door uit te gaan van een concreet onderzoek als vertrekpunt bij het zoeken naar benutting (fasenaanpak; lees: intensiteitsmeting);
- b. door uit te gaan van potentiële gebruikers, zoals ambtenaren en bestuurders (wat is hun zoekgedrag met betrekking tot informatie?);
- c. door uit te gaan van interactie (bijvoorbeeld analyse van standpunten en de onderbouwing daarvan; en van debatten); auteurs die veel op hebben met 'boundary work' benadrukken overigens deze derde mogelijkheid.

Achtereenvolgens wordt op elke manier ingegaan.

a. We gaan eerst uit van onderzoek, de eerstgenoemde aanpak. Hoe bereikt een onderzoeksrapportage van toegepast onderzoek in een organisatie de doelgroep van beleidvoerders en eventueel netwerkpartners? Als een in opdracht van een overheid uitgevoerd onderzoek wel conclusies en daaraan gerelateerde handelingsperspectieven biedt maar er wordt nauwelijks iets van in beschouwing genomen door een ontvanger (lees een gevisiteerde) dan spreekt men van *onderbenutting* of *ondergebruik*. De termen gebruik en benutting gelden hier als synoniem en worden hier door elkaar gebruikt.

Non-gebruik komt ook voor. *Non-gebruik* betekent dat een onderzoek al bij binnenkomst de la in gaat en er niet meer uitkomt. Daarvan bestaan meerdere vormen (Kerseboom, Leeuw en Elte, 2007; Kerseboom, 2008).

De manier waarop informatie wordt verzameld, is nog al eens bepalend voor de bruikbaarheid. Neem onderzoek naar de doeltreffendheid van beleid in de vorm van prestaties die worden afgezet tegen oorspronkelijke contracten of doelstellingen. Als dergelijk onderzoek in een interbestuurlijke setting regelmatig plaatsvindt, roept dat strategisch gedrag op bij personen die aantallen kunnen beïnvloeden, waarmee de meetmethoden aan waarde inboeten. Daarmee

hebben we de eerste vorm van non-gebruik te pakken: verkeerde of 'besmette' informatie ontvangen.

Een tweede vorm van non-gebruik houdt in dat beschikbare beleidsinformatie genegeerd wordt. Bij voorbeeld, er wordt selectief gewinkeld waarbij wat bruikbaar is voor voornemens en argumentaties gebruikt wordt en wat daarbij niet past genegeerd wordt. Hier kan sprake zijn van onethisch handelen of bedrog. Een derde vorm van non-gebruik is het nemen van beleidsbeslissingen zonder onderbouwing. Soms wil een kabinet iets ondernemen maar blijft onduidelijk waarom en wat het daarmee wil bereiken. De commissie-Dijsselbloem die vernieuwingen in het onderwijs bestudeerde trof ook gebrekkige onderbouwingen aan (Kerseboom, 2008: 45).

Als een onderzoek wordt aangewend voor legitimatie van iets waarover het onderzoek niet gaat en niet veel zegt, wordt gesproken van *misbruik* van onderzoek (Köbben en Tromp, 1999; Köbben, 2003). Over dit laatste hebben we het hier verder niet. We gaan dus uit van de veronderstelling dat misbruik van visitatierapporten hier geen punt van onderzoek is. Als een 'gebruiker' gaat shoppen in een rapport om te zien of er iets van zijn gading te vinden is, is de kans groot dat sprake is van '*selectief gebruik*': het ene meenemen in beschouwing en wat niet past 'weglaten'.

De vaststelling van de mate of intensiteit van benutting geschiedt vaak aan de hand van *een fasenconcept* ('stages'; Van de Vall, 1980). Denk maar eens aan de volgende fasen. Wordt een rapportage van een onderzoek ontvangen en komt het bij de goede behandelaar terecht of niet?; wordt vervolgens ook van de inhoud van het rapport kennis genomen of niet?; wordt het rapport daarna bestudeerd en verwerkt of niet?; wordt de inhoud van het rapport dan besproken of niet?; worden naar aanleiding van het rapport standpunten ingenomen en naar derden gecommuniceerd voor debat of vaststelling?; etc. Een dergelijke fasenbenadering kan laten zien dat het doorsijpelen van een rapport snel of traag gaat of dat een rapport ergens in een bepaald verwerkingsstadium 'blijft hangen', bij voorbeeld omdat de politieke-bestuurlijke urgentie om er iets mee te doen er niet is of verlopen is. Het onderzoek van Mark van de Vall (1980) naar benutting van 120 toegepaste onderzoeken door overheidsorganisaties laat zien hoe de fasenaanpak is uitgevoerd en wat zijn onderzoek naar gebruik van de 120 projecten opleverde. Wie wil begrijpen of sprake is van ondergebruik of onderbenutting van visitatierapporten kan deze aanpak dus toepassen.

Er bestaan verfijndere benuttingsladders. Van de Vall is niet de enige die denkt in termen van fasen of niveaus van benutting. Van Hoesel (2003) onderkent vier fasen: 1. reacties in het veld, de pers en de politiek; 2. verwerking van het resultaat door beleidsmakers in het beleidsproces; 3. invoering van nieuw beleid; 4. effecten van dit nieuwe beleid. Deze indeling achten we te grofmazig. Een verfijndere indeling is de volgende *benuttingsladder*:

- 1 ontvangst van het rapport ('weet hebben van ..');
- 2 waarneming dat ambtenaren de rapportinhoud op hoofdlijnen kennen;
- 3 interne discussie over de inhoud van het onderzoeksrapport door relevante 'gebruikers';
- 4 pleitbezorging ('advocacy') doordat sommigen intern of in een netwerk met het onderzoek in de hand 'gebruik' aan de orde stellen;
5. in een ambtelijk stuk naar het onderzoek verwijzen ('citation');
6. adoptie doordat het onderzoek meegenomen is in de opstelling, en de keuze of besluitvorming beïnvloed heeft (eventuele verwijzing in een bestuurlijk gedragen stuk);
7. 'impact' doordat het onderzoek tot wezenlijke actie aanzette;
8. determinatie: doordat het onderzoek de keuze (nagenoeg) bepaald heeft (verg. Landry e.a., 2003).

Is onderbenutting daadwerkelijk waargenomen?

Onderbenutting is waargenomen maar wel in een onderzoek waarin een 'strengere' methode is toegepast. In een onderzoek van Mulder e.a. (1991) is namelijk het accent gelegd op schriftelijke bronnen om *aanwijsbaar* gebruik van evaluatieonderzoeken bij de rijksoverheid waar te nemen. In hun onderzoek is bevestigd dat evaluatierapporten (direct en instrumenteel) bruikbaar waren en is gelet op de context voor het gebruik. Hiervoor zijn indicatoren gebruikt. Uit dat onderzoek bleek dat voor minder dan de helft van de rapporten sprake was van aanwijsbaar schriftelijk gebruik. Mulder staat niet alleen in de vaststelling van ondergebruik of meer en minder gebruik. Van de Vall's onderzoek is hiervan doordrenkt.

Als er van een zekere mate van benutting sprake is, is vervolgens de vraag: in welke vorm? Voor het vaststellen van de feitelijke vorm van (positieve) benutting wordt onder meer gebruik gemaakt van de volgende begrippen: *direct* en *indirect gebruik* (of directe en indirecte benutting); *gebruik op de lange of korte termijn*; *instrumenteel of conceptueel gebruik* (Korsten, 1983). Ook andere begrippen worden gebruikt, waarover verderop meer.

Onder 'direct gebruik' wordt verstaan dat een rapport rechtstreeks op de bureaus van de gevisiteerden komt en verwerkt wordt in beleidsvoornemens. 'Indirect gebruik' daarentegen impliceert dat een rapport een tijd blijft liggen en onderdeel wordt van het recenserend forum en/of discussie met belangengroepen of anderen. Het gebruik vindt als het ware via een omweg plaats. Na zekere tijd wordt de informatie dan met andere informatie samengebracht in een notitie of beleidsvoorstel. De kennisverwerking krijgt zo een soort *percolator-effect*, deze sijpelt geleidelijk met andere informatie door; op kortere of langere termijn. Carol Weiss (1980) noemde dit '*knowledge creep*'. Is er ook sluipend gebruik van visitatierapporten geweest?

Een rapport kan instrumenteel gebruikt worden of conceptueel. Instrumenteel gebruik verwijst naar maatregelen die alleen of voornamelijk zijn terug te voeren op een onderzoeksrapport. Bij voorbeeld een rapport brengt een wantoestand aan het licht en daarop wordt de directeur van een instelling op non-actief gezet en het 'lek gedicht'. Conceptueel gebruik impliceert dat de kijker op een vraagstuk verandert en misschien wel leidt tot een paradigmaswitch in beleid ('reframing'; Korsten, 1988; Schön & Rein, 1994). Iets dergelijks is wel voorgekomen met rapporten uit de onderwijssociologie. Rapporten uit die discipline zijn ten tijde van het ministerschap van de onderwijssocioloog Van Kemenade wel gebruikt om (meer en anders) te gaan denken in termen van (sociale) gelijkheid en ongelijkheid van kansen van leerlingen (Korsten, 1983).

Terwijl de impliciete norm bij toegepast (opdracht)onderzoek vaak is: direct, instrumenteel gebruik op de korte termijn blijkt de praktijk nogal eens dat onderzoek juist indirect, conceptueel en op langere termijn doorwerkt in beleid (Korsten, 1983). Wat is bij visitatierapporten het geval? Aannemelijk is dat hier sprake is van volgende te verwachten direct, instrumenteel gebruik op kortere termijn. De onderwijsinspectie komt immers kijken wat met het visitatierapport gebeurd is.

Er is – zoals gesteld – niet alleen onderzoek naar gebruik van prestatie-informatie ('the use of performance information') dat uitgaat van het onderzoek zelf en daarna de benutting in fasen volgt, de routing (direct, indirect gebruik) of de termijn waarop gebruik plaatsvindt maar ook specifiek onderzoek onder bestuurders (onder b genoemd; Ter Bogt, Askim) en onderzoek naar benutting vanuit een interactieperspectief (aanpak c; Moynihan).

b. Bij onderzoek naar gedrag van bestuurders en eventueel ambtenaren gaat dus om een vertrekpunt vanuit potentiële of werkelijke gebruikers, bijvoorbeeld door te kijken naar hun informatiezoekgedrag en hun vaardigheden om om te gaan met prestatie-informatie. Ter Bogt

(2001, 2003, 2004) deed in Nederland onderzoek onder wethouders naar hun (non)gebruik van uitkomsten van prestatiemeting. Hij stelde vast dat wethouders dergelijke veelal kwantitatieve informatie, bijvoorbeeld uit benchmarkings, maar weinig gebruiken, op de wethouders financiën na. Ze geven de voorkeur aan mondelinge contacten met ambtenaren en anderen, die de essentie van uitkomsten uit prestatiemeting kunnen weergeven en direct deze informatie kunnen interpreteren en in een bredere context kunnen plaatsen. Dat werkt sneller en biedt meer. Een teleurstellende bevinding voor fanaten op leergebied, dus voor al degenen die menen dat prestatierapporten onmiddellijk op grote schaal benut worden. Van de andere kant een begrijpelijk resultaat want Mintzberg had al vastgesteld dat vele leidinggevendenden in organisaties minder studieuze zijn dan gedacht; ze halen hun informatie mondeling en telefonisch bij vertrouwelingen in de organisatie. Begrijpelijk ook, want benchmarkrapporten roepen soms meer vragen op dan ze beantwoorden (Leeuw, 2003).

Maar dit is niet het hele verhaal. De uitkomsten van benchmark-studies wordt toch vaak door (gemeente)bestuurders wel van belang gevonden (Ter Bogt, 2005: 51). Het toegekende belang betreft dan niet zozeer de waardering van de precieze informatie over 'goed' of 'slecht' presteren en zeker niet de achtergronden daarvan, want die zijn in benchmark-studies vaak niet uitgewerkt. Maar de scores kunnen in vergelijkend perspectief toch iets zeggen. Is de score van de organisatie in vergelijking met andere organisaties voldoende of beter dan is de kans op geen actie ('non-gebruik') groot. 'Van wat voldoende of meer dan dat is, moet je blijkbaar afblijven'. Zeker als daar bij komt dat de benchmark over een politiek niet erg gevoelig en prioritair thema gaat, zoals met gemeentelijke dienstverlening nogal eens het geval is (omdat de kwaliteit van de dienstverlening voldoende of ruim voldoende is; Abma en Korsten, 2009a). Maar stel dat de scores onvoldoende of matig is dan kan bezinning op een benchmark-studie wel volgen. 'Dat kan aanleiding zijn tot het verrichten van nader onderzoek naar de achtergronden van een bepaalde score in de eigen organisatie' (Ter Bogt, 2005: 51). Bij voorbeeld door bij andere organisaties navraag te doen of door nader onderzoek te doen. De gebruiksvorm luidt dan: *de ene studie roept een andere studie op*. Het is ook mogelijk dan nagegaan wordt welke politieke keuzen een andere organisatie maakte.

Onderzoek naar benchmarking haalt de relevantie van het begrip 'gebruikswaarde' van onderzoek naar voren (zie ook Van de Vall, 1980). Veel geïnterviewden uit het onderzoek van Ter Bogt geven aan 'dat de gegevens in benchmarkstudies lang niet altijd goed te interpreteren en werkelijk vergelijkbaar zijn' (Ter Bogt, 2005: 51). De gebruikswaarde van een studie is gering als een benchmarkstudie de gebruiker voor leesbaarheidsproblemen plaatst (over opzet, methode van onderzoek, conclusies; gebruik van jargon), de bevindingen uit een studie moeilijk of niet te interpreteren zijn, en een studie geen verbeteroriëntatie biedt en geen concrete uitvoerbare aanbevelingen. Daardoor ligt de functionaliteit van benchmark-studies onder vuur, aldus Ter Bogt (2005: 52).

Askim (2007) deed onderzoek naar informatiezoekgedrag van Noorse gemeentebestuurders ('councillors') en de vraag of deze bestuurders de informatie gebruiken om een politieke keuze in de politiek-bestuurlijke arena te ondersteunen of te verwerpen. Daartoe formuleert hij een aantal contraire hypothesen vanuit een behoeftoperspectief (a) en een vaardighedenperspectief (b).

H1a luidt: Gebruik van prestatie-informatie (PI) is het grootst onder 'backbenchers' in het bestuur, dus – bot gesteld - onder de tweede garnituur. De contraire H1b-hypothese luidt: Gebruik van PI is het grootste onder 'frontbenchers', dus onder de bestuurlijke voormannen en - vrouwen.

H2a luidt: Gebruik van PI is het hoogste onder leden van oppositiefracties. De tegengestelde H2b luidt: Gebruik van PI is het hoogste onder leden van coalitiefracties.

Dit zijn slechts voorbeelden. Zijn onderzoek laat onder meer zien dat backbenchers en burgemeesters PI minder gebruiken dan frontbenchers (2007: 166). Het gebruik wordt niet beïnvloed door het feit of bestuurders deel uitmaken van de oppositie of de machtscoalitie. Een achtergrondvariabele blijkt wel relevant. Het gebruik van PI is het laagste onder 'the best-educated councillors' (2007: 167). Politieke ervaring maakt blijkbaar uit maar in ene verrassende richting. Hoe meer politieke ervaring in het lokaal bestuur hoe minder gebruik er van prestatiemeting gebruik wordt gemaakt. Veteranen gebruik PI minder dan jonkies. Deze uitkomst sluit aan op een bevinding van Melkers & Willoughby (2005) die stellen: 'the longer one's employment in government, the less one used performance information'. Hoe is dit te verklaren? De ervaren en hoger opgeleide bestuurders weten dat je beleid niet zo maar veranderd met een boekje met prestatie-informatie in de hand. Er is veel continuïteit (Askim, 2007: 168). Gebruik van PI blijkt overigens ook te maken te hebben met de stijl van gemeentebestuurders. Machtsgerichte bestuurders en politici relativeren prestatie-informatie, ombudsachtige types halen andere informatie elders, terwijl de compromiszoekers onder politici zich richten op het interne spel in bestuurlijke arena's. Vertegenwoordigers van deze stijlen hebben niet veel op met prestatie-informatie. Hier zitten niet de hongerige gebruikers van benchmarkinformatie of van andere prestatiemetingen.

c. Dan was er ook nog een derde benadering om benutting van prestatie-informatie in beeld te krijgen, niet door niet uit te gaan van het onderzoek als onafhankelijke variabele (fasenaanpak enz.) noch van bestuurders of ambtenaren (zoekgedrag van informatie in kaart brengen) maar door uit te gaan vanuit een interactief dialoogperspectief. Moynihan zegt in zijn boek *'The Dynamics of Performance Management'* (2008) dat performance management systemen op de veronderstelling berusten dat als er prestatie-informatie bestaat deze vanzelf wel gebruikt zal worden en zal leiden tot betere beslissingen. Zo makkelijk verloopt die causaliteit echter niet. Evenmin is het zo dat bestuurders altijd en vanzelfsprekend en op eigen initiatief wel verguld zijn met prestatiemetingen. Prestatiemetingen zeggen doorgaans niets over de context waarop de feiten slaan en de implementatiemogelijkheden van uitkomsten. Hij prefereert daarom denken in termen van interacties tussen relevante belanghebbenden. Zijn stelling is dat bestuurders belangen hebben, zoals een positie behouden of een beleidsvoorkeur bevestigd willen zien of het budget willen vergroten. Of een belang om een bestuur of bestuurder zo snel mogelijk weg te krijgen of om beleid te veranderen. En prestatie-informatie wordt in dat kader genegeerd of geprezen. Anders gesteld, gebruikt wordt datgene wat in de kraam te pas komt en als iets niet welgevallig is wordt het genegeerd. Prestatiegegevens worden dus in arena's bezien en vervolgens geïnterpreteerd, gebruikt, dan wel gekritiseerd. En dat gebeurt in geschreven teksten (nota's, memo's) en in oraal verband (tijdens presentaties en persconferenties, in debatten). Moynihan:

'How individuals perceive performance data will depend on their individual background and beliefs and the institutional role they fill' (Moynihan, 2008: 16).

Verder:

'The various actors involved in the policy process can be expected to use performance information consistent with their interests. Agency actors will contribute to the dialogue in ways that protect agency preferences and budgets' (2008: 17).

De selectie van informatie richt zich op wat bruikbaar is. 'The interactive dialogue perspective acknowledges that political preferences will dominate decisions', but that does not mean that performance data is not used'. We moeten dus beseffen dat uitkomsten van prestatiemetingen niet in een sprookjeswereld vol naïviteit terecht komen.

Conclusie

Deze klassieke benadering vanuit de 'utilization research' om bij het opsporen van benutting slechts uit te gaan vanuit onderzoek en rechtstreekse instrumentele benutting of doorwerking te verwachten, is als te instrumentalistisch gekritiseerd. Benutting kent vele andere vormen en

functies. Hoppe (2008) waarschuwt terecht. Nadere studie leert ook dat veel recente onderzoek de verwachting van direct, instrumenteel gebruik van kennis sterk tempert. Deze benadering houdt te weinig rekening met belangen, met machtsverhoudingen en interactie (Hoppe, 2008: 15; De Kool, 2007a, 2007b, 2007c, 2008). De interactieve benadering (zoals geformuleerd door Moynihan, 2008) houdt daar al meer rekening mee.

Wie de omvang en vorm van benutting van visitatierapporten bestudeert, kan volgens drie *complementaire* benaderingen te werken: uitgaan van de rapporten en de routing volgen (report analysis; eerder benadering a genoemd), uitgaan van het zoekgedrag van ambtenaren en bestuurders (search analysis; b), of een analyse maken van arenagedrag en dus onder meer van gebruik van rapporten en argumenten in debatten en rekening houden met belangen en macht (interactive analysis; c). De Kool (2007a,b en c; 2008) komt in verband met een studie naar interbestuurlijke monitoring met een driedeling in benaderingen die hij aanduidt als de rationalistische, sociale en politieke benadering. Report analysis schaarst hij onder de rationalistische optiek.

Tegenover al deze verbreding in optiek staat dat er wel degelijk visitaties en toegepast onderzoeken voorkomen die bedoeld zijn om direct te benutten. Een toezichthouder als de inspectie hoger onderwijs ziet zelfs toe op direct gebruik van visitaties van het wetenschappelijk onderwijs.

‘When a review report comes out, the recipients of review recommendations must consider action on it, and this action has to fit in the existing practices, beliefs and traditions’. En: ‘The review system has some in-built mechanisms to ensure that *some* action is taken’, aldus Jeliaskova (2002: 3).

We moeten dus niet doorslaan met de relativering van direct gebruik op korte termijn. Het visitatiesysteem in het w.o. dwingt in feite tot minstens enige bezinning aan de kant van de gevisiteerden.

Verklaringen van afwezige of geringe benutting in de klassieke benadering- tweede thema

Het als eerste behandelde thema in de ‘utilization research’ betref de mate van benutting of gebruik en de vorm van gebruik, het tweede themacomplex gaat over de *verklaring* van geringe of juist krachtige benutting, of van minder of meer benutting (of gebruik). In de literatuur zijn drie clusters aan theorieën te vinden voor een verklaring van onderbenutting of juist sterke benutting (Caplan, 1979; Hutjes & Cuisinier, 1982; Korsten, 1983; Van Oijen, Snellen en Van Westerlaak, 1984):

- het eerste cluster verklaringen betreft de ‘*knowledge constraints*’-verklaringen,
- het tweede cluster betreft de ‘*policy constraints*’-verklaringen, en
- het derde cluster betreft de ‘*two communities*’-verklaringen.

Korsten (1983) stelt dat deze drie verklaringsrichtingen blijken de literatuur op drie niveaus zijn te onderzoeken:

- het projectniveau van individuele onderzoeken (is juist dit onderzoek niet benut en waarom dan?);
- het structurele niveau, dus dat van de organisatie (bijv. ligt het niet benutten aan het management van het onderzoeksinstituut waarin geen plaats is voor bezinning op gebruik?; idem, geen bezinning bij de overheidsorganisatie?);
- en het culturele niveau (bijv. is niet benutting terug te voeren op de houding en opvattingen van onderzoekers of beleidvoerders?).

Figuur 2: *Verklaringsmogelijkheden voor geringe benutting van onderzoek in overheidsbeleid (Korsten, 1983)*

	Niveaus van analyse van mate van gebruik
--	--

		Niveau van onderzoeksproject	Structureel of organisatieniveau	Cultureel niveau in organisatie of sector
Verklaringen	a. Onderzoek en de onderzoeker	1a: factor 'kwaliteit van onderzoek'; reputatie onderzoeker;	2a: bijv. in onderzoeksinstituut geen sprake van reflectie op bruikbaar onderzoek?	3a: houding en opvattingen van onderzoekers: desinteresse in gebruik?; gering politiek gevoel
	b. Beleid en de beleidvoerders	1b: bijv. resultaten te laag ingebracht in gebruikersorganisatie?	2b: verbinding tussen kennisproductie en gebruik; organisatie van bezinning op onderzoek is gebruikelijk?	3b: houding en opvattingen van ambtenaren
	c. Communicatie	1c: bijv. taal waarin bevindingen worden geformuleerd; ontmoeten gebruikers onderzoekers tussentijds?; begeleidingscommissie?	2c: bijv. samenstelling begeleidingscommissie; 'knowledge utilization as a networking process'	3c:

De meeste studies naar de verklaring van meer of minder benutting zijn verricht op projectniveau en betreffen de cellen 1a, 1b en 1c. Ze gaan voor een aantal onderzoekingen na wat de reden is voor matig of juist sterk gebruik van bevindingen of meer of minder gebruik. Een studie van Van de Vall (1980) zit duidelijk in deze hoek. Ook zijn hier te plaatsen onderzoek naar gebruik van resultaten van evaluatieonderzoek of benchmarkings. Dat de andere niveaus overigens van belang zijn, blijkt bij Hutjes en Cuisinier (1982).

Bij deze figuur plaatsen we overigens twee kanttekeningen. Ten eerste is soms onduidelijk in welk cluster een verklarende factor precies zit. Ten tweede oogt de figuur statisch. In werkelijkheid kunnen factoren in elkaar grijpen of kan de ene factor bevorderend of remmend werken op een andere die dreigende onderbenutting of juist sterke benutting naderbij brengt.

Wie zicht wil krijgen op het procesverloop rond een onderzoek dat uiteindelijk wel of niet benut wordt door een opdrachtgevende overheidsorganisatie kan een *projectbiografie* opstellen. Zo is het bij voorbeeld mogelijk de onderbenutting van zeg vijf evaluatieonderzoekingen te bestuderen door van elk een biografie op te stellen waarin alle potentiële variabelen uit figuur 2 *in de tijd* aan bod kunnen komen. In een projectbiografie komt het onderzoek X aan bod vanaf het moment dat de opdracht nog niet eens geformuleerd was tot – stel - een jaar na het verschijnen van de onderzoeksresultaten. We noemen dit de *biografische benuttingsanalyse*. Deze aanpak is overigens nog vrijwel nooit toegepast. Een aanzet is ooit overigens gegeven door Ig Snellen.

We lopen de clusters langs. De figuur heeft namelijk als positief kenmerk dat de driedeling in clusters internationaal erkend is.

Eerste verklarend cluster: Factoren aan de kant van onderzoeker en onderzoek

De verklaring voor meer of minder gebruik of van onderbenutting zoeken aan de kant van onderzoek en de onderzoekers ('knowledge constraints') impliceert dat zwakke benutting of onderbenutting wordt toegeschreven aan een gebrekkige kwaliteit van het onderzoek of van de onderzoekers. Het onderzoek wordt niet benut omdat de onderzoekers bij voorbeeld het verkeerde hebben onderzocht en/of niet de goede waarnemingsmethoden kozen of anderen methodische zwakten aan de dag legden, zoals een gebrekkig concluderende kwaliteit of het ontbreken van aanbevelingen of van aanbevelingen die niet goed aansluiten op de bevindingen (Korsten, 1983).

Is gebrekkige kwaliteit van onderzoek wel aan de orde? Is er bewijs voor dat dergelijke factoren een rol spelen als verklaring? Nogal eens wordt in dit verband benchmarking op de korrel genomen. In het rapport *'Wil tot verschil; gemeenten in 2015'* (Commissie-Bovens, 2006) wordt betoogd dat benchmarking gemeentebesturen aan banden legt. Twee nadelen worden gearticuleerd. Benchmarkings werken nivellerend in plaats van onderscheidend en ze miskennen de zegeningen van verschillen tussen gemeenten ('wij lopen achter ten opzichte van het benchmarkgemiddelde'). Benchmarking belemmert dus de pluriformiteit in gemeenteland. De eigen omstandigheden, lokale kennis en afwegingen worden door benchmarkings weggemoffeld als middel tot verantwoording van een gemeentebestuur naar de gemeenteraad en samenleving. Dat veronderstelt dat benchmarking ook daadwerkelijk gebruikt wordt en tot keuzen leidt. Is daar grond voor? Ter Bogt(2005) is kritisch over benchmarking. Benchmarking, dus organisatievergelijking, kan zinvol zijn om prestaties te verbeteren, zo stelt hij, maar de gegevens en conclusies van benchmarkrapporten zijn niet altijd goed te interpreteren. De achtergrond van scoreverschillen tussen gemeenten wordt ook niet duidelijk. Gemeente X scoort ruim voldoende in een benchmark maar waarom? Gemeente Y scoort onvoldoende maar waarom? Ligt dat aan de fysieke kenmerken van het territorium, aan het gedrag of de opvattingen van burgers of aan beleid of aan iets anders? Dergelijke informatie ontbreekt of is schaars in benchmarkrapportages. Daarmee raakt de gebruikswaarde van benchmarkings ernstig aangetast. Er wordt in de rapportages ook geen brug geslagen naar 'hoe' beleidvoerders de conclusies zouden kunnen gebruiken. De vertaalslag naar beleid wordt niet of niet voldoende gemaakt. Toch maakt onderzoek van Ter Bogt ook wel duidelijk dat bestuurders een zekere interesse hebben in benchmarkings. Logisch want er kan ook een *glamoureffect* van uitgaan. Gebruik is vaak niet geheel afwezig. Op de gebruiksladder worden wel enkele stappen op treden gezet.

Welke inzichten zijn er nog meer? De manier waarop het probleem in *onderzoekbare termen* is geformuleerd, blijkt van betekenis voor de benutting van latere bevindingen. Er moet dus een inhoudelijk rijke *vertaling* plaatsvinden van een beleidsproblematiek in een onderzoeksvraagstelling, een expliciete doelstelling en een onderzoeksstrategie (Hutjes en Cuisinier, 1982). De doelstelling moet duidelijk maken wat men waarom wil weten en dus waaraan een bijdrage geleverd wordt. Mulder e.a. (1991) stelden vast dat in de helft van onderzochte evaluatierapporten een doelstelling echter niet te vinden is.

Het heeft ook een gunstig effect op 'benutting in beleid' als onderzoekers een grote hoeveelheid potentieel *verklarende variabelen* voor een verschijnsel gebruiken? Breed verklaren is beter dan smal verklaren? Ellemers heeft verdedigd dat het beter is om variabelen in beschouwing te nemen die ertoe doen en die je met overheidsbeleid kunt beïnvloeden ('*manipuleerbare variabelen*') dan dat een onderzoeker veel verklaart maar er weinig te veranderen valt (zie ook Ellemers, 1976; Hupe, 1984). Als men weet dat wapenbezit leidt tot meer geweld heb je daar meer aan dat dat je weet dat erfelijkheid van invloed is op geweld. Want aan wapenbezit is wat te doen en aan de genetische kenmerken van mensen komen, is niet goed mogelijk of niet toegestaan. Terwijl de genetische factor misschien wel het meest verklaart. Dit is het dilemma van 'veel kunnen verklaren en weinig kunnen veranderen' genoemd versus 'veel kunnen veranderen maar weinig kunnen verklaren'.

De relevantie van de geleverde informatie speelt een rol bij het gebruik van een onderzoek. Bij benchmarking van dienstverlening zal onderzoek dus iets moeten zeggen over *de kern* van de dienstverlening (tijdigheid, klantvriendelijkheid, of de dienst kwalitatief goed is, enz.) en niet slechts iets over de openingstijden van het loket. Met een benchmark over marginale aspecten van dienstverlening zal weinig gebeuren.

Een *geloofwaardig* verricht onderzoek zal beter benut worden. Onderzoek is geloofwaardiger als bij de informatieverzameling gebruik is gemaakt van *meerdere waarnemingsbronnen*. 'Triangulatie', het over elkaar leggen van informatie uit verschillende bronnen, is blijkbaar gunstig (Swanborn, 1999). Ook is positief als onderzoek zich positief verhoudt tot bestaand beleid. Onderzoek vergezeld van aanbevelingen, die het bestaande beleid op de kop zetten, wordt minder gebruikt. Wat niet uitkomt, wordt blijkbaar makkelijk ter zijde gelegd (Mulder e.a., 1991; Weiss & Bucuvalas, 1980).

Aan een onderzoek moet in het algemeen voldoende *methodisch-technische kwaliteit* ten grondslag liggen wil het tot gebruik of benutting komen. Een te lage kwaliteit in dit opzicht kan een reden zijn om het rapport niet te gebruiken (Mulder e.a., 1991). Overigens worden 'zwakkere onderzoeken' soms toch gebruikt, bijvoorbeeld als ondersteuning van beleid blijkt (Hutjes en Cuisinier, 1982).

Er zijn nog enkele voor de hand liggende factoren te noemen. *Zo is timing* van belang. Een onderzoek moet verschijnen op een tijdstip dat het ook gebruikt kan worden. Een korte doorlooptijd van het onderzoek zal het gebruik ook bevorderen (Mulder e.a., 1991). Een andere factor is de *presentatie* van het onderzoek. De rapportage moet in het algemeen aansluiten op het begrip van de gebruiker(s) (Swanborn, 1999). Concrete en *duidelijke* bevindingen, conclusies en aanbevelingen zijn van betekenis om gebruik te stimuleren, zeker als ze de *aandacht* van ambtenaren, politici en anderen weten te trekken (Mulder e.a., 1991). Onderzoek dat *draagvlak* heeft, zal kans maken om gebruikt te worden (Hupe en Van der Meer, 2006) maar onderzoek dat met een uitdagende of controversiële aanbeveling komt, kan ook gebruikt worden om debat te stimuleren of zich tegen af te zetten.

We springen nu van het onderzoek naar de onderzoeker(s). Literatuurstudie wijst verder uit dat er onderzoek is gedaan naar de vraag of de *reputatie* van onderzoekers uitmaakt voor de mate van benutting van onderzoeksuitkomsten (Weiss & Bucuvalas, 1980). Een gunstige conditie is over het algemeen: gereputeerde onderzoekers die een goed onderzoek verrichten, daarover excellent en leesbaar in fasen rapporteren met bruikbare ('beleidsrelevante') bevindingen komen en op een 'goed' tijdstip met bevindingen naar buiten komen (Weiss & Bucuvalas, 1980). Voor onderzoek kan men 'visitatie' lezen. Vertaling naar visitaties doet de stelling opkomen: maakt de reputatie van de visitatiecommissie uit voor de benutting van het visitatierapport aan de kant van de gevisiteerden? Het verwachte antwoord is 'ja, dat maakt uit': hoe gereputeerder de commissie is en hoe meer de commissie in het 'veld' erkend wordt, hoe groter de kans op benutting van de inhoud (en boodschap) van visitatierapporten. Maar het is zelden de enige factor. Ook de kwaliteit van het onderzoek (lees zelfevaluatie en visitatierapport) zelf maakt uit voor de benutting.

Wie verdere inspiratie wil opdoen over factoren die er bij benutting in de Nederlandse context toe doen, kan onder meer terecht bij Van de Vall en Van Dijkum-De Jong (1979) en Van de Vall (1980). Van de Vall verrichtte een kwantitatief onderzoek naar benutting van 120 onderzoeken die alle in opdracht waren uitgevoerd en waarbij sprake was van gebruik van methoden en technieken van sociaalwetenschappelijk onderzoek. Het ging onder meer om toegepast onderzoek op het vlak van ruimtelijke ordening. Steeds waren de resultaten neergelegd in een rapport. Van de Vall komt tot de volgende bevindingen.

Onderzoek dat werd uitgevoerd door *interne* onderzoekers van een overheidsorganisatie werd meer benut dan (uitbesteed) onderzoek dat door externe onderzoekers was uitgevoerd (1979: 19). Een redelijke mate van betrokkenheid van de interne onderzoeker van toegepast onderzoek bij beleidsmaatregelen had ook een gunstig effect op de benutting van een rapport in de beleidsvorming (Van de Vall en Van Dijkum-De Jong, 1979: 20). Gebruik van theorieën en

abstracte begrippen in het onderzoek had daarentegen een ongunstig effect (1979: 23). Onderzoek mag methodologisch gezien wel nauwgezet zijn als het maar een ruimte oriëntatie heeft op *de informatiebehoefte van beleidsvoerders* (1979: 25). Onderzoek moet dus beleidsrelevant zijn en de onderzoeker moet gevoel hebben voor wat speelt bij de beleidsvoerders en wat raakt aan zijn onderzoek, zodat hij daarop kan inspelen (bij voorbeeld door publicatie van bevindingen op een moment dat die bevindingen er nog toe kunnen doen).

Niet alleen de onderzoeker en het onderzoek werkt in op benutting, ook rapportagefactoren blijken – zoals eerder gesteld – relevant voor het begrijpen van de benutting van onderzoek in beleid. Dikke rapporten worden minder benut dan minder dikke (Van de Vall en Van Dijkum-De Jong, 1979: 31). Het is in het algemeen ook gunstig als onderzoekers goed nadenken over de te hanteren communicatiestrategie en dus over de verspreiding. Alleen een rapport uitbrengen is ongunstig. Het tevens tussentijds informeren van de opdrachtgever en eventuele doelgroepen apart bedienen, is juist gunstig, evenals het gedoseerd – in stukjes en beetjes – brengen van de bevindingen. Aan de kant van beleidsvoerders moet de ontvankelijkheid en aandacht voor het onderzoek blijikbaar kunnen groeien. De onderzoeker die weet te bereiken dat beleidvoerders zelfs uitzien naar het onderzoeksrapport heeft goed gewerkt. Dat zal des te beter lukken als de onderzoeker duidelijke conclusies uit het onderzoek weet te trekken.

Onderzoeken waarvan bevindingen een rel veroorzaken of protest uitlokken, blijken beter benut te worden dan andere (Van de Vall, 1980: 179). Dat komt omdat het onderzoek dan ‘in de krant komt’ en actoren in het veld zich direct op het onderzoek storten (ze leggen een verbinding met macht en belangen). Dan is de kans groot dat parlementariërs schriftelijke of mondelinge vragen aan de betrokken bewindspersoon gaan stellen en het bij de beleidvoerders direct ‘alle hens aan dek is’ (vragen moeten worden beantwoord, enz.).

Conclusie: er is dus geen sprake van monocausale werking van een goed onderzoek dat positief benut wordt. Wie als onderzoeker wil dat het onderzoeksrapport gelezen wordt zal de ruimte moeten krijgen en grijpen om tal van handelingen te verrichten om het onderzoek in het beleidsnetwerk ‘te brengen’ en een communicatiestrategie moeten uitvoeren. De kans is groot dat onderzoek dat geen protest uitlokt, onbenut neervalt als een natte krant op de grond, tenzij andere actoren het oppakken. Dat oppakken heeft een onderzoeker niet helemaal in de hand.

Tweede verklarend cluster: Policy constraints

Wie de verklaring in de hoek van de overheidsorganisatie, het beleidsnetwerk en de beleidvoerders zoekt, komt uit bij de manier waarop beleidsprocessen en andere processen verlopen aan de kant van de ontvangende partij, de ambtenaren in grote organisaties en de bestuurders. Deze factoren zijn wel getypeerd als ‘policy constraints theories’.

Het type organisatie blijkt er in het algemeen toe te doen. Is de organisatieleiding al langer bezig met prestatie meting of niet? Is prestatie meting gerelateerd aan de bestuurlijke ambitie of niet? Het interne ambitieniveau blijkt van invloed op benutting: hoe meer ambitie, hoe meer neiging om ook naar prestaties te kijken (Van Helden en Brouwer, 2005). Het maakt uit als de leidinggevendenden tonen dat ze onderzoek stimuleren om ervan te leren en ze de processen van bezinning ook stimuleren. Mulder e.a. (1991) sluiten hierop aan. De aanwezigheid van procedureel-organisatorische voorzieningen ten aanzien van het starten van onderzoek, de uitvoering ervan en afhandeling van beleidsevaluatieonderzoek draagt bij aan de benutting van onderzoeksbevindingen. Hoe minder dergelijke voorzieningen voor onderzoeksprogrammering, bespreking en andere bezinning aanwezig zijn, hoe meer benutting afhankelijk wordt van ad hoc situaties of toeval. En als er ook nog eens vol hoon en

sceptis door ambtelijke leidinggevenden gesproken wordt over evaluaties of benchmarks zegt dat veel: de kans is gering dat er veel met onderzoeksbevindingen gebeurt.

Derksen stemt hiermee in. Veel van de factoren uit dit cluster die van invloed zijn op benutting hebben te maken met *'verbindingen'* (Derksen, 2011: 118). Naarmate een onderzoek meer aansluit op het werk- en onderzoeksprogramma van een overheidsorganisatie ('if any'), neemt de kans op benutting toe. Dan is immers sprake van *geïnstitutionaliseerde ontvankelijkheid* voor het onderzoek. Maar het is niet alleen een zaak van regels en procedures. Naarmate gebruikers meer betrokken zijn bij onderzoek, en het onderzoek dus van begin tot einde kunnen volgen, zullen ze zich er minder van distantiëren (Tijssen, 1988). Naarmate ambtenaren en bestuurders meer veranderingsgezind zijn en meer open staan voor innovatie, zal de kans op benutting ook toenemen (Swanborn, 1999). Zijn ambtenaren weinig veranderingsgezind, betrek ze dan tijdig bij het onderzoek want dan zal hun weerstand vermoedelijk afnemen.

Onderzoekers hebben er ook voordeel bij als ze de ambtenaren en bestuurders kunnen *begrijpen*. Van der Meer e.a. (2000) stelden vast dat de benutting van evaluatieonderzoek niet zozeer afhangt van de evaluatie zelf maar vooral bepaald wordt door de directe en indirecte aansluiting van de evaluatie op de manieren van denken en doen van actoren in de overheidsorganisatie. In *'Can Governments learn?'* komt een vergelijkbare bevinding naar voren. Daarin is ingegaan op de vraag onder welke voorwaarden organisatorisch leren plaatsvindt en beïnvloed kan worden door evaluaties. Aangegeven wordt dat een continue inspanning nodig is om bruikbare informatie voort te brengen en samen te werken met gebruikers (Leeuw e.a., 1994). Ook hier wordt dus het belang benadrukt van bruikbare kennis, van verbindingen, van netwerken.

Een andere factor betreft het niveau waarop een onderzoek binnenkomt. Als een (in opdracht verricht) onderzoek laag wordt ingebracht in de opdrachtverlenende rijksorganisatie is dat een conditie die gebruik ernstig bemoeilijkt; het is dus een barrière in een agenderingstraject (Korsten, 1983). Immers, een onderzoek moet dan een lange route doorlopen alvorens dit onderzoek via *'samenvatting op samenvatting'* in een departement of andere organisatie naar boven komt en uiteindelijk belandt op het bureau van degenen die *'er echt over gaan'*. Wie wil dat een onderzoek direct en snel benut wordt, moet zorgen dat het hoog wordt ingebracht en ook de netwerkpartners bereikt.

Institutionele kaders en ontmoetingsplekken, verbindingen, bruikbare kennis, zorgen dat ze de bevindingen begrijpen, rapporten hoog inbrengen, leidinggevenden die ambitie hebben en willen innoveren; is er nog meer? Het is mogelijk dat de beleidvoerders qua opleiding niet erg vertrouwd zijn met het lezen en interpreteren van onderzoeksuitkomsten. Op het eerste gezicht is dit een triviale, bijna *'ongeloofwaardige'* factor. Toch is de werking ervan aangetoond. Als men juristen een sociaal-wetenschappelijk onderzoek naar abortusbrieven laat lezen kan dat bij voorbeeld iets krijgen van *'schuring'* en *'onbegrip'*. Juristen zijn over het algemeen niet vertrouwd met sociaal-wetenschappelijk onderzoek en/of begrijpen het niet of nauwelijks. Juristen weten niets van regressieanalyse; het woord *'probleemstelling'* of *'operationalisering'* is hun doorgaans ook vreemd. Ze vertrouwen sociaal-wetenschappelijk onderzoek daarom ook minder. Volgens Duits onderzoek zijn de meeste juristen niet vertrouwd met sociaal-wetenschappelijke methoden en technieken van onderzoek en maakt dat uit voor benutting van onderzoek (Friedrich, 1970; Korsten, 1983).

Verbindingen vergen een bepaalde cultuur. Daar willen we ook nog eens op wijzen (Derksen, 2011: 118). Zie cel 3b in figuur 2. En hoe is het dan met cel 2b: is de structuur van de organisatie van belang voor de benutting van onderzoeksbevindingen? *'Structure matters'*. We wezen al op

de institutionele factor. De aanwezigheid van 'grenswerkers' binnen een departement doet er ook toe, stelt Derksen die als 'chief scientist' werkzaam was bij het ministerie van VROM en dus van nabij heeft kunnen zien of zijn rol zoden aan de dijk zette (Derksen, 2011: 119). Grenswerkers kunnen discussies over onderzoeksbevindingen aanzwengelen en bloot leggen waar onderzoek vergeten dreigt te worden.

Zijn er meer positieve condities in dit cluster voor benutting van onderzoeksbevindingen te vinden? 'Timing' is van belang, het is al gezegd. De kennis moeten op het juiste tijdstip worden aangeleverd in de ambtelijke organisatie (Derksen, 2011: 118). De agenderingstheorie van John Kingdon (2003) biedt een voedingsbodem hiervoor. Onderzoeken en visitaties zijn volgens deze theorie effectiever als ze passen in een klik van drie stromen of die klik bevorderen: de verbinding tussen een reservoir aan problemen, oplossingen en politiek geschikt tijdstippen. Ook synchronie blijkt relevant. 'Vaak zijn visitaties effectiever wanneer de organisatie erin slaagt ze synchroon te laten lopen met andere organisatorische trajecten', aldus Schillemans (2009: 163) die onderzoek deed naar verantwoordingsprocessen.

Tot zover dit cluster.

Derde verklarend cluster: Gescheiden werelden, gebrekkige communicatie

De verklaring zoeken in de communicatie ('two communities theories') stoot op twee werelden, van onderzoekers en beleidvoerders, die elkaar niet (goed) verstaan. Hier gaat het om de interactie. Wellicht hebben onderzoekers niet in de gaten dat een onderzoek voor moment x moet verschijnen omdat dan een debat aanstaande is over een bepaald verschijnsel waar het onderzoek ook betrekking op heeft. Een onderzoek moet op het geschikte moment verschijnen gezien de politiek-bestuurlijke context. Om de kans op elkaar niet goed begrijpen te verkleinen worden vaak begeleidingscommissies ingesteld die letten op de kwaliteit van het onderzoek, de leesbaarheid van de rapportage en de onderzoekers doordringen van 'betere en slechtere momenten van publicatie en openbaarmaking van uitkomsten'. Wordt dit in onderzoek uit de 'utilization research' bevestigd?

Als een begeleidingscommissie het toegepast onderzoek begeleidt, blijkt dat een gunstig effect op de benutting te hebben, volgens het onderzoek van Van de Vall (1980: 169). Begrijpelijk want in een dergelijke commissie zijn vaak (opdrachtgevende) ambtenaren opgenomen, ambtenaren die later met het onderzoek moeten 'werken' en representanten uit het onderzoeksveld. Zo'n commissie is dus ook een hulpmiddel voor onderzoekers die de commissieleden ontmoeten om zicht te krijgen en houden wat in het veld en in de overheidsorganisatie of politieke arena speelt.

Ook blijkt dat het hanteren van (communicatieve) terugkoppelingsstrategieën een gunstig effect heeft op de benutting van onderzoeksresultaten in de beleidsvorming (Van de Vall en Van Dijkum-De Jong, 1979; Van de Vall, 1980). Voorbeelden van deze positieve strategieën zijn: a. het uitbrengen van tussenresultaten; b. het op een geschikt moment presenteren van onderzoeksresultaten, namelijk als de kans op politiek-bestuurlijke en ambtelijke aandacht groot is; c. het benutten van verschillende communicatiekanalen (niet alleen schriftelijke rapportages maar ook mondelinge presentaties, etc.); d. het gedoseerd verstrekken van bevindingen (en niet alles tegelijk) waardoor de aandacht gevangen wordt en een onderzoek 'in de week gezet wordt' en men ernaar uitzielt; e. het verstrekken van informatie over onderzoeksbevindingen aan verschillende doelgroepen, die deze informatie kunnen benutten om het bestuur te bestoken en eventueel 'lawaaï' kunnen veroorzaken (1979: 34).

Dergelijke factoren zijn ook in verband te brengen met de verklaring van eventuele onderbenutting van visitatierapporten. Bij voorbeeld: Heeft een visitatiecommissie wel feedbackstrategieën gehanteerd? Is het resultaat aan doelgroepen gepresenteerd?

De betekenis van verbinding en communicatie bleek ook uit ander onderzoek. Dat wijst op de betekenis van georganiseerde bezinning, waarbij kennisgebruikers geholpen worden om na te denken over wat onderzoek voor hen kan betekenen. Reflectie werkt drempelverlagend en brengt onderzoeksbevindingen als het ware naderbij. Askim e.a. (2006) onderzochten de leereffecten bij Noorse gemeenten. Het leereffect bleek groter onder gemeenten die hadden meegedaan aan benchmarkonderzoek en in netwerken georganiseerd waren, zeker als het netwerk heterogeen samengesteld was. De betekenis van netwerken blijkt ook bij Yin & Gwaltney (1981). Ze keken naar de manier waarop onderzoeksideoën op het federale niveau in de V.S. bruikbaar konden worden. Wetenschappelijke kennis kan het meest direct gebruikt worden als er een communicatienetwerk opgezet en onderhouden wordt. Het netwerk dient te bestaan uit kennisproducenten en gebruikers en het netwerk moet continu functioneren.

Welk van de drie verklaringsclusters heeft het primaat?

Welke verklaring van deze drie clusters (factoren aan de onderzoekskant belemmeren benutting, of factoren aan de beleidskant of factoren in de interactie?) is de belangrijkste om onderbenutting te begrijpen? Onderzoek laat zien dat niet zelden alle drie de verklaringen betekenis hebben (Korsten, 1983; Van Oijen, Snellen en Van Westerlaak, 1984) maar Caplan (1979) heeft nadrukkelijk gewezen op de laatstgenoemde verklaring van 'gescheiden werelden' vol onbegrip. Dat we de drie verklaringsrichtingen in het oog moeten houden blijkt ook uit ander onderzoek. Onderzoek van Dal Santo e.a. (2002) naar verspreiding en gebruik van onderzoekingen bij regionale overheidsdiensten in de Verenigde Staten wijst uit dat de belangrijkste factoren die het gebruik van onderzoek beïnvloeden zijn: kenmerken van de onderzoeker, kenmerken van de 'ontvangende' organisatie en de daarin actieve medewerkers en de communicatie. Specifieke projectomstandigheden spelen ook een rol.

Belangrijk is te onderkennen dat niet slechts de betekenis van drie clusters verklaringen voor ondergebruik of onderbenutting van (toegepast) sociaal-wetenschappelijk onderzoek als zodanig aan de orde is maar juist de '*use of performance information*'. Ook hier drie clusters ter verklaring van ondergebruik van bevindingen uit prestatiemetingen? Op dit vlak is onderzoek verricht (Van Dooren en Van de Walle, 2008; Van Dooren e.a., 2010; Taylor, 2011; Moynihan e.a., 2012). Dat onderzoek leert dat de aanwezigheid van transformationele leiders de benutting van visitatie-uitkomsten positief beïnvloedt. Denk bij dergelijke leiders aan onderwijsdecanen. Waarom? Omdat transformationele leiders op beleidsverandering in de vorm van verbetering gericht zijn en daarom medewerkers kunnen en zullen aansporen om kennis te nemen van uitkomsten van prestatiemetingen. Dat gedrag kan versterkt worden door verbeterdoelen te stellen ('goal clarity') en bij te dragen aan een organisatiecultuur waarin meting en kwaliteitsverbetering van belang geacht worden (Moynihan, Pandey & Wright, 2012). Australisch onderzoek (Taylor, 2011) maakt duidelijk dat 'the use of performance information' niet alleen afhangt van een op prestatieverbetering gerichte leiderschapsstijl maar ook positief beïnvloedt wordt door het scheppen van een structuur met organisatieroutines gericht op meting en verbetering (waarin data over prestaties ook beschikbaar zijn; wie wil verbeteren moet over gegevens kunnen beschikken), en het creëren van een cultuur van benutting (medewerkers moeten ontvankelijk zijn voor prestatiemeting en uitkomsten).

Deze factoren (leiderschapsstijl, doelgerichtheid, structuur en cultuur) horen thuis in het eerder als 'policy constraints' aangeduide cluster van verklaringen voor ondergebruik. Bleek ook iets te achterhalen over de opstelling van politieke bestuurders? Taylor signaleert 'a reported lack of interest amongst Australian politicians on performance measurement' (Taylor,

2011: 1329). Deze bevinding slaat echter vooral op benchmarking, een aanpak van prestatiemeting.

Sectoraal en internationaal onderzoek

Naar benutting van benchmarkingonderzoek is onderzoek gedaan. Dorsch & Yasin (1998) bieden een raamwerk en Nelissen en De Goede (1999) treden in hun voetsporen. Van Helden en Brouwer (2005) richtten zich op de bedrijfsvergelijking zuiveringsbeheer (bvz), een benchmarking. De relatie tussen benchmarking van bvz en prestatieverbetering bleek in hun onderzoek niet overtuigend aanwezig. Tillema (2006) blijft optimistisch. Ze onderzocht de bvz bij vier waterschappen en komt tot de vaststelling dat benchmarkinformatie zeker een rol *kan* spelen bij het realiseren van prestatieverbeteringen maar dan moeten wel enkele hindernissen worden overwonnen.

Visitaties

Nu terug naar de visitaties. We moeten rekening houden met gebruik in allerlei vormen. Hatry (2008) spreekt in het kader van hoe prestatie-informatie wordt gebruikt immers over *'The Many Faces of Use'*. We gaan er hier vanuit dat bij visitatierapporten in het w.o. veelal sprake was van directe en instrumentele benutting van bevindingen op de korte termijn (ook: De Lancer Julnes, 2008: 58-72) en niet of minder van indirecte en conceptuele benutting op langere termijn.

Waarom? Het onderzoek heeft duidelijk gemaakt dat de gevisiteerden vooral werk maakten van het wegwerken van 'onvoldoendes' en minder deden aan verbetering van wat al voldoende was. Dat betekent dat zekere onderbenutting van bevindingen aan de orde was. Hoe die te verklaren? Dat ondergebruik kan vermoedelijk niet verklaard worden uit het elkaar niet verstaan ('two communities') want de visitatiecommissies waren deels bememd met vakgenoten. En de decanen konden een conceptrapport van commentaar voorzien. De 'two communities'- verklaring krijgt volgens ons geen voet aan de grond. De verklaring zoeken aan de kant van de 'knowledge constraints' heeft ook maar beperkte verklarende kracht. De visitatierapporten waren niet slecht. Weliswaar waren de visitatierapporten geen diagnostische en remediale rapporten van een adviesbureau maar ze waren voldoende richtinggevend om gebruikt te kunnen worden. Als onderbenutting plaatsvond moet die vermoedelijk vooral bij de gevisiteerden gezocht worden. Knelpunten in de vorm van 'policy constraints' veroorzaken onderbenutting of ondergebruik. Gevisiteerden hebben vermoedelijk weinig trek in een brede verbeteringsoperatie die ook gericht is op verbeteren wat al voldoende is. Maar extern toezicht op de visitatie zorgt voor tegendruk. Als er geen extern toezicht plaatsvindt, is de kans op ondergebruik van visitatierapporten aanzienlijk groter.

Conclusie: we nemen aan dat visitaties buiten het w.o. zullen laten zien dat sprake is van een zekere mate van benutting die de vorm aanneemt van: direct, instrumenteel gebruik op de kortere termijn. Maar ook van onderbenutting. De kans dat een beleid als gevolg van de benutting van visitatie-uitkomsten opnieuw herontworpen wordt, is echter klein. Van radicale innovaties zal, naar analogie van het wetenschappelijk onderwijs, geen sprake zijn. Wie in een visitatierapport 'voldoendes' scoort, verbetert doorgaans weinig maar vanuit 'onvoldoendes' worden wel degelijk verbeterinitiatieven genomen. Om een metafoor te gebruiken: als visitatie van een bakkerij aangeeft dat de kwaliteit van het deeg voor bepaalde producten niet voldoende is, weet de bakker wel wat hem te doen staat: zorgen voor beter deeg. Hij zal niet onmiddellijk het hele assortiment gaan vernieuwen of de totale productiewijze omgooien.

Dat was het klassiek perspectief, een uit drie perspectieven. Wellicht is echter nog een ander (vierde) perspectief te kiezen. Dat perspectief ondergraaft de hier geformuleerde verwachting. Onder anderen Dubnick (2005) en Swierstra en Tonkens (2008: 121) vertolken dit meer kritische perspectief. Zij zetten kanttekeningen bij een eenvoudige stap van meting naar

benutting en vervolgens kwaliteitsverbetering van het gemetene; misschien treedt er wel kwaliteitsverlaging op, zo stellen zij. Immers, alle bijdragen aan de metingen en ook bezinning op uitkomsten gaat af van het werken aan primaire processen. De betrokkenen die de meting moesten ondergaan, hebben niet veel trek om zich in te laten met de uitkomsten. Wat echt moet verbeteren, hebben betrokkenen vaak zelf al in gang gezet.

Deze gereserveerde houding treft men inderdaad in de praktijk wel hier en daar – en dus niet overall - aan onder docenten, maar deze informatie komt uit het middelbaar beroepsonderwijs, niet uit het w.o.. Een van ons was in 2012 betrokken bij een audit van een Roc in het midden van het land, waar dit verschijnsel zich voordeed ('Een interne audit? Nee, liever niet'). Vanuit dit perspectief van Swierstra en Tonkens wordt dus ook rekening gehouden met machtsverschillen, met verzet.

Bezinning op uitkomsten van visitaties is in dit perspectief van Swierstra en Tonkens zeker niet alleen een cognitief-rationalistische zaak maar meer een zaak van interactie en sentimenten. Beleid is niet alleen een zaak van uitdenken en benutten maar ook van uitvechten en dus onder andere van deprioriteren, negeren, uitstellen, de baas spelen, ketelmuziek maken, symboolpolitiek bedrijven, etc. (zie ook Hoppe, 2008; 2012). Hoe overtuigend dit ook klinkt, dit perspectief kan vooralsnog onze verwachting niet ondergraven. Maar het is goed om rekening te houden met scepsis over het gebruik van visitatierapporten en 'moves' om schijngevolgen voor te wenden (nader te bezien: De Lancer Julness & Holzer, 2001; Landry e.a., 2003; Mark & Henry, 2004; Johnson e.a., 2009). Daarmee raken we al bijna aan het fenomeen van bedrog.

11 Vier kwaliteiten van visitaties

Wie visitatieprocessen bestudeert, zoals in het voorgaande is gebeurd, kan vier kwaliteitsaspecten aan visitaties onderscheiden en daarmee komen tot vier terreinen van kwaliteitsoordelen. We noemen hier die vier kwaliteitsgebieden en zullen verderop de verwachtingen onderbrengen in elk van deze vier onderzoeksgebieden.

De kwaliteitsgebieden bij visiteren die aanleiding zijn voor verder onderzoek zijn:

1. De ontwerpqualiteit van de visitatie.
2. De beoordelingskwaliteit in de vorm van een kader dat de visitatiecommissie hanteert.
3. De proceskwaliteit.
4. De concluderende kwaliteit en bruikbaarheid.

Ad 1 De (visitatie)ontwerpqualiteit heeft betrekking op beschrijving en waardering van doel, opzet en grenzen van het visitatiesysteem en -proces. We rekenen er de beginselen van behoorlijk visiteren toe en dus ook de volgende keuzen op ontwerpvlak.

Is visiteren in een visitatiebestel in de kern beoordelen en dus ook evalueren of is het (op transformatie gericht) meedenken met de gevisiteerde over de toekomstige taken en opgaven (welke dat zijn)? En als het beoordelen is, is het dan beoordelen van de bestaande bestuurlijke opgaven en/of wettelijke (taak)opdrachten die de gevisiteerde moet vervullen (lees: het bestaande beleid)?

Is de visitatie in een bepaalde sector of branche gericht op publieke verantwoording en/of op leren (lees: verbeteren in de vorm van repareren of corrigeren) en/of spelen andere doelstellingen een rol, zoals veronderstelde bijdragen aan een systeeminnovatie? Is in het ontwerp van een visitatieaanpak sprake van een consistente uitwerking in het licht van het centrale oogmerk (dus van verantwoorden, leren of ...)?

Geschied de visitatie in de context van een visitatiebestel van interorganisationele aard (van 'multi-level governance'), bij voorbeeld omdat een minister of ander politiek bestuur, al of niet op wettelijke grondslag, tot visitatie aanzet of verplicht? En/of is een brancheorganisatie bij de visitatieopzet betrokken en verplicht die leden tot deelname (zoals Aedes bij corporaties doet)? Is er een sanctie op niet-deelname aan een visitatie (zoals het vervallen van de bekostiging voor een w.o.-opleiding)? Is er eventueel een positieve prikkel op deelname en een positieve uitkomst (extra budget bij positieve scores)? Is sprake van visitatie van slechts één organisatie of van een gelijktijdige visitatie van een aantal vergelijkbare organisaties (of eenheden, zelfde opleidingen, etc.)? Hoe worden deze aspecten gewaardeerd?

Is sprake van een organisatie die de start van het proces organiseert (bestuurders in wetenschappelijk onderwijs zelf, in VSNU-verband)? Is sprake van een regisseur zodat het stuk gespeeld kan worden, met andere woorden zodat duidelijk is of er een protocol is en een beoordelingskader (denk aan de VSNU)?

Is wel of niet sprake van de aanwezigheid van een visitatiecommissie die wel of niet bestaat uit vakgenoten en onafhankelijk opereert? Is sprake van wel of niet medewerking van de gevisiteerde organisatie aan de dataverzameling en eventueel van een zelfstudie? Is sprake van het wel of niet ter plaatse gaan kijken en spreken met betrokkenen: site visit? Is sprake van het wel of niet halen van informatie van buiten de organisatie over de gevisiteerde als check op de informatieverstrekking door de gevisiteerde? Is er een mogelijkheid aan de gevisiteerde om commentaar te leveren op een conceptrapport? Is het eindrapport wel of niet openbaar?

Kan een visitatierapport 'in de wind wegwaaien'? Of: hoe belangrijk is een positieve of negatieve score in een visitatierapport? Zo ja, voor wie is dit rapport met scores belangrijk? Is wel of niet sprake van een toezichthouder op de kwaliteit van de visitatie, zoals een inspectie?

Het gaat bij deze vragen niet alleen om de beschrijving maar ook om draagvlakaspecten en waardering.

De volgende verwachtingen reken we tot de ontwerp kwaliteit:

- *de congruentieverwachting* (nr. 1);
- *de verwachting van contextafhankelijkheid* (nr. 2);
- *de verwachting van visitatiekenmerken als randvoorwaarde voor succes* (nr. 4);
- *de verwachting van slijtage in de externe kwaliteitszorg en de roep om evolutie* (nr. 15).
- *de verwachting van de begrensde uitdaging van visitatieobjecten* (nr. 16).
- *de verwachting van de begrensde verantwoording* (nr. 17).

Wie deze verwachtingen beziet, ziet tevens dat er meer vragen gesteld zijn dan waarop de verwachtingen betrekking hebben. Anders gesteld, niet voor elke vraag is een verwacht antwoord in de vorm van een verwachte uitkomst geformuleerd. We mogen dan ook aannemen dat een onderzoek naar het visiteren in andere sectoren, zoals van gemeentebesturen, nieuwe verwachtingen zal opleveren.

Zijn deze verwachtingen alle gelijkwaardig of zijn er bij die belangrijker zijn dan andere? We hebben eerder aangegeven dat de congruentieverwachting (nr. 1) voor ons hier belangrijk is (Korsten, 2004a,b). De congruentieverwachting gaat ervan uit dat de kwaliteit van visitatie contextonafhankelijk is en dat als elf kernbeginselen aanwezig zijn in een visitatieopzet een dergelijk visitatiebestel ook deugdelijk en succesvol zal zijn. We hebben echter een contrasterende verwachting geformuleerd (nr. 2). Die houdt in dat visiteren juist wel contextafhankelijk is en dat er 'meerdere wegen naar Rome leiden'. Anders gesteld, ook visitatieprocessen waarbij een of meer beginselen afwezig zijn en eventueel enkele andere zijn

toegevoegd zouden tot een succesvolle visitatie kunnen leiden. Logisch eigenlijk, want in tal van contexten ontbreekt een inspectie die de kwaliteit van de visitatie in meta-evaluatief perspectief beoordeelt. En dan zou elders - volgens de congruentieverwachting - dus geen succesvolle visitatie mogelijk zijn. Dat gaat ver, maar onderzoek moet nader uitwijzen wat kenmerken van visitaties in andere sectoren zijn en of dergelijke visitaties bij het al of niet ontbreken van een of meer (ontwerp-) beginselen toch succesvol zijn. Dat succes hangt niet alleen af van de ontwerp kwaliteit maar ook van de drie andere kwaliteitselementen: de kaderkwaliteit, de proceskwaliteit en de concluderende kwaliteit. Die komen nu achtereenvolgens aan bod.

Ad 2 De beoordelings(kader)kwaliteit slaat op het specifieke beoordelingskader ('frame'), de focus om naar de te visiteren organisatie te kijken.

Vragen zijn in dit verband: Hoe ziet dat beoordelingskader eruit? Gaat het om een ordening naar taakvelden en naar niveaus van besturen, zoals het strategische, tactische en operationele niveau, waarbij vervolgens over elke 'cel' in de vier bij drie matrix een uitspraak gedaan wordt? Of is sprake van een geheel andere indeling om grip te krijgen op de integraliteit van de organisatie? Is het kader gericht op een retrospectief-evaluatieve visitatie (terugblikkend op wat van beleid terecht komt?) en/of op een prospectieve visitatie (gericht op toekomstig bestuur, op de goede dingen doen)? Brengt het beoordelingskader de organisatie al met al wel of niet integraal in beeld?

Als het om een prospectieve evaluatie gaat, zal een beoordelingskader wellicht ontbreken. Maar wat is dan de aard van de meedenkoperatie met de gevisiteerde(n)?

In geval van een (groten)deels retrospectieve visitatie: is het beoordelingskader opgesteld met of zonder steun van de gevisiteerden? Erkennen de visitatiecommissies en de gevisiteerden het kader wel of niet? Is het beoordelingskader in het publieke forum wel of niet omstreden? Ontbreekt er iets in het kader of is het juist veel te omvattend en complex? Houdt het kader (on)voldoende of maar matig rekening met het functioneren van organisaties in een of meer netwerken?

Bevat het kader oordelen over elementen of aspecten van de gevisiteerde organisaties in termen van goed, voldoende, onvoldoende, niet van toepassing?

Is het kader wel of niet begrijpelijk en hanteerbaar voor zowel commissie als gevisiteerden? Is het kader goed geoperationaliseerd en bestaat daarover geen discussie? Houdt de inhoudsopgave van de zelfstudie goed verband met het beoordelingskader zodat de commissie de vertaling naar de inhoudsopgave van een eindrapport eenvoudig kan maken?

Spoort de meetambitie om een integraal beeld van een organisatie te krijgen uiteindelijk wel of niet met de praktijk van visiteren? Worden ambities met visiteren waargemaakt?

Enkele verwachtingen hebben betrekking op kaderkwaliteit? We noemen:

- *de verwachting van het immer uitblijvende integraal kader voor beoordeling* (het 'rupsje-nooit-genoege'-effect; nr. 8);
- *de verwachting van de visitatie-illusie* (het fragmentatie-effect van checklist-denken; nr. 9).

Ook hier geldt weer dat er meer vragen zijn gesteld dan in verwachtingen beantwoord worden.

Ad 3 Bij de (visitatie)proceskwaliteit gaat het om het feitelijk procesverloop van visitatieprocessen.

Lag aan het procesverloop wel of niet een proef ten grondslag? Is die wel of niet bevallen? Heeft die wel of niet tot aanpassing geleid? Voldoen de verschillende elementen in de praktijk:

de fasering, en het optreden van de visitatiecommissie, de kwaliteit van de zelfstudie, de site visit?

Is het procesverloop in de ogen van de visitatiecommissie en de gevisiteerden wel of niet acceptabel? Is het proces qua te nemen stappen wel of te uitgebreid? Is het wel of niet te bureaucratisch doordat het veel inzet vraagt in termen van personeel en tijdsbesteding van gevisiteerden? Is het proces ondoelmatig? Laat het proces strategisch gedrag toe van gevisiteerden en zo ja, welk? Heeft dit ingewerkt op de inhoud van het eindrapport en daarmee op de geldigheid van de analyse en de betrouwbaarheid?

De volgende verwachtingen hebben betrekking op proceskwaliteit bij visiteren.

- *de doelmatigheidsverwachting* (nr. 3);
- *de verwachting van de grote proceskracht van het protocol* (nr. 7);
- *de verwachting van strategisch gedrag bij visiteren* (nr. 10);
- *de verwachting van het inburgeringseffect* (visiteren wordt routine; nr. 11).

Ook nu geldt weer: er zijn meer vragen geformuleerd dan antwoorden zijn voorzien in de vorm van verwachtingen.

Ad 4 De vierde kwaliteit ('concluderende kwaliteit en bruikbaarheid') heeft betrekking op de aanwezigheid van conclusies in het visitatierapport, op de aanbevelende kracht en de erkenning daarvan. Dit kwaliteitsgebied omvat ook wat Van de Vall (1980) ooit de implementaire validiteit noemde, dus de bruikbaarheid.

Hierbij is te denken aan de volgende vragen: Geeft de rapportage over toepassing van het beoordelingskader oordelen over slechts een gevisiteerde organisatie of over meerdere organisaties (zoals vergelijkbare opleidingen aan een aantal universiteiten)? In geval van vergelijking, is sprake van benchmarking en daarmee van 'ranking'?

Is het eindrapport volledig? Is het eindrapport consistent? Zijn de conclusies helder en dus begrijpelijk geformuleerd?

Is sprake van geldige en betrouwbare bevindingen in een intersubjectieve aanpak waarbij de visitatiecommissie (mogelijk op basis van een zelfstudie en site visits en ondervraging) tot een oordeel komt? Is het beeld van de gevisiteerde organisatie in het rapport herkenbaar en wordt het erkend?

Bevat het eindrapport voldoende conclusies voor direct en instrumenteel gebruik in de organisatie? Is het eindrapport bruikbaar voor reflectie? Bevat het eindrapport aanbevelingen of zodanig handelingsgerichte conclusies dat het rapport als bruikbaar voor verbetering kan worden gezien?

Is het visitatierapport (eventueel in conceptvorm) in een of meer ambtelijke arena besproken of niet? Zo ja, op welke wijze? Is het visitatierapport in een bestuurlijke arena besproken of niet? Heeft een dagelijks bestuur wel of geen besluiten genomen over het visitatierapport in de vorm van beleidsvoorstellen? Heeft een algemeen bestuur beleidsvoorstellen naar aanleiding van ene visitatierapport vastgesteld of niet?

Wordt er wel of geen keurmerk toegekend na een positieve beoordeling?

De volgende verwachtingen hebben betrekking op concluderende kwaliteit.

- *de concluderende kracht als succesindicator* (nr. 5);
- *de verwachting van de uitblijvende aanbevelingen* (nr.6);
- *de verwachting van de benuttingsimpuls* (nr. 12);
- *de verwachting van gezwollendheid* (het plofkip-effect; nr. 13);

- *de verwachting van niet-collectief verbeteren* (het 'op-zichzelf'-effect; nr. 14).

Wie de vier kwaliteitsdimensies van visiteren beziet, zal kunnen stellen dat een visitatie positief is die positief scoort op vier gebieden: de ontwerp kwaliteit, de kaderkwaliteit, de proceskwaliteit en de concluderende kwaliteit en bruikbaarheid.

Literatuur

- Abma, K. en A.F.A. Korsten, Burgers beoordelen hun gemeente. Waar zijn gemeenten volgens hun burgers goed in en waarin minder goed?, in: *Tijdschrift voor Public Controlling*, april 2009c, pp. 35-39.
- Abma, K. en A.F.A. Korsten, *Gemeenten in rapportcijfers*, Eburon, Delft, 2009a.
- Abma, K. en A.F.A. Korsten, Normenkaders bij bestuurskrachtmetingen - Een beeld van de ideale gemeente?, in: *Bestuurswetenschappen*, jrg. 62, 2008, nr. 6, pp. 42-67.
- Abma, K., *Beoordelen van gemeenten*, Wolf Legal Publishers, Nijmegen, 2012.
- Abma, K., Bestuurskracht en kwaliteit, in: *Openbaar bestuur*, jrg. 19, nr. 10, oktober 2009b.
- Abma, K., Bestuurskrachtmeting inclusief visitatie, in: A. Eversdijk en J. Schutgens (red.), *Openbaar bestuur op expeditie - Verkenningen van bestuurlijke vernieuwingspogingen* - Liber Amicorum voor prof. dr. Arno F.A. Korsten, Boom Lemma, Den Haag, 2010, pp. 18-39.
- Algemene Rekenkamer, *Bestuur op afstand in beeld*- Tien jaar onderzoek van de Algemene rekenkamer naar instellingen op afstand van het Rijk, Sdu, Den Haag, 2008a.
- Algemene Rekenkamer, *Kwaliteitsbewaking in het hoger onderwijs in Nederland en Vlaanderen*, Verslag aan de Tweede Kamer der Staten-Generaal en het Vlaams Parlement, Sdu Uitgevers, Den Haag, 2008b.
- Algemene Rekenkamer, *Kwaliteitszorg Hoger Onderwijs*, Den Haag, 2000.
- Algemene Rekenkamer, *Onregelmatigheden bekostiging (hoger) onderwijs*, Den Haag, 2003.
- Amaral, A.M.S.C., The US Accreditation System and the CRE's Quality Audits - A Comparative Study, in: *Quality Assurance in Education*, vol. 6, 1998, nr. 4, pp. 184-196.
- Argyris, C., *Leren in en door organisaties*, Scriptum, Schiedam, 1996.
- Argyris, C., *On Organizational Learning*, Blackwell Publishing, Malden, USA, 1999.
- Argyris, C. & D. Schön, *Organizational Learning: A Theory of Action Perspective*, Addison-Wesley, Reading, 1978.
- Askim, J., Determinants of Performance Information Utilization in Political Decision Making, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 125-140.
- Askim, L., A. Johnson & K.A. Christophersen, Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks, in: *Journal of Public Administration Research and Theory*, vol. 18, 2008, nr. 2, pp. 297-320.
- Askim, L., *Local Government by Numbers - Who Makes Use of Performance Information, When and For What Purposes?*, University of Oslo, Oslo, 2007.
- Bagchus, R., *Onderwijs, Cultuur en Wetenschappen - Profiel van een ministerie*, Departementale studies nr. 6, VUGA, Den Haag, 1995.
- Bagchus, R., *Onderwijs, Cultuur en Wetenschappen*, in: Twist, M.J.W. van, R. Bagchus en F.O.M. Verhaak (red.), *Kerndepartementen op afstand?*, Eburon, Delft, 1996, pp. 147-167.
- Behn, R.D., *Rethinking democratic accountability*, Brookings, Washington, 2001.
- Behn, R.D., Why measure performance? - Different Purposes Require Different Measures, in: *Public Administration Review*, 2003, vol. 63, nr. 5, pp. 586-606.
- Bekker, R., *Marathonlopers rond het Binnenhof - Topambtenaren bij het rijk 1970-2010*, Boom/Lemma, 2012.
- Bekkers, V.J.J.M. e.a., *Doorwerking van strategische beleidsadviesing*, Erasmus Universiteit, Rotterdam, 2004.
- Bekkers, V.J.J.M., *Beleid in beweging*, Boom/Lemma, Den Haag, 2007.
- Bekkers, V.J.J.M., M. Fenger en V. Homburg, Doorwerkingspatronen van strategische adviesing, in: *Bestuurswetenschappen*, 2005, nr. 5, pp. 431-450.
- Bekkers, V.J.J.M., *Nieuwe vormen van sturing en informatisering*, Eburon, Delft, 1994.
- Bemelmans-Videc, M.L., B. Perrin & J. Lonsdale (eds.), *Making Accountability Work: Dilemmas for Evaluation and for Audit*, Transaction Publ., Edison (NJ), 2007.

- Benneworth, P., H. de Boer, L. Cremonini, B. Jongbloed, L. Leisyte, H. Vossensteyn & E. de Weert, *Quality related funding, performance agreements and profiling in higher education: An international comparative study*, CHEPS Universiteit Twente, Enschede, 2011.
- Berkel, H. van, De visitatie-illusie, in: *Tijdschrift voor hoger Onderwijs en Management (THEMA)*, 2000, nr. 5, pp. 25-29.
- Beus, J. de, Te veel transparantie werkt niet, laat politici de ruimte om te besturen, in: *NRC*, 24 juni 2006.
- Bogt, H.J. ter, P. Jansen en S. Tillema (red.), *Met het oog op de publieke zaak*, RUG, Groningen, 2010.
- Bogt, H.J. ter, Bronnen van controlinformatie voor politieke bestuurders, in: *Maandblad voor Accountancy en Bedrijfseconomie*, jrg. 75, 2001a, nr. 12, pp. 525-538.
- Bogt, H.J. ter, De doelmatigheid van outputgerichte managementinformatie, in: *Beleidswetenschap*, jrg. 16, 2002, nr. 2, pp. 114-144.
- Bogt, H.J. ter, en J.G. van Helden, Gebruik van outputinformatie bij gemeentelijke overheden, in: *Bestuurskunde*, 1994, nr. 1, pp. 2-13.
- Bogt, H.J. ter, Gebruik van outputinformatie door gemeentebestuurders, in: *Overheidsmanagement*, 2000, nr. 7/8, pp. 183-187.
- Bogt, H.J. ter, *Neo-institutionele economie, management control en verzelfstandiging*, Labyrint, Cappelle a.d. IJssel, 1997.
- Bogt, H.J. ter, Performance evaluation styles in governmental organizations: how do professional managers facilitate politicians' work?, in: *Management Accounting Research*, vol. 14, 2003, nr. 4, pp. 311-332.
- Bogt, H.J. ter, Politicians and output-oriented performance evaluation in municipalities, in: *The European Accounting Review*, vol. 10, 2001b, pp. 621-643.
- Bogt, H.J. ter, Politieke bestuurders, productbegrotingen en prestatiebeoordelingen bij gemeenten, in: *Maandblad voor Accountancy en Bedrijfseconomie*, jrg. 73, 1999, nr. 12, pp. 671-683.
- Bogt, H.J., ter, *Managementvernieuwing bij de overheid - Mooie woorden of echte daden?*, Groningen, 2005 (oratie).
- Bogt, H.J., ter, Prestatiegegevens: nuttig voor sturing van gemeentelijke taken, in: *Public Controlling*, juni 2004, pp. 52-58.
- Boomkens, R., *Topkisch en slow science – Kritiek van de academische rede*, Uitgeverij Van Gennep, Amsterdam, 2008.
- Borst, R., C. Lako en M.S. de Vries, Is prestatie management in de publieke sector toepasbaar? – Een vergelijkend onderzoek naar de opvattingen onder ambtenaren, in: *Bestuurswetenschappen*, jrg. 66, 2012, nr. 4, pp. 14-33.
- Bouckaert, G. & J. Halligan, *Managing performance. International comparisons*, Routledge, London, 2008.
- Bouckaert, G. & J. Halligan, Performance and performance management, in: B. Guy Peters & J. Pierre (eds.), *Handbook of Public Policy*, Sage, Londen, 2006, pp. 443-460.
- Bouckaert, G. & W. van Dooren, Performance Measurement and Management in Public Sector Organizations, in: Boivard, T. & E. Löffler (eds.), *Public Management and governance*, Routledge, Londen, 2004, pp. 127-137.
- Bouckaert, G., B. De Peuter en W. van Dooren, *Met en vergelijken van lokale bestuurlijke ontwikkeling: een monitoringsysteem voor het lokaal bestuur in Vlaanderen*, Die Keure, Brugge, 2003.
- Bouckaert, G., Prestaties en prestatie management in de publieke sector, in: *Tijdschrift voor Economie en Management*, 2006, nr. 3, pp. 237-265.
- Bouckaert, G., Renewing Public Sector Performance Management, in: Jann, W., M. Röber & H. Wollmann (Hrsg.), *Public Management: Grundlagen, Wirkungen, Kritik*, Sigma, Berlijn, 2006, pp. 119-133.
- Bovens, M. en Th. Schillemans (red.), *Handboek publieke verantwoording*, Lemma, Den Haag, 2009.
- Bressers, J.Th.A., P. de Jong, P.J. Klok en A.F.A. Korsten (red.), *Beleidsinstrumenten bestuurskundig beschouwd*, Van Gorcum, Assen, 1993.
- Brink, G. van den, Th. Jansen en D. Pessers (red.), *Beroepszeer- Waarom Nederland niet goed werkt*, Boom, Amsterdam, 2005.
- Bruijn, H. de, en E. ten Heuvelhof, *Handhaving- Het spel tussen inspecteur en inspectee*, Lemma, Utrecht, 2005.
- Bruijn, H. de, en G.J. van Helden, A Pea for Dialogue Driven Performance-based Management Systems: Balance from the Dutch Public Sector, in: *Financial Accountability & Management*, vol. 22, 2006, nr. 4, pp. 405-423.

Bruijn, H. de, en J. van Helden, Effectief prestatie management bij professionele organisaties in de publieke sector, in: *Overheidsmanagement*, febr. 2007, nr. 1, pp. 20-26.

Bruijn, H. de, *Framing- Over de macht van taal in de politiek*, Atlas, Amsterdam, 2011.

Bruijn, H. de, *Managers & professionals*, Academic Service/Sdu, Den Haag, 2008.

Bruijn, H. de, *Managing Performance in the Public Sector*, Routledge, Londen, 2007.

Bruijn, H. de, *Prestatiemeting in de publieke sector*, Lemma, Utrecht, 2001.

Bruinsma, F., De ondraaglijke lichtheid van de rechtenstudie, in: *Nederlands Juristenblad*, jrg. 28, 2000, pp. 1371-1374.

Burg, W. van der Burg, De regulering van professionals: twee botsende perspectieven, in: Jansen, Th., G. van den Brink en J. Kole (red.), *Beroepstrots- Een ongekende kracht*, Boom, Amsterdam, 2009, pp. 168-182.

Camps, Th.W.A., *Het besloten concilie der deskundigen*, Lemma, Utrecht, 1995 (diss.).

Caniëls, M.C.J., *Lerende organisaties in de kenniseconomie*, Open Universiteit, Heerlen, 2011.

Caplan, N., The Two-Communities Theory and Knowledge Utilization, in: *American Behavioral Scientist*, vol. 22, 1979, nr. 3.

Carter, R., Maximizing the Use of Evaluation Results, in: Wholey, J.S., H.P. Hatry & K.E. Newcomer (eds.), *Handbook of Program Evaluation*, Jossey-Bass Publ., San Francisco, 1994, pp. 576-590.

Casey, D., *Managing Learning in Organizations*, Open University Press, Buckingham, 1993.

Charbonneau, E. & F. Bellavance, Blame Avoidance in Public Reporting: Evidence from a Provincially Mandated Municipal Performance Measurement Regime, in: *Public Performance & Management Review*, vol. 35, march 2012, nr. 3, pp. 399-422.

Commissie Accreditatie Hoger Onderwijs, *Prikkelen, presteren, profileren*, Eindrapport, Den Haag, 2001.

Commissie Diplomerings Langstudeerders Hogeschool INHolland (commissie-Leers), *Veel ruimte, weinig rekenschap*, Haarlem, september 2010.

Commissie Toekomst Lokaal Bestuur (Cie-Bovens), *Wil tot Verschil - Gemeenten in 2015*, Vereniging van Nederlandse Gemeenten, Den Haag, 2006.

Commissie-Korsten, *De kunst van het balanceren - Visitatierapport hbo-opleidingen bestuurskunde en overheidsmanagement*, Den Haag, 2003.

Commissie-Korsten, *Ontwikkeling door visie - Midterm review-commissie opleidingen bestuurskunde aan de Erasmus Universiteit*, Rotterdam, 2007.

Daemen, H. en Th. Van der Krogt, Vier functies van internationale accreditatie, in: Kickert, W.J.M., L. Cachet, F.B. van der Meer en L. Schaap, *Liefde voor het openbaar bestuur en liefde voor de bestuurskunde*, Eburon, Delft, 2007, pp. 105-119.

Dal Santo, T., S. Goldberg, P. Choice & M.J. Austin, Exploratory Research in Public Social Service Agencies – An Assessment of Dissemination and Utilization, in: *Journal of Sociology and Social Welfare*, vol. 29, 2002, nr. 4, pp. 59-81.

Denhardt, K.D. & M.P. Aristigueta, Performance Management Systems: Providing Accountability and Challenging Collaboration, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 106-125.

Derksen, W., A.F.A. Korsten en A.F.M. Bertrand (red.), *De praktijk van onderzoek*, Wolters-Noordhoff, Groningen, 1988.

Derksen, W., *Kennis en beleid verbinden*, Boom/Lemma, Den Haag, 2011.

Dijk, A. van, en H. Woldendorp, Sturen in complexe organisaties, in: A. van Dijk, G. Smid en H. Woldendorp (red.), *Sturen in complexe organisaties- Theorie en praktijk van het hoger onderwijs*, Assen, 1993, pp. 11-21.

Dill, D.D., Accreditation, Assessment, Anarchy? The Evolution of Academic Quality Assurance in the US, in: Brennan, J. e.a., *Standards and Qualities in Higher Education*, Jessica Kingsley Publ., Londen, 1997, pp. 15-44.

Dittrich, K., *Accreditatie en de mogelijke gevolgen ervan*, speech HvA/UvA, 21 okt. 2003.

Dittrich, K., Het nieuwe Nederlandse accreditatiestelsel, in: TH&MA- *Tijdschrift voor Hoger onderwijs en Management*, 2010, nr. 4, pp. 23-26.

Doelen, F.C.J. van der, Legitiem en effectief overheidsbeleid: geven en nemen, in: *Beleid en Maatschappij*, jrg. 18, 1991, nr. 2, pp. 90-100.

Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008.

Dooren, W. van, G. Bouckaert & J. Halligan, *Performance Management in the Public Sector*, Routledge, Londen, 2010.

- Dooren, W. van, *Performance Measurement in the Flemish Public Sector*, Instituut voor de Overheid, Leuven, 2006.
- Dorsch, J.J. & M.M. Yasin, A Framework for Benchmarking in the Public Sector, in: *International Journal of Public Sector Management*, vol. 11, 1998, nr. 2/3, pp. 91-115.
- Douma, Th., *Accreditatie: een (zelf)diagnose - Beschouwing over de werking van het stelsel*, Sdu, Den Haag, 2004.
- Drooglever Fortuijn, E., *Onderwijsbeleid – Maatschappelijke functies en strategische keuzen*, Aksant, Amsterdam, 2003.
- Dubnick, M., Accountability and the Promise of Performance, in Search of the Mechanisms, in: *Public Performance and Management Review*, vol. 28, 2005, nr. 3, pp. 376-417.
- Ehren, M., Het spel der neveneffecten. Strategisch gedrag van toezichthouders en scholen, in: Leeuw, F.L., J.S. Kerseboom en R. Elte (red.), *Turven, tellen toetsen*, Boom Juridische uitgevers, Den Haag, 2007, pp. 195-203.
- Ellemers, J.E., Veel kunnen verklaren of iets kunnen veranderen: krachtige versus manipuleerbare variabelen, in: *Beleid en Maatschappij*, nov. 1976, pp. 281-291.
- Elsacker, W.J., *Roles of performance management in local government*, EPM Sneek, Groningen, 2007.
- Elte, R., Het ministerie van Onderwijs en Wetenschappen, in: Breunese, J. en L.J. Roborgh (red.), *Ministeries van algemeen bestuur*, Spruyt, Van Mantgem en De Does, Leiden, 1989, pp. 123-139.
- Elte, R., Winst- en verliesrekening van bureaucratisering van het toezicht op onderwijs, in: Leeuw, F.L., J.S. Kerseboom en R. Elte (red.), *Turven, tellen toetsen*, Boom Juridische uitgevers, Den Haag, 2007, pp. 179-193.
- Engbersen, G., Ontmantel de visitatie-industrie, in: *Beleid en Maatschappij*, 2010, nr. 2, pp. 161-164.
- Engbersen, G., *Fatale remedies – Over onbedoelde gevolgen van beleid en kennis*, Pallas Publications, Amsterdam University Press, Amsterdam, 2009.
- Engelen, E., De manager heeft gewonnen, in: *NRC*, 18 juli 2012, p. 15.
- Erasmus Universiteit Rotterdam– Rijksuniversiteit Leiden, *Zelfstudie Bestuurskunde ten behoeve van Visitatie onderwijs 1997*, Rotterdam/Leiden,
- Fijnaut, C., H. Boutellier e.a., *Veel te winnen- Visitatiecommissie Veiligheid Eindhoven*, Gemeente Eindhoven, 2008.
- Foppen J.W., *Gistend beleid – Veertig jaar universitaire onderwijspolitiek*, Vuga, Den Haag, 1989.
- Frederiks, M.M.H., *Beslissen over kwaliteit: verklaring van besluitvorming over het gebruik van onderwijsvisitatierapporten in Nederlandse universiteiten*, CHEPS/De Tijdstroom, Enschede, 1996.
- Frey, B.S. & M. Osterloh, Gaat de wetenschap ten onder aan de beoordeelziekte evaluïtis?, in: *NRC*, 29 november 2008.
- Frey, B.S. & M. Osterloh, Onderzoeksevaluaties: verborgen kosten, twijfelachtige voordelen en betere alternatieven, in: Jansen, Th., G. van den Brink en J. Kole (red.), *Beroepstrots- Een ongekende kracht*, Boom, Amsterdam, 2009, pp. 194-219.
- Friedrich, H., *Staatliche Verwaltung und Wissenschaft*, Europäische Verlagsanstalt, Frankfurt am Main, 1970.
- Gardner, H., M. Csikszentmihalyi & W. Damon, Het GoodWork Project®: enkele resultaten van vijftien jaar onderzoek, in: Jansen, Th., G. van den Brink en J. Kole (red.), *Beroepstrots- Een ongekende kracht*, Boom, Amsterdam, 2009, pp. 48-77.
- Garré, P., Externe verantwoording, kwaliteitsverbetering en –verankering in het hoger onderwijs. Accreditering: hefboom of keurslijf, droom of illusie?, in: *Tijdschrift voor Onderwijsrecht en Onderwijsbeleid*, sept.-okt. 2002, nr. 1, pp. 53-60.
- Goedegebuure, L., O.A.M. Maassen en D.F. Westerheijden (eds.), *Peer Review and Performance Indicators – Quality Assessment in British and Dutch Higher Education*, Cheps, Enschede, 1990.
- Goossens, J., P. Fröhling en M. Laman, *De taakuitvoering van visitatiecommissies in het hoger onderwijs 1999-2001*, Inspectie voor het Onderwijs, mei 2001 (concept).
- Greener, I., *Public Management – A Critical Text*, Palgrave MacMillan, Londen, 2009.
- Greve, H., *Organizational learning from performance feedback: A behavioral perspective on innovation and change*, Cambridge University Press, Cambridge, 2003.
- Grinten, T. van der, en P.L. Meurs, Publieke verantwoording in de gezondheidszorg, in: Bakker, W. en K. Yesilkagit (red.), *Publieke verantwoording*, Regimes van inzicht en rekenschap bij de uitvoering van publieke taken, Boom, Amsterdam, 2005.
- Grit, K., *Economisering als probleem – Een studie naar de bedrijfsmatige stad en de ondernemende universiteit*, Van Gorcum, Assen, 2000.

Grooteman, E., Interdepartementale beleidsvisitaties – Instrument voor het evalueren van beleidsontwikkeling, in: Knaap, P. van der, en A. Schilder (red.), *Resultaatgericht sturen en evalueren*, Sdu, Den Haag, 2004, pp. 215-222.

Habermas, J., *Theorie des kommunikativen Handelns*, Surhkamp, Frankfurt, 1987.

Hakvoort, J. en H. Klaassen, Benchmarking: resultaten vergelijken, in: Knaap, P. van der, en A. Schilder (red.), *Resultaatgericht sturen en evalueren*, Sdu, Den Haag, 2004, pp. 205-215.

Halachmi, A., Performance Measurement, Accountability, and Improved Performance, in: *Public Performance and Management Review*, vol. 25, 2002b, nr. 4, pp. 370-374.

Halachmi, A., Who Gets What, When and How: Performance Measures for Accountability?; for Improved Performance?, in: *International Review of Public Administration*, vol. 7, 2002a, nr. 1, pp. 85-95.

Hanney, S.R., M.A. Gonzalez-Blok, M.J. Buxton & M. Kogan, The utilization of health research in policy-making: concepts, examples and methods of assessment, in: *Health Research and Policy Systems*, vol. 1, 2003, nr. 3, pp. 1-28.

Hart, P. 't, J. Hufen en M.J. van Duin, De lerende overhead: mogelijkheden en grenzen van een modieuze metafoor, in: *Beleid en Maatschappij*, jrg. 15, 1988, nr. 2, pp. 83-103.

Haselbekke, A.G.J., H.L. Klaassen, A.P. Ros en R.J. in 't Veld, *Prestaties tellen – Kengetallen als instrument voor een bedrijfsmatig(er) bestuur en beheer van decentrale overheden*, VNG Uitgeverij, Den Haag, 1990.

Hatry, H., Epilogue: The Many Faces of Use, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 227-241.

Hatry, H., Performance Measurement Principles and Techniques: An Overview for Local Government, in: Kearny, R. & E.V. Berman (ed.), *Public Sector Performance – Management, Motivation and Measurement*, ASPA, Westview Press, Boulder, 1999, pp. 304-329.

Haverland, M. en D. Yanow, Hoe overleef ik discussies over bestuurskundig onderzoek?, in: *Bestuurskunde*, 2010, nr. 4, jrg. 19, pp. 87-95.

Heiber, H., *Messung von Forschungsleistungen der Hochschulen*, Nomos Verlagsgesellschaft, Baden-Baden, 1983.

Helden, G. van, *BBI in de praktijk*, Groningen, 1999.

Helden, G.J. van, & A. Johnson, A comparative analysis in the development of performance-based management systems in Dutch and Norwegian local government, in: *International Public Management Journal*, vol. 5, 2002, nr. 1, pp. 75-95.

Helden, G.J. van, & H.J. ter Bogt, The Application of Businesslike Planning & Control in Local government: A Field Study of Eight Dutch Municipalities, in: *Local Government Studies*, vol. 27, 2001, nr. 1, pp. 61-86.

Helden, G.J. van, en S. Tillema, In Search of a Benchmarking Theory for the Public Sector, in: *Financial Accountability & Management*, vol. 21, 2005, nr. 3, pp. 237-261.

Helden, G.J. van, en E. Jansen, New public management bij gemeenten, in: *Overheidsmanagement*, jrg. 15, 2002, nr. 3, pp. 64-68.

Helden, G.J. van, en N. Brouwer, Benchmarking en prestatieverbetering in de publieke sector: ervaringen met de bedrijfsvergelijking zuiveringsbeheer, in: *MAB*, jrg. 79, 2005, nr. 11, pp. 574-583.

Helden, G.J. van, en S. Tillema, Naar een benchmarkingtheorie voor de publieke sector, in: *Tijdschrift voor Openbare Financien*, jrg. 36, 2004, nr. 3, pp. 99-113.

Helden, G.J. van, Normering prestatie-indicatoren, begin- of eindpunt van doelmatigheid, in: *B&G*, september 1997, pp. 8-12.

Helden, G.J. van, S. Tillema, M. Kuppens en J.W.C. Dekking, Werkt benchmarking bij waterschappen?, in: *Bestuurskunde*, jrg. 14, maart 2005, nr. 2, pp. 30-39.

Helderman, J.K. en M.E. Honingh, *Systeemtoezicht- Een onderzoek naar de condities*

Henry, G. & M. Mark, Beyond use: Understanding Evaluation's Influence on Attitudes and actions, in: *American Journal of Evaluation*, vol. 24, 2003, nr. 3.

Hill, M., *The Public Policy Process*, Pearson, Harlow, 2009 (5^e editie).

Hoesel, P. van, *Beleidsonderzoek als professie- Een lang gekoesterd ideaal*, Stichting Economisch Instituut voor het Midden- en Kleinbedrijf, Zoetermeer, 2003 (oratie)

Hogwood, B.W. & B.G. Peters, *Policy Dynamics*, Wheatsheaf Books Ltd., Brighton, 1983.

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

Hood, C., Public Service Management by Numbers: Why Does it Vary? Where has it Come From? What are the Gaps and the Puzzles?, in: *Public Money and Management*, vol. 27, 2007a, nr. 2, pp. 95-102.

Hood, C., R. Dixon & C. Beeston, Rating the Rankings: Assessing International Rankings of Public Service Performance, in: *International Public Management Journal*, vol. 11, 2008, nr. 3, pp. 298-328.

Hood, C., *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government*, Princeton University Press, Princeton, 2011.

Hood, C., What Happens when Transparency meets Blame-Avoidance?, in: *Public Management Review*, vol. 9, 2007b, nr. 2, pp. 191-210.

Hoppe, R. e.a., *Beleidsnota's die (door)werken*, Coutinho, Bussum, 1998.

Hoppe, R. en A. Peterse (red.), *Bouwstenen voor argumentatieve beleidsanalyse*, Elsevier, Den Haag, 1998.

Hoppe, R. en S. Huijs, *Werk op de grens tussen wetenschap en beleid: paradoxen en dilemma's*, RMNO, Den Haag, 2003.

Hoppe, R. en W. Halffman, Wetenschappelijke beleidsadvisering in Nederland. Trends en ontwikkelingen, in: *Beleidswetenschap*, jrg. 18, 2004, nr. 1, pp. 31-61.

Hoppe, R., 'Anorexia consulta': afslanking infrastructuur rijksdienst, deel 2, in: *Beleid en Maatschappij*, jrg. 34, 2007, nr.4, pp. 238-250.

Hoppe, R., From 'knowledge use' to 'boundary work'; sketch of an emerging new agenda for inquiry into science-policy interaction, in: Veld, R.J. in 't (ed.), *Knowledge Democracy: Consequences for Science, Politics and Media*, Heidelberg, 2010, pp. 169-186.

Hoppe, R., Na 'doorwerking' naar 'grenzenwerk', in: *Bestuurskunde*, 2008, nr. 2, pp. 15-26.

Hoppe, R., *The Governance of Problems – Puzzling, Powering, Participation*, Policy Press, Bristol, 2011.

Hoppe, R., *Van flipperkast naar grensverkeer- Veranderende visies op de relatie tussen wetenschap en beleid*, AWT, Den Haag, 2002.

Hoppe, R., Waarom uitvechten uitdenken overtroeft, in: *Bestuurskunde*, jrg. 21, 2012, nr. 2, pp. 7-18.

Hoppe, R., Werken op de grens tussen wetenschap en politiek: naar een typologie van grensarrangementen, in: *Beleidswetenschap*, jrg. 17, 2003, nr. 2, pp. 144-170.

Huisman, H., Hoe verander je de stijl van management bij resultaatgerichte organisaties?, in: *Bestuurskunde*, jrg. 13, nov. 2004, nr. 7, pp. 301-305.

Huisman, J. & H. Theisens, Marktwerking in het hoger onderwijs, in: *Beleid & Maatschappij*, jrg. 28, 2001, nr. 3, pp. 165-177.

Huisman, J. en F. de Vijlder, Hoger onderwijs en beleid in Nederland, in: *Bestuurskunde*, jrg. 8, 1999, nr. 8, pp. 357-367.

Huisman, J., De regulering van het opleidingsaanbod. Een slingerbeweging tussen overheidsplanning en zelfregulering, in: *Beleidswetenschap*, jrg. 11, 1997, nr. 2, pp. 122-142.

Huisman, J., *Differentiation, diversity and dependency in higher education*, CHEPS, Utrecht, 1995.

Huisman, P. en F. de Vijlder, *Sectorstudie toezicht hoger onderwijs*, WRR, Den Haag, 2012 (Webpublicatie 64).

Huitema, D., M. Jeliaskova en D.F.Westerheijden, Fasen, niveaus en cirkels bij beleidsontwikkeling- Hogeronderwijsbeleid en kwaliteitsinterne milieuzorg, in: *Beleidswetenschap*, jrg. 17, 2003, nr. 1, 29-51.

Hulshof, M., J. Goossens en J. Meulesteen, Een archeologisch punt ... De kwaliteit van visitatierapporten, in: Hulshof, M., B. Bulder e.a. (red.), *Toezichthouder: lam of leeuw?*, Inspectie van het Onderwijs, Den Haag, 2003, pp. 21-37.

Hupe, P.L. en F.B. van der Meer, Cruciale koppelingen. Over doorwerking van emancipatie-effectrapportages in overheidsprocessen, in: *Beleidswetenschap*, jrg. 20, 2006, nr. 2, pp. 3-21.

Hupe, P.L., De mythe van de beleidsrelevantie, in: *Beleid en Maatschappij*, oktober 1984, pp. 278-294.

Hupe, P.L., *Overheidsbeleid als politiek*, Van Gorcum, Assen, 2007.

Huysman, M. en R. van der Vlist, Naar een 'organizational learning'-benadering van de lerende organisatie, in: *Gedrag & Organisatie*, jrg. 11, okt. 1998, nr. 5, pp. 219-232.

Inspectie van het Onderwijs, *Onderwijsverslag 2000*, Utrecht, 2001.

Jansen, Th., G. van den Brink en J. Kole (red.), *Beroepstrots- Een ongekende kracht*, Boom, Amsterdam, 2009.

Jansen, Th., G. van den Brink en R. Kneyber (red.), *Gezagsdragers – De publieke zaak op zoek naar haar verdedigers*, Boom, Amsterdam, 2012.

Jeliaskova, M.I., Running the maze: interpreting external review recommendations, in: *Quality of Higher Education*, vol. 8, 2002, nr. 1, pp. 89-96.

Jeliaskova, M. & D. Westerheijden, The Netherlands: A Leader in Quality Assurance Follows the Accreditation Trend, in: Schwarz, S. en D. Westerheijden (eds.), *Accreditation and Evaluation in the European Higher Education Area*, Kluwer Academic Publ., Dordrecht, 2004, pp. 323-346.

- Jeliazkova, M. & D.F. Westerheijden (red.), *Het zichtbare eindresultaat. Doorwerking van onderwijsvisitaties in hogeronderwijsinstellingen*, Algemene Rekenkamer, Den Haag, 17 februari 2000.
- Jeliazkova, M., & D.F. Westerheijden, Systemic adaptation to a changing environment- Towards a next generation of quality assurance models, in: *Higher Education*, vol. 44, 2002, nr. 3-4, pp. 433-448.
- Jenniskens, I.G.M., *Governmental Steering and Curriculum Innovation – A Comparative Study of the Relationship between Governmental Steering Instruments and Innovation in Higher Education Curricula*, Elsevier/De Tijdstroom, Maarssen, 1997.
- Jensma, F., Interview met Grahame Lock 'De meeste mensen kunnen ook toe met minder weten', in: *NRC*, 23 febr. 2008.
- Johnson, K., L.O. Greenesid, S.A. Toial, J.A.King, F. Lawrenz & B. Volkov, Research on Evaluation Use – A Review of the Empirical Literature from 1986 to 2005, in: *American Journal of Evaluation*, 2009, pp. 307-410.
- Jongbloed, B.W.A., Prestatiesubsidiëring in het hoger onderwijs: een internationale vergelijking, in: *Openbare Uitgaven*, 1999, nr. 4, pp. 212-221.
- Karsten, S., A.B. Dijkstra en R.J.Bosker, Openbaarmaking van prestatiegegevens in het basis- en voortgezet onderwijs- Een balans na zeven jaar, in: *Pedagogische Studiën*, jrg. 81, 2004, pp. 397-403.
- Katholieke Universiteit Brabant, Faculteit der Sociale Wetenschappen, *Zelfstudie opleiding Beleids- en Organisatiewetenschappen ten behoeve van de Onderwijsvisitatie Bestuurs- en Beleidswetenschappen in 1997*, Tilburg, 1997.
- Kearny, R. & E.V. Berman (ed.), *Public Sector Performance – Management, Motivation and Measurement*, ASPA, Westview Press, Boulder, 1999.
- Kerseboom, J., F. Leeuw en R. Elte, Een mozaïek van toezicht, inspectie, handhaving en evaluatie in tien thema's, in: Leeuw, F.L., J.S. Kerseboom en R. Elte (red.), *Turven, tellen, toetsen*, Boom, Den Haag, 2007, pp. 9-29.
- Kerseboom, J., Leren van evalueren, in: *Bestuurskunde*, jrg. 17, 2008, nr. 2, pp.41-47.
- Kingdon, J.W., *Agendas, Alternatives and Public Policies*, Longman, New York, 2003
- Kirkhart, K., Reconceptualizing evaluation use: an integrated theory of influence, in: *New Directions for Evaluation*, Jossey-Bass, San Francisco, 2000, pp. 5-23.
- Kluijtmans, F. (red.), *Leerboek HRM*, Noordhoff Uitgevers, Groningen, 2010.
- Knaap, P. van der, *'Lerende' overheid, intelligent beleid*, Phaedrus, Den Haag, 1997.
- Knaap, P. van der, A. Korsten, K. Termeer en M. van Twist (red.), *Trajectmanagement*, Lemma, Utrecht, 2004.
- Knippenberg en W. van der Ham, *Een bron van aanhoudende zorg- 75 Jaar ministerie van Onderwijs (Kunsten) en Wetenschappen: 1918-1993*, Van Gorcum, Assen, 1994.
- Köbben, A.J.F. en H. Tromp, *De onwelkome boodschap- Of hoe de vrijheid van wetenschap bedreigd wordt*, Uitgeverij Jan Mets, Amsterdam, 1999.
- Köbben, A.J.F., *Het gevecht met de engel - Over verheffende en minder verheffende aspecten van het wetenschapsbedrijf*, Mets en Schilt, Amsterdam, 2003.
- Kool, D. de, *Monitoring in beeld*, Erasmus Universiteit Rotterdam, Rotterdam, 2007a.
- Kool, D. de, Monitoring in de publieke sector: De rapportage Water in beeld, in: *Bestuurswetenschappen*, 2011, nr. 1, pp. 21-39.
- Kool, D. de, Monitoring in het openbaar bestuur, in: *Overheidsmanagement*, nr. 10, 2007b, pp. 6-10.
- Kool, D. de, Monitoring in beeld, in: *Bestuurskunde*, 2007c, nr. 2, pp. 92-103.
- Kool, D. de, Rational, Political and Cultural Uses of Performance Monitors: the Case of Dutch Urban policy, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 174-192.
- Korsten, A.F.A. en M. Schoenmaker, Gezond bestuur en bestuurlijke probleemgemeenten: Tijd voor de bestuurschout?, in: *Bestuurswetenschappen*, jrg. 65, 2011, nr. 3, pp. 28-49.
- Korsten, A.F.A. en P. de Goede (red.), *Bouwen aan vertrouwen in het openbaar bestuur*, Elsevier, Den Haag, 2006d.
- Korsten, A.F.A. en R. Hoppe, Van beleidswetenschap naar kennissamenleving, in: *Beleidswetenschap*, jrg. 20, 2006e, nr. 4, pp. 34-73.
- Korsten, A.F.A. en Th.A.J. Toonen (red.), *Bestuurskunde - Hoofdfiguren en kernthema's*, Stenfert Kroese, Leiden, 1993.
- Korsten, A.F.A., Beoordeling - De betekenis van beleidsnetwerken voor beoordelen van provinciale visitatiecommissies, in: Cörvers, R. en C. Dieperink (red.), *Milieubeleid als sturingsopgave*, Open Universiteit/Universiteit Utrecht, 2011, pp. 201-211.

- Korsten, A.F.A., *Bestuurskunde als avontuur*, Kluwer, Deventer, 1988.
- Korsten, A.F.A., *Deugdelijk bestuur*, Open Universiteit Nederland, Heerlen, 2010b (afscheidsrede).
- Korsten, A.F.A., *Grote klasse - Op zoek naar excellente ambtenaren en leiderschap*, Beljon + Westerterp, Heerlen, 2005a.
- Korsten, A.F.A., *Maatstaven voor de beoordeling van overheidshandelen*, Open Universiteit, Heerlen, 14 maart 2005b (www.arnokorsten.nl)
- Korsten, A.F.A., Visitatie goed instrument voor continue verbetering, in: *Provincie en Gemeente*, jrg.55, mei 2004c, nr. 4, pp. 3-5.
- Korsten, A.F.A., Visiteren van gemeentebesturen, in: *Bestuurswetenschappen*, jrg. 58, aug. 2004bd, nr. 4, pp. 305-324.
- Korsten, A.F.A., *Visiteren van gemeenten - Evaluatie van de toepassing van de bestuurskrachtmonitor in Limburgse gemeenten*, Open Universiteit Nederland, Heerlen/Maastricht, 2004a.
- Korsten, A.F.A., Visiteren, in: *Binnenlands Bestuur*, 2003a, nr. 29-30, p.21.
- Korsten, A.F.A., *Wat is goed genoeg?*, Kobra, Amsterdam, 1983.
- Kuhlmann, S., J. Bögumil & H. Wollmann (Hrsg.), *Leistungsmessung und -vergleich in Politik und Verwaltung*, VS Verlag, Wiesbaden, 2004.
- Kwikkers, P.C., D.F. Westerheijden, N. Verduyts, E.V. Schalkwijk, C.J. Peels en J.A. Frederik, *Accreditatie en kwaliteitszorg, De accreditatie van hoger onderwijs in Nederland en Vlaanderen*, Sdu Uitgevers, Den Haag, 2011.
- Mark, M.M. & H.T. Henry, The Mechanisms and Outcomes of Evaluation Influence, in: *Evaluation*, vol. 10, 2004, nr. 1, pp. 35-57.
- Laegreid, P. P.G. Roness & K. Rubecksen, Performance Information and Performance Steering: Integrated System or Loose Coupling?, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 42-58.
- Laman, M., Façades van Potemkin? – Kwaliteitsverbetering door visitaties, in: Hulshof, M., B. Bulder e.a. (red.), *Toezichthouder: lam of leeuw?*, Inspectie van het Onderwijs, Den Haag, 2003, pp. 37-53.
- Lancer Julnes, P. de, Performance Measurement Beyond Instrumental Use, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 58-72.
- Lancer Julnes, P. de & M. Holzer, Promoting the Utilization of Performance Measures in Public Organizations: an Empirical Study of Factor Affecting Adoption and Implementation, in: *Public Administration Review*, vol. 61, 2001, nr. 6, pp. 693-705.
- Landry, R., M. Lamary & N. Amara, The Extent and Determinants of the Utilization of University Research in government Agencies, in: *Public Administration Review*, vol. 63, 2003, nr. 2, pp. 192-205.
- Leeuw, F.L. en H. van de Graaf, Onderzoek naar beleidstheorieën: uitgangspunten, methodische aspecten en praktische relevantie, in: *Beleid en Maatschappij*, jrg. 15, 1988, nr. 1, pp. 1-7.
- Leeuw, F.L., Diminishing returns in evaluation? - Challenges for educational evaluators, in: *Lifelong Learning in Europe*, vol. 9, 2004, nr. 2, pp. 106-114.
- Leeuw, F.L., J.S. Kerseboom en R. Elte (red.), *Turven, tellen, toetsen*, Boom, Den Haag, 2007.
- Leeuw, F.L., R. Rist & R.C. Sonnichsen (eds.), *Can Governments Learn?*, Transaction books, Lexington, 1994.
- Leeuw, F.L., Reconstructing Program Theories: Methods available and Problems to be Solved, in: *American Journal of Evaluation*, vol. 24, 2003, nr.1, pp. 5-20.
- Leeuw, F.L., Toezicht in het hoger onderwijs- Denkkaders voor een woelige tijd, in: *Th&Ma - Tijdschrift voor Hoger Onderwijs en Management*, 2003, nr. 3, pp. 18-23.
- Leeuw, F.L., Visitatie, inspectie en accreditering in het hoger onderwijs, in: *TH&MA- Tijdschrift voor Hoger Onderwijs en Management*, 2001, nr. 5, pp. 44-49.
- Lock, G., De onmenselijke maat en de 'crisis in het onderwijs', Inleiding op het symposium van het BON, 4 april 2009.
- Lock, G., The Collectivisation of Dutch Universities, in: *Minerva*, vol. 27, 1989, nr. 2/3, pp. 157-176.
- Lips, A.M.B., *Autonomie in kwaliteit, Ambigüiteit in bestuurlijke communicatie over de ontwikkeling van kwaliteitszorg in het Hoger Onderwijs*, Eburon, Delft, 1996.
- Luts, M., W. Van Dooren en G. Bouckaert, *Internationale rangschikkingen gerangschikt – Een meta-analyse van rangschikkingen van publieke sectoren*, Steunpunt beleidsrelevant onderzoek- bestuurlijke organisatie Vlaanderen, Leuven, 2008.
- Maarse, J.A.M., *Marktvorming in de zorg*, Maastricht University Press, Maastricht, 2011.

- Mark, M. & G. Henry, The Mechanisms and Outcomes of Evaluation Influence, in: *Evaluation*, vol. 10, 2004, nr. 1, pp. 33-57.
- McDaniel, O.C., *The effects of government policies on higher education- In search of alternative steering methods*, Vuga, Den Haag, 1997.
- Meer, F.B. van der, G.J. de Vries en G.A.N. Vissers, Evaluatie en leerprocessen bij de overheid: de rol van institutionele condities, in: *Beleidswetenschap*, jrg. 14, 2000, nr. 3, pp. 253-277.
- Meurs, P.L., E.K. Schrijvers en G.H. de Vries, *Leren van de praktijk – Gebruik van lokale kennis en ervaring voor beleid*, Amsterdam University Press, Amsterdam, 2006.
- Meijerink, R., Een gulden eeuw voor universitair onderwijs?, in: *Bestuurskunde*, jrg. 8, 1999, nr. 8, pp. 377-385.
- Mertens, F., *Inspecteren - Toezicht door inspecties*, Sdu Uitgevers, Den Haag, 2011.
- Mertens, F.J.H., Kwaliteitsbewaking in het Nederlands hoger onderwijs: enkele persistente problemen van het Nederlands hoger onderwijs, in: *Tijdschrift voor Hoger Onderwijs*, jrg. 17, april 1999, nr. 1.
- Mertens, F.J.H., Vertrouwen en controle in het hoger onderwijs, in: Gansewinkel, H., L. Römken en W. Spee (red.), *Complexiteit. Hoera! Kansen en krachten in het hedendaagse onderwijs*, Interstudie NDO, Arnhem, 2008, pp. 103-110.
- Ministerie van O&W, *Hoger onderwijs: autonomie en kwaliteit*, TK 1985-1986, nr. 19253, nr. 1-2.
- Ministerie van OC &W, *Bestuurlijke hantering kwaliteitszorg*. Brief aan het bestuur van de HBO-Raad en de VSNU, Zoetermeer, 3 dec. 1998.
- Ministerie van OC&W, *Keur aan kwaliteit – HOOP 2000*, TK 1999-2000, Den Haag, 2000.
- Ministerie van OC&W, *Kwaliteitszorg hoger onderwijs*, vergaderjaar 1999-2000, 27 010, nrs. 1-2, Sdu Uitgevers, Den Haag, 2000.
- Montfort, C. van, Verantwoording aan departementen, in: Bovens, M. en Th. Schillemans (red.), *Handboek publieke verantwoording*, Lemma, Den Haag, 2009, pp. 81-101.
- Moynihan, D.P., *The Dynamics of Performance Management*, Georgetown University Press, Washington D.C., 2008.
- Moynihan, D.P., S.K. Pandey & B.E. Wright, Setting the Tabler: How Transformational Leadership Fosters Performance Information Use, in: *Public Administration*, vol. 22, jan. 2012, nr. 1, pp. 143-165.
- Mulder, H.P., G. Walraven e.a., Het gebruik van beleidsevaluatie-onderzoek bij de rijksoverheid, in: *Beleidswetenschap*, 1991, nr. 3, pp. 203-277.
- Nationale Accreditatie Organisatie, *Uitgangspunten accreditatie en accreditatiekaders*, Den Haag, maart 2003.
- Nelissen, N.J.M. en P.J.M. de Goede, Benchmarking: vergelijken en verbeteren, in: *Bestuurskunde*, jrg. 8, 1999, nr. 2, pp. 54-70.
- Newman, D.L., R.D. Brown & L.A. Braskamp, Communication Theory and the Utilization of Evaluation, in: Braskamp, L. & R.D. Brown (eds.), *Utilization of Evaluative Information*, Jossey-Bass Publ., San Francisco, 1980, pp. 29-37.
- Noordegraaf, M. en J. Sterrenburg, Administratieve lasten voor publieke professionals: ficties en feiten, in: Jansen, Th., G. van den Brink en J. Kole (red.), *Beroepstrots- Een ongekende kracht*, Boom, Amsterdam, 2009, pp. 115-129.
- Noordegraaf, M. en J. Sterrenburg, Publieke professionals en verantwoordingsdruk, in: Bovens, M. en Th. Schillemans, *Handboek publieke verantwoording*, Lemma, Den Haag, 2009, pp. 231-255.
- Noordegraaf, M., Prestatiemanagement, in: Noordegraaf, M., K. Geuijen en A. Meijer (red.), *Handboek Publiek Management*, Boom/Lemma, Den Haag, 2011, pp. 365-397.
- Noordegraaf, M. en W.J.M. Kickert, Sturen op afstand in het hoger onderwijs, in: Kickert, W.J.M. (red.), *Veranderingen in management en organisatie bij de rijksoverheid*, Samsom, Alphen, 1993, pp. 77-90.
- Noordegraaf, M., *Management in het publieke domein*, Uitgeverij Coutinho, Bussum, 2004.
- Noordegraaf, M., *Professioneel bestuur*, Lemma, Utrecht, 2008.
- NVAO, *Accreditatiekaders*, website NVAO; www. NVAO.nl.
- Onderwijsraad, *Bureaucratisering in het onderwijs*, Den Haag, 2004.
- Onderwijsraad, *Hoger onderwijs voor de toekomst*, Den Haag, 2010.
- Oijen, P. van, I.Th.M. Snellen en J.M. van Westerlaak, *Ambtenaren en onderzoekers*, Leiden, 1984.
- Patton, M.Q., *Utilization-focused Evaluation*, Sage, Thousand Oaks, 1997.
- Pollitt, C., X. Girre e.a., *Performance or Compliance? – Performance Audit and Public Management in Five Countries*, Oxford University Press, Oxford, 1990.
- Power, M., *The Audit Explosion*, Demos, Oxford, 1994.
- Power, M., *The Audit Society - Rituals of verification*, Oxford University Press, Oxford, 1999.

Power, M., The Theory about the Audit Explosion, in: Ferlie, E., L.E. Lynn jr & C. Pollitt (eds.), *The Oxford Handbook of Public Management*, Oxford University Press, Oxford, 2005, pp. 326-346.

Prick, L., *Drammen, dreigen, draaien – Hoe het onderwijs twintig jaar lang vernieuwd werd*, Mets en Schilt, Amsterdam, 2006.

Pröpper, I.M.A.M., Het gebruik van beleidsevaluatie: terugkoppeling, in: Hoogerwerf, A. (red.), *Overheidsbeleid*, Samsom, Alphen a.d. Rijn, 1993, pp. 180-198.

Pröpper, I.M.A.M. en M.S. de Vries, Presentatie en gebruik van evaluatie-onderzoek, in: Bressers, J.T.A. en A. Hoogerwerf (red.), *Beleidsevaluatie*, Samsom, Alphen a.d. Rijn, 1995, pp. 158-177.

Quality Assurance Netherlands Universities (QANU), *Assessment of degree programmes, Public Administration*, Utrecht, September 2005.

Quality Assurance Netherlands Universities (QANU), *Onderwijsvisitatie Bedrijfswetenschappen*, Utrecht, 2004.

Quality Assurance Netherlands Universities (QANU), *QANU-KADER*, februari 2004, www.qanu.nl.

Quality Assurance Netherlands Universities (QANU), *Onderwijsvisitatie Politicologie*, Utrecht, 2004.

Raad voor het openbaar bestuur, *Presteren door leren – Benchmarken in het binnenlands bestuur*, Rob, Den Haag, 2002.

Rein, M. & D.A.Schön, Frame-reflective policy discourse, in: *Beleidsanalyse*, 1986, nr. 4, pp. 4-18.

Rein, M., Reframing Problematic Policies, in: Moran, M., M. Rein & R.E. Goodin (eds.), *The Oxford Handbook of Public Policy*, Oxford UP, Oxford, 2006, pp. 389-409.

Reitsma, A.B., *Kritische succesfactoren bij visitatie en kwaliteitsverbetering van P&A opleidingen*, Open Universiteit, Heerlen, 2003. (diss.).

Rossi, P.H. & H.E. Freeman, *Evaluation – A systematic approach*, Sage, Newbury Park, 1989.

Scheele, J.P., P.A.M. Maassen & D.F. Westerheijden (eds.), *To be Continued Follow-UP of Quality Assurance in Higher Education*, Elsevier/De Tijdstroom, Maarsssen, 1998.

Schillemans, Th. en M. Bovens, Publieke verantwoording 2.0: sober maar scherp, in: Bovens, M. en Th. Schillemans (red.), *Handboek publieke verantwoording*, Lemma, Den Haag, 2009, pp. 275-295.

Schillemans, Th., *Verantwoording in de schaduw van de macht - Horizontale verantwoording bij zelfstandige uitvoeringsorganisaties*, Lemma, Den Haag, 2007.

Schillemans, Th., Visitaties als vorm van verantwoording, in: Bovens, M. en Th. Schillemans (red.), *Handboek publieke verantwoording*, Lemma, Den Haag, 2009, pp. 161-181.

Schoenmaker, M., *Bestuurlijk gedonder*, Wolf Publ., Nijmegen, 2011.

Schön, D.A. & M. Rein, *Frame Reflection – Towards the Resolution of Intractable Policy Controversies*, Basic Books, New York, 1994.

Schutgens, J.M.L.R., A.J.C. Maessen en A.F.A. Korsten, *Beoordeling van gemeenten – Financiële verdiepingsonderzoeken en bestuurskrachtmetingen van gemeenten*, Eburon, Delft, 2009.

Schwarz, S. & D.F. Westerheijden (eds.), *Accreditation and Evaluation in the European Higher Education Area*, Kluwer Academic Publ., Dordrecht, 2004.

Scott, J., *Seeing like a State – How certain schemes to improve the human condition have failed*, Yale University Press, New Haven, 1996.

Shulba, L.M. & J.B. Cousins, Evaluation use: theory, research and practice since 1986, in: *Evaluation Practice*, vol. 18, 1997, nr. 3, pp. 195-208.

Smid, G., *Goed werk vraagt om goed besturen*, Open Universiteit, Heerlen, 2012 (oratie).

Slob, A., Sturen op kwaliteit van evaluatieonderzoek, in: Knaap, P. van der, en A. Schilder (red.), *Resultaatgericht sturen en evalueren*, Sdu, Den Haag, 2004, pp.139-145.

Smith, H. e.a., *Benchmarking and Threshold Standards in Higher Education*, Kogan Page, London, 2000.

Smith, P., On the unintended consequences of publishing performance data in the public sector, in: *International Journal of Public Administration*, 1995, nr. 18, pp. 377-e.v.

Steijn, B., *Carrièrejager of dienaar van de publieke zaak – Over ambtenaren en hun motivatie*, Erasmusuniversiteit, Rotterdam, 2006.

Steijn, B., Er leeft onvrede over managers in de publieke sector, in: Brink, G. van den, Th. Jansen en D. Pessers (red.), *Beroepszeer*, Boom, Amsterdam, 2005, pp. 132-143.

Swanborn, P.G., *Evalueren*, Boom, Amsterdam, 1999.

Swieringa, J. en A.F.M. Wierdsma, *Op weg naar een lerende organisatie*, Wolters Noordhoff, Groningen, 1990.

Swierstra, T. en E. Tonkens (red.), *De beste de baas? Prestatie, respect en solidariteit in een meritocratie*, Amsterdam University Press, Amsterdam, 2008.

Taylor, J., Factors Influencing the Use of Performance Information for Decision Making in Australian State Agencies, in: *Public Administration*, vol.89, 2011, nr. 4, pp. 1316-1335.

Thiel, S. van, en F.L. Leeuw, De prestatieparadox in de publieke sector, in: *Beleidswetenschap*, jrg. 17, 2003, nr. 2, pp. 123-144.

Tijdelijke adviescommissie Beoordeling Projectvoorstellen Studeerbaarheidsfonds (BPS), *3000x beter – Slotrapport Studeerbaarheidsfonds*, Ministerie van OC en W, Zoetermeer, 1999.

Tijssen, I., *Kwaliteit noodt tot meer gebruik- Bruikbaarheid van sociaal-wetenschappelijke onderzoeksresultaten volgens arts-bestuurders*, Quickprint, Nijmegen, 1988.

Tillema, S., Gebruik van benchmarkinformatie in de publieke sector, in: *MAB*, vol. 80, 2006, nr. 6, pp. 300-308.

Tillema, S. en G.J. van Helden, *Van vergelijking naar verbetering? Onderzoek naar de actiebereidheid van waterschappen naar aanleiding van de bedrijfsvergelijking Zuiveringsbeheer 1999*, Groningen, 2003.

Tonkens, E., *Mondige burgers, getemde professionals*, Van Gennip, Amsterdam, 2008.

Tops, P., *Veiligheid: iedereen een zorg? Visitatiecommissie veiligheidsbeleid in Tilburg*, Gemeente Tilburg, Tilburg, 2006.

Twist, M.J.W. van, R. Bagchus en F.O.M. Verhaak (red.), *Kerndepartementen op afstand?*, Eburon, Delft, 1996.

Universiteit Twente, Faculteit Bestuurskunde, *Zelfevaluatie Opleiding Bestuurskunde*, mei 1997.

Vall, M. van de, *Sociaal beleidsonderzoek*, Samsom, Alphen, 1980.

Vall, M. van de, en B. van Dijkum-De Jong, *Het rendement van sociaal beleidsonderzoek – Een onderzoek naar de voorwaarden voor beleidsrelevantie*, Samsom, Alphen a.d. Rijn, 1979.

Veld, R.J. in 't, De sturingsconceptie van paars, in: Becker, F., W. van Henekeler, M. Sie Dhian Ho en B. Tromp (red.), *Zeven jaar paars*, De Arbeiderspers, Amsterdam, 2001a, pp. 164-197.

Veld, R.J. in 't, Onderwijs: bevrijding uit de beleidsgevangenis, in: Bussemaker, J. en R. van der Ploeg (red.), *Leven na Paars?*, Prometheus, Amsterdam, 2001b, pp. 85-97.

Veld, R.J. in 't, *Kennisdemocratie- Opkomend stormtij*, Academic Service/ Sdu, Den Haag, 2010.

Veld, R.J. in 't, *Een vlucht naar Isfahan?*, Vuga, Den Haag, 1984.

Veld, R.J. in 't, en P. van der Knaap (red.), *Dynamische bestuurskunde*, Phaedrus, Den Haag, 1994.

Verhoeven, J. en K. de Wit, Hoger onderwijs en beleid in Europa, in: *Bestuurskunde*, jrg. 8, 1999, nr. 8, pp. 346-357.

Visitatiecommissie Rechtspraak, *Rapport visitatie gerechten*, Raad voor de Rechtspraak, Den Haag, 2006 en 2010.

Visitatiecommissie Vlaamse steden (Commissie-Korsten-Reynaert/IDEA), *Syntheserapport Visitaties stedenfonds 2011 - Over het Stedenfonds en het samenspel tussen de Vlaamse overheid en haar centrumsteden*, IDEA Consult, nov. 2011 (op www.thuisindestad.be).

Vries, J. de, 'Bildung is iets voor in het weekend' - Interview met Grahame Lock, in: *De Groene Amsterdammer*, 22 augustus 2008.

Vries, J. de, *Paars en de managementstaat – Het eerste Kabinet-Kok (1994-1998)*, Garant, Amersfoort, 2002.

Vries, M.S. de, *De ontmythologisering van het openbaar bestuur*, Radboud Universiteit, Nijmegen, 2005.

VSNU, *Assessment of Research Quality - Protocol 1998*, VSNU, Utrecht, 1998.

VSNU, *Assessment of Research Quality- Public Administration – Political Science – Communication Science*, Utrecht, 2002.

VSNU, *Assessment of Research Quality- Rechtsgeleerdheid*, VSNU, Utrecht, 2002.

VSNU, *Onderwijsvisitatie Economie*, Utrecht, juli 2003.

VSNU, *Quality Assessment of Research - Onderzoekbeoordeling – Protocol 1994*, VSNU, Utrecht, 1994.

VSNU, *Quality Assessment of Research – Rechtsgeleerdheid*, VSNU, Utrecht, 1996.

VSNU, *Quality Measurement Made tot Measure: Protocol for External Assessment of Educational Programs 2000-2005*, Utrecht, 1999.

VSNU-visitatiecommissie Bestuurs-, Beleids- en Politieke Wetenschappen, *Onderwijsvisitatie – Bestuurs-, beleids- en politieke wetenschappen*, VSNU, Utrecht, 1992.

VSNU-visitatiecommissie Bestuurskunde, *Onderwijsvisitatie – De kwaliteit van het universitair onderwijs – Bestuurskunde*, VSNU, Utrecht, 1998.

Vught, F.A. van, en D.F. Westerheijden, Toward a general theory of quality assessment in higher education, in: *Higher Education*, jrg. 28, 1994, pp. 355-371.

- Walle, S. van de, & A. Roberts, Publishing Performance Information: An Illusion of control?, in: Dooren, W. van, en S. van de Walle (eds.), *Performance Information in the Public Sector- How it is used*, Palgrave MacMillan, Londen, 2008, pp. 211-227.
- Walle, S. van de, *Perceptions of Administrative Performance: The Key to Trust in Government?*, KULeuven, Leuven, 2004.
- Walle, S. van de, Zullen beter werkende overheidsdiensten bijdragen tot vertrouwensherstel?, in: Korsten, A. en P. de Goede (red.), *Bouwen aan vertrouwen in het openbaar bestuur – Diagnoses en remedies*, Elsevier, Den Haag, 2006, pp. 125-139.
- Webber, D.J., The Distribution and Use of Policy Knowledge in the Policy Process, in: *Knowledge and Policy*, vol. 4, 1991, nr. 4, pp. 6-35.
- Weibel, A., K. Rost & M. Osterloh, Pay for performance in the public sector- Benefits and (hidden) costs, in: *Journal of Public Administration Research and Theory*, vol. 20, 2009, nr. 2, pp. 387-412.
- Weiss, C.H. (ed.), *Using Social Science Research in Public Policy-Making*, Lexington Books, Lexington, 1977.
- Weiss, C.H. (with M. Bucuvalas), *Social Science Research and Decision-making*, Columbia University Press, New York, 1980.
- Weiss, C.H., E. Murphy-Graham & S. Birkeland, An Alternative Route to Policy Influence, in: *American Journal of Evaluation*, vol. 26, maart 2005, nr. 1, pp. 12-30.
- Weiss, C.H., Evaluation Research in a Political Context, in: Struening, E. & M. Guttentag (eds.), *Handbook of Evaluation Research*, Sage, Beverly Hills, 1976, pp. 13-27.
- Weiss, C.H., Knowledge Creep and Decision Accretion, in: *Knowledge: Creation, Diffusion, Utilization*, 1980, pp. 381-404.
- Weiss, C.H., Research for Policy's Sake: the Enlightenment Function of Social Science, in: *Policy Analysis*, vol. 3, 1977, nr. 4, pp. 531-545.
- Weiss, C.H., The Many Meanings of Research Utilization, in: *Public Administration Review*, sept/oct. 1979, pp. 426-431.
- Westerheijden, D.F. e.a., *Changing contexts of quality assessment, Recent trends in West European Higher Education*, CHEPS/Lemma, Utrecht, 1994.
- Westerheijden, D.F., A solid base for decisions: Use of the VSNU research evaluations in Dutch universities, in: *Higher Education*, vol. 33, 1997, nr. 4, pp. 397-413.
- Westerheijden, D.F., B. Stensaker en M.J. Rosa (red.), *Quality Assurance in Higher Education, Trends in Regulation, Translation and Transformation*, Springer, Dordrecht, 2007.
- Westerheijden, D.F., Quality assessment in Dutch higher education: Balancing improvement and accountability, in: *European Journal for Education Law and Policy*, vol. 1, 1997, pp. 81-90.
- Weusthof, P.J.M., *De interne kwaliteitszorg in het wetenschappelijk onderwijs*, Lemma, Utrecht, 1994.
- Wildavsky, A., The Self-Evaluating Organization, in: Kearny, R. & E.V. Berman (ed.), *Public Sector Performance – Management, Motivation and Measurement*, ASPA, Westview Press, Boulder, 1999, pp. 273-293.
- Winter, H.B. en M. Herweijer, Evaluatie van beleid, in: *Strategie en beleid in de publieke sector*, Samsom, Alphen, dec. 1995, F2120, pp. 1-26.
- Yin, R.K. & M.K. Gwaltney, Knowledge Utilization as a Networking Process, in: *Knowledge: Creation, Diffusion, Utilization*, vol. 2, 1981, nr. 4, pp. 555-580.

Over de auteurs

Prof.dr. A.F.A. Korsten is honorair hoogleraar Bestuurskunde van de lagere overheden bij de Capaciteitsgroep Methoden en Grondslagen van het Recht aan de Universiteit Maastricht (fac. rechtsgeleerdheid) en emeritus hoogleraar Bestuurskunde aan de Open Universiteit (fac. managementwetenschappen). Publiceerde tal van boeken en artikelen. Was in 2003 voorzitter van een visitatiecommissie ter visitatie van de hbo-opleidingen op het gebied van bestuurskunde/overheidsmanagement en in 2007 voorzitter van de 'midterm review'-commissie die een tussentijdse evaluatie uitvoerde van de wo- opleiding bestuurskunde van de Erasmus Universiteit te Rotterdam vanuit o.m. de vraag van de faculteitsdecaan: is voldoende verbetering zichtbaar na de vorige visitatie? Was in 2010-2011 voorzitter van negen visitatiecommissies ter visitatie van Vlaamse grote steden.

Drs. J. Schutgens is van opleiding bestuurskundige. Was jarenlang in verschillende functies werkzaam in het lokaal bestuur, waaronder als gemeentesecretaris. Is promovendus, met A. Korsten als (beoogd) promotor. Onderwerp: visiteren van gemeentebesturen. Publiceerde met met K. Abma en A.Korsten *Bestuurskracht van gemeenten* (Eburon, 2007) en met A. Korsten en A. Maessen het boek *Beoordeling van gemeenten* (Eburon, Delft 2009). Redigeerde met A. Eversdijk *Openbaar bestuur op expeditie* (Eburon, 2010). Publiceerde met anderen in 2007 een artikel in het tijdschrift Bestuurskunde.