

Analyse van besluitvormingsprocessen bij de overheid

Een toetsingsprocedure voor expertkennis in overheidsbeslissingen

A. Bruinsma en G.P.A. Braam*

Summary

*Analysis of decision making in public administration
A method to evaluate expert-knowledge*

It is to be expected that, in the near future, public administration will enlist computerized systems to support decision making. These so-called expertsystems contain human knowledge about the problem domain.

As for public administration, this means that the stored knowledge should be justifiable in the light of the policy followed; in other words, not be biased by prejudices and personal judgements. We think that, particularly in many domains of public policy, the quality of expert-knowledge is dubious.

Therefore, in this article we intend to outline a method which makes it possible, first, to establish to what extent the decisions of public administrators reflect expert-knowledge, and second, to uncover improper criteria possibly underlying these decisions.

The quintessence of the method lies in the tracing of decision criteria from empirical data. To accomplish this, we need a model of the decision process. Next we transform the criteria into productions (if . . . then rules), and put them to a statistical test, borrowed from the field of meteorology. We applied this method on a set of decisions made by a local administrative agency about subsidy requests made by primary schools.

The most important finding was, that the proposed procedure seemed to be viable. Among other things, we came across the finding that the improper criterium 'prestige' influences the decisions made by a local administrative agency.

* Onderafdeling der Bestuurskunde, Technische Hogeschool Twente. Met dank aan L. Elting en M. Even voor het ter beschikking stellen van het materiaal, en aan P.A.Th.M. Geurts voor de eerste bewerking van dit materiaal.

1. Inleiding

In dit artikel wordt een methode geschetst waarmee het mogelijk is ten eerste de mate waarin beslissingen expertkennis weerspiegelen vast te stellen, en ten tweede eventuele oneigenlijke criteria in deze beslissingen bloot te leggen. De methode is te gebruiken voor onderzoek op beslissingsdomeinen waarin de overheid beslist over wensen van burgers. We zullen haar hier toepassen op beslissingen van een lokale overheid over geldvragen van openbare lagere scholen ten behoeve van leermiddelen.

2. Vaardigheid en kennis bij beslissingen: de kwaliteit van expertkennis

Het hoe en wat van de menselijke besluitvorming komt in toenemende mate in de belangstelling te staan. Dit wordt mede veroorzaakt door ontwikkelingen op het gebied van de informatica en de kunstmatige intelligentie. Deze ontwikkelingen hebben het mogelijk gemaakt expertsystemen te ontwerpen. Dat wil zeggen, computersystemen die de mens professioneel ondersteunen in de besluitvorming (cf. Shortliffe, 1977).

Ook in de sfeer van de overheid liggen dergelijke ontwikkelingen in het verschiep; ontwikkelingen die vanuit de sociale wetenschappen nauwkeurig zullen moeten worden gevolgd. We lichten dit nader toe.

In een expertsysteem wordt menselijke kennis opgeslagen. Dit kan 'objectieve' kennis zijn zoals wetenschappelijk vastgestelde verbanden of door wet en jurisprudentie voorgeschreven regels, maar ook op ervaringen gebaseerde 'subjectieve' kennis (cf. Simon, 1981). Deze ervaringskennis is heuristisch, vaak onzeker (Feigenbaum, 1979) en moeilijk toegankelijk (Breuker en Wielinga, 1983). Op grond hiervan valt te betwijfelen of de ervaringskennis die door ambtenaren wordt aangedragen wel altijd expertkennis is.

We veronderstellen dat de door ambtenaren aangedragen kennis gekleurd zal zijn door eigen visies, vooroordelen, alsmede groepsinvloeden (cf. Inbar, 1979; Simon, 1976; Steinbruner, 1974). Voor zover dit het geval is, worden officiële beleidsuitgangspunten gebrekkig in het geding gebracht. Dat hiervan sprake is, blijkt uit sociaal-wetenschappelijk onderzoek. Zo zijn er in de verzorgingsfeer aanwijzingen dat beslissingen over toewijzing van gezinszorg (Coolen, 1982) en toelating tot verzorgingstehuizen (Coolen en Van Pelt, 1985) eenzijdigheden en onregelmatigheden vertonen. Ook invloeds-onderzoek op terreinen als waterstaatkundige voorzieningen (Braam, 1973) en sport (Bos en Braam, 1980) wijst in deze richting. Het recente voorbeeld van het RSV-debacle heeft iets van de gebrekkige besluitvorming – een ge-

brekigheid die zeer algemeen is — in de openbaarheid gebracht.

Daarmee is de bedoeling van dit artikel gegeven: we willen de vraag *in hoeverre (wettelijk) voorgeschreven uitgangspunten een rol spelen in de beslissingen, en omgekeerd in hoeverre andere uitgangspunten ten onrechte een rol spelen*, methodisch beantwoorden.

We beperken ons hierbij tot een bepaald type van beslissingen, te weten *beslissingen van de overheid over vragen van burgers*.

2.1. Een methode voor het evalueren van expertkennis (in hoofdlijnen)

In de evaluatieprocedure die we zullen schetsen zijn een tweetal stappen te onderscheiden:

1. Het vaststellen van gebruikte uitgangspunten

Dit geschiedt op basis van hypothesen naar aanleiding van officieel bekende beleidsuitgangspunten, en verder op basis van in empirisch onderzoek gevonden 'onjuiste uitgangspunten'.

Of deze uitgangspunten wel of niet in de feitelijke beslissingen tot gelding komen, gaan we na aan de hand van kenmerken van de 'vragende' burger (omgevingskenmerken).

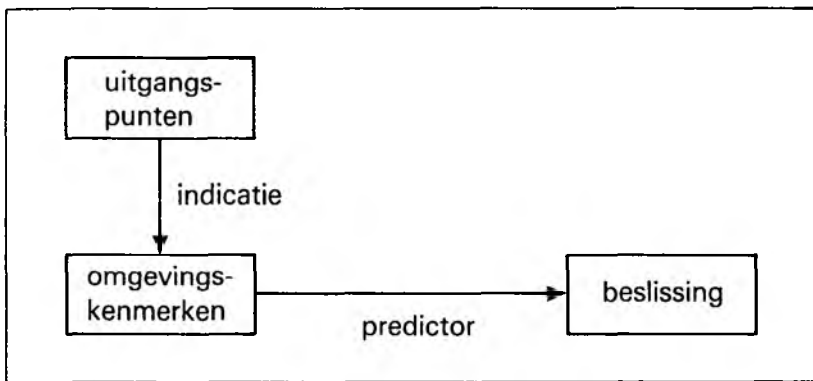
De nu volgende stap is relevant met betrekking tot de ontwikkeling van een expertstelsel.

2. De evaluatie in engere zin

Op basis van de gevonden uitgangspunten formuleren we beslissingsregels. Deze hebben de vorm van een *productie: als* conditie *dan* actie. Door de beslissingsregels te formuleren in termen van producties sluiten we zo nauw mogelijk aan bij de wijze waarop kennis veelal wordt gerepresenteerd in expertsystemen.

In de beslissingsregels treden de omgevingskenmerken op in het conditiegedeelte, en de feitelijke beslissingen in het actiegedeelte. We gaan na in hoeverre deze kenmerken beschouwd kunnen worden als *predictoren* voor de actie (overheidsbeslissing).

We mogen van een dergelijke evaluatieprocedure verwachten dat zij aanwijzingen verschaft voor mogelijke correcties op de expertkennis. Met name kan het de beslisser duidelijk worden gemaakt welke uitgangspunten sterker in het geding moeten worden gebracht, en welke juist niet.



Figuur 1. Samenhang tussen omgevingskenmerken, (beleids)uitgangspunten en overheidsbeslissingen

3. Het vaststellen van gebruikte uitgangspunten in de besluitvorming (op basis van omgevingskenmerken en beslissingsuitkomsten)

Zoals we hebben aangekondigd bestaat de eerste stap van de hier voorgestelde procedure uit het vaststellen van uitgangspunten die in de beslissingen feitelijk zijn gebruikt. Dit betekent dat we moeten bezien hoe de niet direct meetbare uitgangspunten kunnen worden gerelateerd aan wel waarneembare grootheden (namelijk omgevingskenmerken en uitkomsten van beslissingen). Hiervoor hebben we een model van een beslissing nodig.

We zullen aangeven dat uit dit model bepaalde eisen volgen voor de wijze waarop we het beslissingsproces moeten waarnemen. Daarna bespreken we hoe op basis van enkele hypothesen omtrent onze gegevens uitgangspunten kunnen worden vastgesteld.

3.1. Het besluitvormingsproces en omgevingskenmerken.

Wat verstaan we onder een beslissing? Voor het antwoord op deze vraag grijpen we terug naar een schema van Simon (1976). In dit schema wordt een beslissing voorgesteld als een keuze tussen verschillende gedragsalternatieven of handelwijzen. (Ter vergelijking, in ons onderzoek heeft de overheid de keuze tussen het al dan niet verstrekken van het geld.) We zullen nu beschrijven hoe de keuze voor een bepaalde handelwijze tot stand komt.

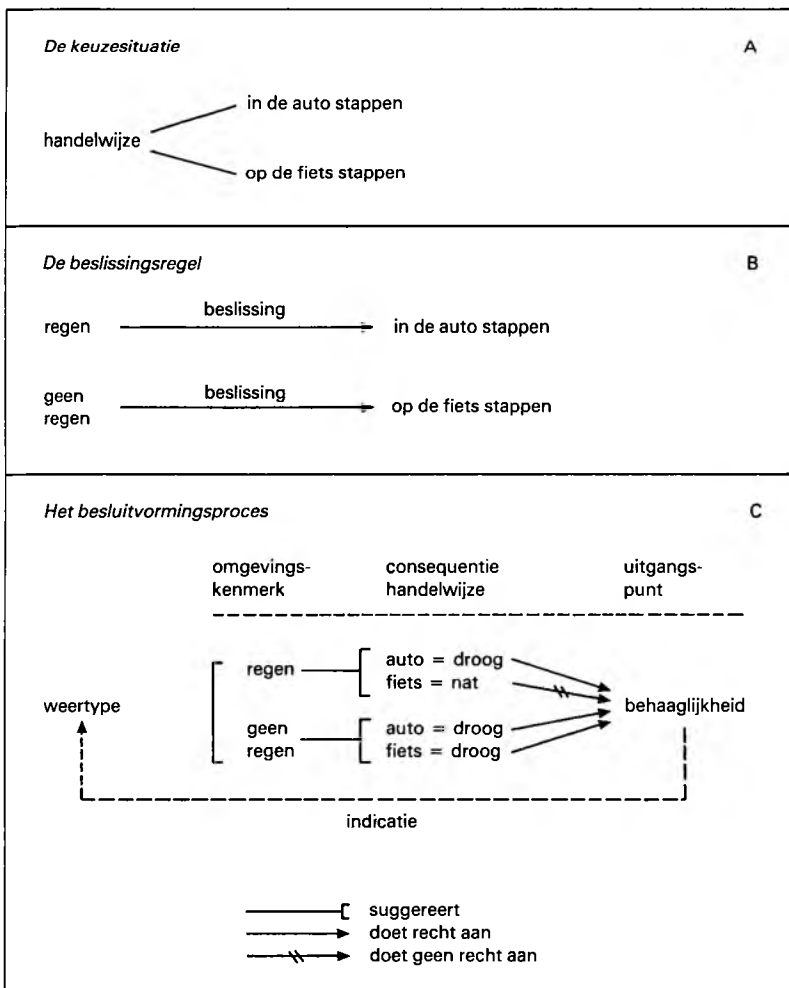
Zoals aangegeven staat een beslisser voor de opgave om uit een aantal handelwijzen één te kiezen. Wat doet de beslisser? Zij of hij begint ermee de con-

sequenties van de verschillende handelwijzen te schatten. Deze consequenties krijgen hun waarde doordat de beslisser bepaalde uitgangspunten bij de beslissing betreft. Zo worden handelwijzen waarvan de consequenties stroken met de uitgangspunten aantrekkelijk gevonden, in tegenstelling tot de handelwijzen waarvan de consequenties botsen met de uitgangspunten. We mogen dus verwachten dat de uiteindelijke keuze zal vallen op de handelwijze die het meest aantrekkelijk wordt gevonden. Deze vrij abstracte beschrijving van het besluitvormingsproces zullen we hierna verduidelijken aan de hand van een eenvoudig voorbeeld.

Dagelijks worden we voortdurend geconfronteerd met situaties waarin een beslissing van ons wordt gevraagd. Bijvoorbeeld, als we iemand willen bezoeken moet een beslissing worden genomen over de wijze waarop we erheen zullen gaan, eenvoudig vanwege het feit dat we ons op verscheidene manieren kunnen verplaatsen. Stel, we hebben de keuze tussen een tweetal vervoersmogelijkheden: de fiets of de auto, en we nemen de beslissing om met de auto te gaan. Op welke manier is deze beslissing tot stand gekomen? Teneinde een beslissing te kunnen nemen over hoe ons te verplaatsen schatten we de consequenties van beide handelwijzen. We doen dit onder andere door te letten op wat er in de omgeving gebeurt. Veronderstel dat er een tweetal omgevingskenmerken mogelijk zijn: het regent of het regent niet. Als we naar buiten kijken en zien dat het regent, dan is een voor de hand liggende conclusie dat we een nat pak krijgen (consequentie) indien we op de fiets stappen (handelwijze), en dat we in de auto droog blijven. Deze twee consequenties zullen we verschillend waarderen, al naar gelang onze uitgangspunten. Streven we bijvoorbeeld een gevoel van behaaglijkheid na dan stappen we bij regenweer in de auto, en niet op de fiets. Immers, een nat pak botst met dit uitgangspunt, in tegenstelling tot droge kleren die ermee stroken. Derhalve, omdat we een gevoel van behaaglijkheid nastreven, verkiezen we als het regent de auto boven de fiets.

In het voorgaande hebben we de gedachtengang weergegeven die schuilt achter de in dit artikel geschetste methode voor het evalueren van expertkennis. Deze gedachtengang is weergegeven in figuur 2, en laat zich beschrijven in een viertal punten die we hierna kort aanduiden.

1. De beslisser die zich in een bepaalde probleemsituatie ziet gesteld, moet een keuze doen uit verschillende handelwijzen (figuur 2A).
2. Het resultaat van het besluitvormingsproces waarlangs de voorkeur voor een bepaalde handelwijze tot stand komt, is samen te vatten in termen van een beslissingsregel of produktie. Bijvoorbeeld:



Figuur 2. Het besluitvormingsproces en omgevingskenmerken

regel weertype als het regent,
dan neem de auto,
anders neem de fiets.

3. De geformuleerde beslissingsregel geeft de samenhang weer die wordt verondersteld tussen omgevingskenmerken en beslissingsuitkomsten. Al-

dus pretendeert zij een voorspelling te zijn van het keuzegedrag van de beslisser (figuur 2B).

4. De achtergrond van de beslissingsregel is een uitgangspunt. Zo vloeit de onder punt 2 gegeven beslissingsregel voort uit het uitgangspunt behaaglijkheid (figuur 2C). (Merk op dat een ander uitgangspunt de aandacht van de beslisser zou hebben gericht op andere omgevingskenmerken, en aldus zou hebben geleid tot een andere beslissingsregel.)

We sluiten deze subparagraaf af met de conclusie dat de niet direct meetbare uitgangspunten zijn af te leiden uit een eventuele samenhang tussen wel direct meetbare omgevingskenmerken en beslissingsuitkomsten. Derhalve, de vraag of de overheid wel of niet terecht gebruik maakt van uitgangspunten in haar beslissingen laat zich beantwoorden indien we *a.* beschikken over een reeks van beslissingsuitkomsten, *b.* weten welke omgevingskenmerken met deze uitkomsten zijn verbonden, en *c.* een procedure hebben om een mogelijke samenhang tussen deze twee op te sporen. In het navolgende komen we hier uitvoerig op terug.

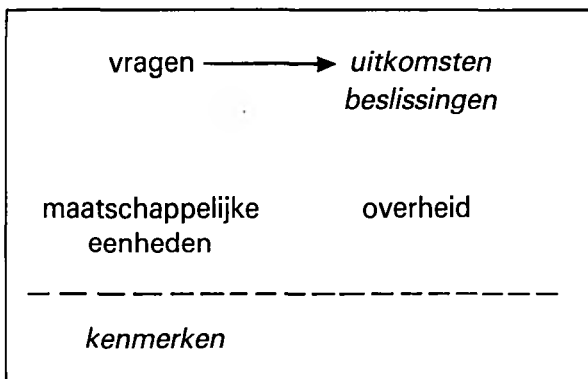
3.2. De waarnemingen: de noodzaak van onafhankelijke gegevens

We zullen thans aandacht besteden aan de wijze waarop de zojuist besproken gegevens (omgevingskenmerken en beslissingsuitkomsten) dienen te worden waargenomen.

Het is gebruikelijk om kennis voor een expertsysteem te vergaren door beslissers, die als expert worden aangemerkt, te vragen hoe ze problemen binnen hun domein tot een oplossing brengen. Voor ons doel is deze raadpleging evenwel niet voldoende. Willen we toetsen in hoeverre de overheid beleidsuitgangspunten betreft bij haar beslissingen, dan dienen we ons tevens te baseren op gegevens die niet door de beslissers zelf zijn vertekend, dat wil zeggen op gegevens die onafhankelijk van de mogelijke vooroordelen van de beslissers zijn vastgesteld.

De beslissingsdomeinen waarop ons onderzoek betrekking heeft, dus vragen van burgers aan de overheid, lenen zich hiertoe. We kunnen ons namelijk wenden tot deze maatschappelijke eenheden, hun kenmerken vaststellen en kijken hoe de overheid over hun vragen heeft beslist. Dit betekent derhalve dat we moeten waarnemen welke processen zich feitelijk hebben afgespeeld (zie figuur 3).

De waarnemingsmethode die we hiertoe zullen gebruiken is ontwikkeld door Braam in zijn studie naar de invloed van bedrijven op de overheid (Braam,



Figuur 3. Vragen van maatschappelijke eenheden en overheidsbeslissingen

1973). Deze methode, de procesmethode, is sedertdien met vrucht toegepast bij ander onderzoek naar de verdeling van invloed (cf. Alferink, 1980; Bos en Braam, 1980; Elting en Even, 1977).

3.3. Het onderzoek en veronderstellingen omtrent uitgangspunten

Ons onderzoek heeft betrekking op 91 geldvragen van 28 openbare lagere scholen aan de lokale overheid van één en dezelfde stad. De scholen wilden dit geld gebruiken om leermiddelen, zoals leesboeken, te bekostigen. In dit verband is het belangrijk te constateren dat de scholen niet zonder meer recht hadden op het gevraagde geld. Dus, als de overheid een krediet verstrekt dan doet zij dit voornamelijk naar aanleiding van de geldvraag van een school. Zoals we hebben gezien, gebeurt deze inwilliging of weigering van een verzoek mede op grond van bepaalde kenmerken van de vrager. Indien we deze kenmerken verbinden met de uitkomst van de overheidsbeslissingen, kunnen we te weten komen in hoeverre de overheid expertkennis in huis heeft (d.w.z. beleidsuitgangspunten bij haar beslissingen betreft, danwel oneigenlijke criteria).

We hebben ons beperkt tot een drietal uitgangspunten waarvan we weten, of redelijkerwijs denken te weten, dat ze van belang zijn geweest voor de uitkomst van de overheidsbeslissingen.

In concreto, we weten dat het bevorderen van gelijke onderwijskansen voor kinderen uit onderscheiden sociale milieus een beleidsuitgangspunt was in de periode waarop het onderzoek betrekking heeft (1972-1975). Daarnaast veronderstellen we dat de beleidsafdelingen van de gemeente zich bij

hun beslissingen tevens hebben laten leiden door een tweetal andere uitgangspunten. Ten eerste: de gemeente zal willen voorkomen dat kinderen in een uitzonderingspositie worden geplaatst. Ten tweede: het is op grond van invloedsonderzoek aannemelijk dat er een zekere gevoeligheid is voor het prestige van de om geld vragende school (cf. Bos en Braam, 1980; Braam, 1973). Uiteraard kan men stellen dat het laatste uitgangspunt, indien het werkt, ten onrechte wordt gehanteerd.

Korthedshalve zullen we in de rest van ons betoog de drie uitgangspunten aanduiden met de termen:

- gelijkheid,
- voorkomen van uitzonderingspositie,
- prestige.

Overeenkomstig het beschreven model, veronderstellen we dat de volgende omgevingskenmerken een indicatie zijn voor deze uitgangspunten, respectievelijk:

- percentage ouders met een laag beroep (LB),
- aantal verwijzingen naar het BLO-onderwijs (BLO),
- groei in aantal leerlingen (SG).

Bijvoorbeeld, indien de overheid positief beslist over geldvragen van scholen die gekenmerkt worden door een hoog percentage ouders met een laag beroep, en negatief over geldvragen van scholen met een laag percentage, dan kunnen we dit verband tussen 'percentage ouders met een laag beroep' en de uitkomst van de beslissingen beschouwen als een indicatie voor het uitgangspunt 'gelijkheid'. Met andere woorden het verband geeft aanleiding tot de formulering van een beslissingsregel:

*regel % laag beroep als hoog percentage laag beroep,
dan positieve beslissing,
anders negatieve beslissing.*

Deze redenering vertoont veel overeenkomsten met het eerder beschreven voorbeeld waarin een beslisser de keuze had tussen twee vervoersmogelijkheden: de auto of de fiets. Een en ander is weergegeven in figuur 4.

3.4. Resultaten

Om vast te stellen of de genoemde uitgangspunten een rol hebben gespeeld bij de overheidsbeslissingen bezien we de verbanden tussen omgevingskenmerken en beslissingsuitkomsten op de manier die in de sociale wetenschappen gangbaar is. We maken hiertoe slechts gebruik van de helft van het mate-

	Omgevingskenmerk		Uitgangspunt
voorbeeld	weertype	[regen geen regen	behaaglijkheid
onderzoek	percentage laag beroep	[hoog laag	gelijkheid
	verwijzingen BLO	[veel weinig	voorkomen van uitzonderingspositie
	schoolgroei	[toename geen toename	prestige

Figuur 4. Overeenkomsten tussen voorbeeld en onderzoek

riaal dat tot onze beschikking staat. Hetgeen wil zeggen dat we de eerste stap van de evaluatieprocedure baseren op 46 waarnemingen. (Op grond van de resultaten formuleren we beslissingsregels die in meer strikte zin worden getoetst op de overige 45 waarnemingen in de tweede stap van de evaluatieprocedure, zie paragraaf 4.) De resultaten van de eerste stap zijn weergegeven in tabel 1. We zien dat de variabelen LB, BLO en SG alle enigszins verband houden met de uitkomst van de overheidsbeslissingen. We concluderen dat de drie uitgangspunten tot op zekere hoogte werkzaam zijn geweest.

Tabel 1. De verbanden tussen omgevingskenmerken en overheidsbeslissingen. LB = percentage ouders met een laag beroep; BLO = aantal leerlingen dat verwezen wordt naar het BLO; SG = groei in aantal leerlingen. $p_p | \text{kenmerk}$ = kans op een positieve beslissing, gegeven een bepaald kenmerk; p = de gemiddelde kans op een positieve beslissing.

Beslissing	Kenmerk						Σ
	LB		BLO		SG		
	hoog	laag	veel	weinig	toename	geen toename	
positief	21	13	18	16	20	14	34
negatief	5	7	4	8	6	6	12
Σ	26	20	22	24	26	20	46
$p_p \text{kenmerk}$.81	.65	.82	.67	.77		
p							.74

In het navolgende pogen we aan dit 'tot op zekere hoogte' een nadere invulling te geven.

4. De evaluatie in engere zin

Het vaststellen van feitelijk gebruikte uitgangspunten zoals boven is gebeurd, verloopt volgens een werkwijze die gangbaar is in de sociale wetenschappen. De met deze werkwijze verkregen uitkomsten zijn voor ons doel echter niet voldoende. Want, we mogen weliswaar zo (beleids)uitgangspunten vinden, de resultaten zijn te ruw om tot praktische aanbevelingen te komen. Aanbevelingen die ertoe leiden dat het gebruik van de uitgangspunten wordt gewijzigd. Om dit mogelijk te maken zullen we een stap verder moeten doen. Hiertoe herformuleren we onze vraag als volgt: *indien we een gevonden verband tussen een omgevingskenmerk en een bepaalde overheidsbeslissing omzetten in een beslissingsregel, wat zal dan de prestatie van deze regel zijn?* Onder een goede prestatie verstaan we dat de in de beslissingsregel genoemde omgevingskenmerken de uitkomst van de overheidsbeslissingen adequaat voorspellen.

Om de predictieve waarde van een beslissingsregel te kunnen evalueren hebben we een maatstaf nodig.

4.1. De prestatie-index *I* als maatstaf voor het bepalen van de prestatie van een beslissingsregel

Uit onderzoek is gebleken dat een aantal gangbare maatstaven zoals de χ^2 -toets, of Yule's *Q* en *Y*, ondergeschikt zijn voor het evalueren van predicties. Voor ons doel zijn deze maatstaven dus in elk geval niet bruikbaar. Andere maatstaven, zoals de correlatie-coëfficiënt, of de ratiotest, blijken zeer gevoelig voor de verhouding tussen positieve en negatieve beslissingen in een reeks gegevens. Dit houdt in dat, indien de verhouding tussen positieve en negatieve beslissingen verschilt voor twee reeksen van gegevens, een vergelijking tussen de sterkte van het verband waarop een beslissingsregel is gebaseerd en de predictieve waarde van deze regel pas na veel inspanning kan worden gemaakt. Dit is te vermijden door te kiezen voor een maatstaf die ongevoelig is voor de genoemde verhouding. Voor een uitgebreide bespreking van deze problematiek verwijzen we naar Woodcock (1975).

Om de predictieve waarde van de beslissingsregel te bepalen, maken we gebruik van een door Kuipers geïntroduceerde maatstaf (Kuipers, 1954). Deze zogenaamde prestatie-index werd en wordt gebruikt om predictoren voor

Geobserveerd	Gegenereerd		Σ
	positief	negatief	
positief	a_1	a_2	n_1
negatief	a_3	a_4	n_1
			n

Figuur 5. Een kruistabel waarop de prestatie-index wordt berekend

weersvoorspellingen te selecteren en te evalueren op hun waarde (cf. Hansen, 1965). (Ook op het wetenschapsdomein van de meteorologie gaat het om beslissingsregels. Deze betreffen hier weersvoorspellingen. Bijvoorbeeld, op grond van het kenmerk luchtdruk beslist men om wel of geen regen te voorspellen.)

De prestatie-index wordt berekend op tabellen zoals gegeven in figuur 5. Dergelijke tabellen verkrijgt men door de beslissingen die een beslissingsregel genereert te vergelijken met de beslissingen zoals die feitelijk zijn genomen. Het aantal malen dat de beslissingen van beide gelijk zijn, is geconcentreerd in de cellen a_1 en a_4 . We noemen dit de *treffers*, en in navolging hiervan de cellen a_2 en a_3 de *missers*. De prestatie-index nu, wordt uitgedrukt in termen van de treffers en de missers. Zij is de som van de volgende twee proporties:

$$I = \frac{a_1}{a_1 + a_2} + \frac{a_4}{a_3 + a_4} - 1 = \frac{a_1}{n p_p} + \frac{a_4}{n(1 - p_p)} - 1 \quad (1)$$

waar:

$n_1 = a_1 + a_2$ het aantal positieve overheidsbeslissingen is;

$n_2 = a_3 + a_4$ het aantal negatieve overheidsbeslissingen is;

n het totale aantal overheidsbeslissingen is;

$p_p = n_1/n$ de gemiddelde kans op een positieve beslissing;

$p_n = (1 - p_p)/n$ de gemiddelde kans op een negatieve beslissing.

De term minus 1 bewerkstelligt dat $-1 \leq I \leq 1$. Met andere woorden, I varieert tussen een volkomen onjuiste en een perfecte predictie (of reproductie). De prestatie-index is ongevoelig voor de verschillen in de marginalen n_1 en n_2 . Zoals we hebben betoogd is dit belangrijk omdat we dan de prestatie van een beslissingsregel in verschillende reeksen *zonder meer* kunnen vergelijken. (In de twee reeksen waarop we ons baseren is de verhouding 34/12 voor de reproductie en 35/10 voor de predictie.)

Het blijkt dat de waarde van I niet alleen afhangt van de predictor die wordt geselecteerd, maar ook van de wijze waarop deze wordt gehanteerd. Het gaat daarbij in het bijzonder om de keuze van de waarde van het 'cutting-point'. Dus om de vraag boven welke waarde van de variabele we een positieve beslissing moeten voorspellen. Het kan worden aangetoond dat I een relatief maximum heeft indien we een positieve beslissing voorspellen als de kans op een positieve beslissing, *gegeven* een geselecteerde predictor, zeg A , groter is dan de gemiddelde kans op een positieve beslissing, dat wil zeggen: als $p_{p|A} > p_p$. Of dit het geval is kunnen we afleiden uit de resultaten die de sociaal-wetenschappelijke werkwijze ons heeft gegeven (zie tabel 1).

4.2. Formulering van beslissingsregels

De beslissingsregels worden geformuleerd overeenkomstig de in de vorige paragraaf beschreven procedure: indien $p_{p|A} > p_p$ voorspellen we een positieve beslissing, anders een negatieve. Tabel 1 geeft aanleiding tot de volgende beslissingsregels:

<i>regel % laag beroep</i>	<i>als percentage laag beroep hoog, dan neem een positieve beslissing, anders neem een negatieve beslissing.</i>
<i>regel verwijzingen BLO</i>	<i>als aantal verwijzingen BLO veel, dan neem een positieve beslissing, anders neem een negatieve beslissing.</i>
<i>regel schoolgroei</i>	<i>als aantal leerlingen toeneemt, dan neem een positieve beslissing, anders neem een negatieve beslissing.</i>

De prestatie-index van deze drie regels is respectievelijk .20, .20 en .09.

Het zal duidelijk zijn dat er veel meer omgevingskenmerken – (beleids)-uitgangspunten – denkbaar zijn. Echter, in dit artikel gaat het ons meer om de aard van de werkwijze dan om gedetailleerde resultaten.

4.3. Toetsing van de geformuleerde beslissingsregels

De waarde van de tot dusverre gevonden beslissingsregels is bepaald aan de hand van de gegevens waaruit ze zijn afgeleid. Dat is natuurlijk zwak. Zoals we hebben aangekondigd voeren we daarom ten slotte een toetsing uit op onze tweede reeks van waarnemingen. Dit betekent dat we de genoemde beslis-

Tabel 2. De gereproduceerde en voorspelde overheidsbeslissingen op grond van regel laag beroep

Geobserveerd	Regel			
	reproductie		predictie	
	positief	negatief	positief	negatief
positief	21	13	22	13
negatief	5	7	6	4

Tabel 3. De resultaten van de analyse (rep = reproductie en pred = predictie)

Variabele	I_{rep}	I_{pred}
LB	.20	.03
BLO	.20	.17
SG	.09	.09
LB × BLO	.25	.20

singsregels gebruiken om de overheidsbeslissingen in deze reeks te voorspellen. Bij wijze van voorbeeld geven we in tabel 2 voor *regel % laag beroep* weer hoe goed zij de overheidsbeslissingen in de eerste reeks reproduceert, en hoe goed zij deze beslissingen in de tweede reeks voorspelt.

De waarde van de prestatie-index voor de predictie is .03. Dit is aanzienlijk lager dan de prestatie-index voor de reproductie, die .20 was. Alle resultaten zijn samengevat in tabel 3.

4.4. De evaluatie van de prestatie van een beslissingsregel

Allereerst merken we op dat de in tabel 3 samengevatte gegevens zijn gebaseerd op weinig waarnemingen (de prestatie-index van de geformuleerde beslissingsregels I_{rep} is gebaseerd op 46 waarnemingen, en de prestatie-index van de voorspellingen I_{pred} op 45 waarnemingen). Als zodanig zullen we de indices niet beoordelen in termen van statistische significantie (dus, wijken I_{rep} en I_{pred} significant af van de in een of andere nulhypothese gestelde waarde?). Aan de door ons gepresenteerde resultaten kunnen dan ook geen al te stellige conclusies worden verbonden. Dit was echter ook niet de bedoeling. Immers, in dit artikel ligt de nadruk op de werkwijze die wordt gevolgd om de aanwezigheid van expertkennis op te sporen, ofwel het gebruik van

uitgangspunten vast te stellen, en niet zozeer op de resultaten zelf.¹

Niettemin zullen we de resultaten interpreteren; dus toch een indicatie geven van de betekenis die we hechten aan de waarde van de prestatie-indices van de beslissingsregels. De interpretatie is als volgt. We stellen dat een beslissingsregel een goede prestatie levert indien zij de overheidsbeslissingen even goed blijkt te voorspellen als te reproduceren. Met andere woorden, wanneer I_{rep} niet noemenswaard afwijkt van I_{pred} , dan concluderen we dat het aan de beslissingsregel ten grondslag liggende uitgangspunt heeft gewerkt.

Dit maakt conclusies op basis van tabel 3 mogelijk. We doen dit op twee manieren. Ten eerste: conclusies binnen de gegevens, en ten tweede: conclusies over de vraagstelling of expertkennis te evalueren is.

1. Conclusies binnen de gegevens

We zien in tabel 3 dat

- regel LB heeft geen predictieve waarde,
- regel BLO heeft predictieve waarde, en
- regel SG heeft predictieve waarde.

Derhalve kunnen we de volgende conclusies trekken:

- het uitgangspunt gelijkheid wordt *ten onrechte niet* gebruikt (*regel % laag beroep*)
- het uitgangspunt bevorderen van gelijke kansen wordt *terecht wel gebruikt* (*regel verwijzingen BLO*)
- het uitgangspunt prestige wordt *ten onrechte wel* gebruikt (*regel school-groei*)

Er is nog uitbreiding van de conclusies mogelijk als we twee regels combineren, namelijk *regel % lager beroep* en *regel verwijzingen BLO*. We zien dan dat de predictieve waarde gelijk blijft, en dat het percentage laag beroep in combinatie met een ander kenmerk wel iets doet.

2. Conclusies met betrekking tot de vraagstelling

Kunnen we nu stellen dat we hebben gemeten in welke opzichten de overheid wel, en in welke opzichten de overheid juist niet heeft gehandeld als een expert? Het antwoord is bevestigend. De conclusies die we binnen het materiaal hebben getrokken stellen ons in staat de beslisser (overheid) duidelijk te maken dat het uitgangspunt gelijkheid sterker in het geding moet worden gebracht, terwijl het uitgangspunt prestige niet mag werken.

We menen met het bovenstaande te hebben aangegeven dat de door ons geschetste evaluatiemethode vruchtbare perspectieven biedt. Dit geldt in het bijzonder tijdens de ontwikkelingsfase van een expertsysteem. Daarbij is het

immers nodig dat we de over het ambtelijke apparaat verspreide expertkennis opsporen en bundelen. Om deze kennis te evalueren zijn echter, zoals we hebben gezien, onafhankelijke gegevens onontbeerlijk. De volgende gang van zaken lijkt nu zinvol.

Tijdens de ontwikkelingsfase van een expertsysteem kunnen we de evaluatieresultaten terugkoppelen naar de 'experts' (de ambtenaren), die de kennis aandragen. Op deze manier kunnen we stapsgewijze komen tot kennis die voldoende kwaliteit heeft om in het expertsysteem te worden opgeslagen.

Noot

1. Om te bepalen of de prestatie-index van een geformuleerde beslissingsregel significant afwijkt van de in een of andere nulhypothese gestelde waarde I_0 , stellen we de volgende hypothese:

$$H_0 : I_{\text{regel}} = I_0$$

$$H_1 : I_{\text{regel}} > I_0 \quad (1)$$

We toetsen deze hypothese met behulp van de t -toets. De toetsgrootte is als volgt:

$$t = \frac{I_{\text{regel}} - I_0}{S} \quad (2)$$

waarin S de standaarddeviatie is van de verdeling van I_{regel} . Deze standaarddeviatie is gedefinieerd als:

$$S = \sqrt{\frac{1}{n} \left(\frac{1}{4p(1-p)} - I_{\text{regel}}^2 \right)} \quad (3)$$

met n = het totale aantal beslissingen (cf. Hanssen, 1965).

Literatuur

- Alferink, E.A.M.A., Tien jaar VARA-ombudsman. Een empirisch onderzoek naar de invloed van de TV-ombudsman en zijn coalitiepartners op het beleid van de overheid. *Sociologische Gids*, 6, 1980, 483-500.
- Bos, J.M., G.P.A. Braam, Invloed van voetbalclubs op de gemeentelijke overheid; een vergelijkend onderzoek bij clubs uit het betaalde voetbal. *Sociologische Gids*, 6, 1980, 456-482.
- Braam, G.P.A., *Invloed van bedrijven op de overheid; een empirische studie over de verdeling van maatschappelijke invloed*. Meppel: Boom, 1973.
- Breuker, J.A., B.J. Wielinga, *Analysis techniques for Knowledge Bases Systems, Part 1 and 2, Report 1.1 and 1.2*. (Esprit Project 12.) Amsterdam, 1983.
- Coolen, J.A.I., Bejaarden en maatschappelijke dienstverlening: een empirische studie over bereik, inhoud en gevolgen van gezinsverzorging. Enschede: TH Twente, 1982.
- Coolen, J.A.I., J.A. van Pelt, Beheersing van een voorziening via het beslissen over hulpvra-

- gen. Onderzoek naar de indicatiestelling voor het verzorgingstehuis. *Tijdschrift voor Gerontologie en Geriatrie*, juni, 1985.
- Elting, L., M. Even, *Onderwijs, invloed en sociale ongelijkheid: een empirisch onderzoek naar de invloed van scholen op het overheidsbeleid*. (Rapport Instituut voor Onderwijskunde, 1977.)
- Feigenbaum, E.A., Themes and case studies of knowledge engineering. In: D. Michie (red.), *Expert systems in the micro-electronic age*. Edinburgh: University Press, 1979.
- Hanssen A.W., An objective method for forecasting thunderstorms in the Netherlands. *Journal of Applied Meteorology*, 4, 1965, 172-177.
- Inbar, M., *Routine decision-making: the future of bureaucracy*, Beverly Hills, London: Sage Publications, 1979.
- Kuipers, W.J.A., Over de noodzakelijkheid van weersvoorspellingen, die ingesteld zijn op de gebruiker. De Bilt: KNMI, 1954.
- Shortliffe, E.H., *Computer-based medical consultations: MYCIN*. New York: Elsevier, 1976.
- Simon, H.A., *Administrative behavior, A study of decision-making processes in administrative organisation*. New York: Free Press, 1976.
- Simon, H.A., *The Sciences of the artificial*. Cambridge: MIT Press, 1981².
- Steinbruner, J.D., *The cybernetic theory of decision: new dimensions of political analysis*. Princeton, NJ: Princeton University Press, 1974.
- Woodcock, F., The evaluation of yes/no forecasts for scientific and administrative purposes. *Monthly Weather Review*, 104, 1976, 1209-1214.