

# WORKING P A P E R

---

## Richting een beleidsevaluatie cultuur in Nederland

Identificatie van knelpunten en  
oplossingsrichtingen

STIJN HOORENS  
WIJA OORTWIJN

WR-277

August 2005

This product is part of the RAND Europe working paper series. RAND working papers are intended to share researchers' latest findings and to solicit additional peer review. This paper has been peer reviewed but not edited. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.

# Voorwoord

---

Dit 'Working Report' is het resultaat van een verkennend onderzoek naar knelpunten in de huidige praktijk van beleidsevaluatie. Het vermoeden bestaat dat burgers en bedrijven vaak worden geconfronteerd met onnodig strenge regels of administratieve rompslomp. Met behulp van evaluatie kan de effectiviteit van deze regelgeving worden ingeschat om vervolgens, waar nodig, te worden aangepast. Het blijkt echter dat ondanks een rijke evaluatiehistorie nog steeds niet gesproken kan worden van een evaluatiecultuur in Nederland. Er bestaan namelijk verschillende knelpunten in de wijze waarop momenteel beleidsevaluatie in de praktijk wordt vormgegeven. Daarbij is het niet alleen belangrijk om het evaluatieproces valide, bruikbaar en betrouwbaar in te richten, maar ook is het noodzakelijk om ervoor te zorgen dat het proces en de uitkomsten geloofwaardig zijn voor de buitenwacht. Om hieraan gehoor te geven zou meer aandacht moeten worden gegeven aan de verschillende visies en daaraan gerelateerde normen en waarden van verschillende actoren. Hierbij is het belangrijk dat niet opnieuw het wiel wordt uitgevonden. Wereldwijd is hierover relevante literatuur verschenen en elk land heeft zijn eigen ervaringen met en visie op het beoordelen van gevoerd beleid. Derhalve pleiten we voor een vervolgstudie gericht op het in kaart brengen van zowel succesvolle als minder succesvolle internationale ervaringen met beleidsevaluatie, waarmee Nederland zijn voordeel kan doen om de huidige beleidspraktijk te verbeteren.

Dit rapport is gebaseerd op literatuuronderzoek en interviews met een groot aantal actoren betrokken bij beleidsevaluatie in Nederland. De lijst met geïnterviewde personen is weergegeven in Bijlage 1. Wij zijn de geïnterviewde personen zeer erkentelijk voor hun waardevolle bijdrage aan en betrokkenheid bij het onderzoek. Voorts willen wij onze collega's Robert Thomson en Tom Ling bedanken voor hun suggesties en commentaar op een eerdere versie van dit rapport.

Het onderzoek is uitgevoerd door RAND Europe, de Europese vestiging van RAND Corporation ([www.rand.org](http://www.rand.org)). RAND Europe is een onafhankelijk instituut voor beleidsonderzoek dat zich inzet voor het verbeteren van beleidsvorming en het informeren van het publieke debat. Voor meer informatie over dit onderzoek kunt u contact opnemen met Stijn Hoorens ([hoorens@rand.org](mailto:hoorens@rand.org)) of Wija Oortwijn ([oortwijn@rand.org](mailto:oortwijn@rand.org)).

RAND Europe  
Newtonweg 1  
2333 CP Leiden  
Nederland  
Tel. (+31) 071-524 5151

# Inhoudsopgave

---

<b>Voorwoord</b> .....	<b>iii</b>
<b>Samenvatting</b> .....	<b>v</b>
<b>HOOFDSTUK 1 Inleiding</b> .....	<b>1</b>
1.1 Zorgen om kwaliteit.....	1
1.2 Richting een beleidsevaluatiecultuur in Nederland .....	2
<b>HOOFDSTUK 2 Welke normatieve eisen worden aan evaluatie gesteld?</b> .....	<b>4</b>
2.1 Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid.....	4
2.2 Eisen aan dekking en kwaliteit van evaluatie.....	4
<b>HOOFDSTUK 3 Hoe wordt evaluatie bij verschillende ministeries vormgegeven?</b> .....	<b>7</b>
<b>HOOFDSTUK 4 Welke knelpunten bestaan er in de evaluatiepraktijk?</b> .....	<b>9</b>
4.1 Bruikbaarheid .....	10
4.2 Validiteit en betrouwbaarheid .....	12
4.3 Geloofwaardigheid .....	14
<b>HOOFDSTUK 5 Naar een betere beleidspraktijk</b> .....	<b>19</b>
5.1 Conclusies.....	19
5.2 Vervolgstappen .....	21
<b>Literatuur</b> .....	<b>23</b>
<b>Bijlage 1: Interviews</b> .....	<b>25</b>
<b>Bijlage 2: Beleidsevaluatie bij ministeries</b> .....	<b>27</b>

# Samenvatting

---

Op basis van dit rapport kan geconcludeerd worden dat er in Nederland een stijgende aandacht voor beleidsverantwoording is, maar dat de coördinatie van beleidsevaluatie en de kwaliteit van evaluatiestudies hier en daar te wensen over laat. Het is duidelijk dat het genereren, gebruiken en rapporteren van beleidsinformatie maatwerk vereist. Wellicht is dit de reden dat beleidsevaluatie veelal als een 'ad hoc' proces wordt ervaren, zonder eenduidige aanpak. Desalniettemin is het van belang dat de overheid een aantal voorwaarden stelt aan de kwaliteit van beleidsinformatie. Kwaliteitseisen betreffen bruikbaarheid, validiteit en betrouwbaarheid.

Gebleken is dat de meeste ministeries voldoende aandacht schenken aan validiteit en betrouwbaarheid van evaluaties, maar dat de objectiviteit en onafhankelijkheid van evaluaties door verschillende stakeholders als knelpunten worden ervaren. Het voornaamste knelpunt in de huidige praktijk is echter de geloofwaardigheid van evaluaties. Het blijkt dat in veel gevallen geen transparante procedure wordt gevolgd en stakeholders hebben het idee dat ze worden buitengesloten. Op basis van literatuuronderzoek en gesprekken met stakeholders hebben wij vervolgens een aantal voorwaarden geïdentificeerd waaraan moet worden voldaan om een geloofwaardige evaluatie te bereiken.

Samengevat blijkt dat de beleidsevaluatiepraktijk niet zonder problemen is. Indien Nederland wil voldoen aan de noodzaak tot betere beleidsverantwoording en kwalitatief betere beleidsevaluatie blijken generieke doch strikte richtlijnen zoals de Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (RPE) niet altijd effectief te zijn. De Algemene Rekenkamer en het kabinet onderkennen dat regels niet vanzelf leiden tot goede beleidsevaluatie. Wij onderstrepen deze bevinding en zijn van mening dat het voor een betere evaluatiepraktijk nodig is dat evaluatie wordt geïntegreerd in de beleidscyclus: er moet een evaluatiecultuur ontstaan, die zich kenmerkt door interdepartementale afstemming, transparantie en geloofwaardigheid, maar waarin wel ruimte wordt gelaten voor flexibiliteit en maatwerk. Inzichten uit het buitenland betreffende succes- en faalfactoren van beleidsevaluatie kunnen hieraan een waardevolle bijdrage leveren.

Evaluatie is een belangrijk instrument om inzicht te krijgen in de kwaliteit van wet- en regelgeving teneinde bestaand beleid te verbeteren en overbodig beleid tegen te gaan. Of, zoals Pawson en Tilley (1997) het formuleren, om na te gaan: “*what works better for whom in what circumstances, and why.*” Geïnspireerd door de noodzaak tot bezuiniging en het verbeteren van de allocatie van overheidsgelden (Rist, 1990), werd beleidsevaluatie in de jaren zeventig van de vorige eeuw in Nederland geïntroduceerd vanuit de Verenigde Staten. Terwijl er tot begin jaren negentig van de vorige eeuw er zo goed als niet aan effectonderzoek gedaan werd (Hoogerwerf, 1977; Leeuw, 1992), registreert het Evaluatie Overzicht Rijksoverheid (EOR) inmiddels zo’n 250 evaluaties per jaar (Ministerie van Financiën, 2004b). Niet alleen in Nederland wordt evaluatie ingezet als instrument voor het verbeteren van de effectiviteit van regelgeving en beleid. Voor alle voorgenomen EU-regelgeving of voorstellen van de Europese Commissie wordt sinds kort een zogenaamd ‘*extended impact assessment*’ (EIA) verplicht gesteld (Commission of the European Communities, 2002).

### 1.1 **Zorgen om kwaliteit**

De huidige evaluatiepraktijk is echter niet zonder problemen - er valt een spanningsveld waar te nemen tussen een toenemende aandacht voor de kwaliteit van beleid enerzijds en de problemen die zich voordoen bij het evalueren van beleid. Verschillende actoren, waaronder de volksvertegenwoordiging, de burger en het bedrijfsleven vragen om meer transparantie, meer draagvlak, een grotere publieke betrokkenheid, evidence-based policymaking, en value-for-money. Door buitenstaanders wordt de totstandkoming en evaluatie van beleid veelal als een ‘black box’ en een ‘ad hoc’ proces ervaren. Daarnaast belichten evaluaties regelmatig - door het ontbreken van een systematische en integrale analyse - slechts één invalshoek. Gevolg is dat evaluaties ter discussie staan en het draagvlak voor de resultaten beperkt is. De eenzijdige benadering van privatisering van de taxibranche en de te optimistische inschatting van de ‘*returns on investment*’ van grote infrastructuurprojecten hebben bijvoorbeeld recentelijk geleid tot een beperkt vertrouwen in de uitkomsten van deze evaluaties. De toenemende bezorgdheid over de kwaliteit van wetgeving en evaluatie blijkt uit tal van recente publicaties en initiatieven, waarbij drie aspecten van de evaluatiecyclus kunnen worden onderscheiden:

### 1. De kwaliteit van beleid/regelgeving

- Een onderzoek van de Algemene Rekenkamer ('Tussen beleid en uitvoering', 2003b) waarin wordt vastgesteld dat de overheid veel aandacht voor het maken van beleid heeft, maar weinig voor de uitvoering en resultaten.
- De oprichting van het programma 'Andere Overheid' onder aanvoering van de Minister van Bestuurlijke Vernieuwing met als doel de hele overheid efficiënter en beter te laten werken (Ministerie van Binnenlandse Zaken, 2005).

### 2. De kwaliteit van evaluaties

- Een meta-evaluatie van het Ministerie van Financiën ('Evaluatie ex post onderzoek; Bijlage bij het eindrapport VBTB-evaluatie', november 2004b) waarin wordt geconcludeerd dat de kwaliteit en met name de betrouwbaarheid van veel evaluatieonderzoeken onder de maat is.
- Het rapport van de Commissie Duivesteijn (december 2004), waaruit blijkt dat bij de besluitvorming rond grote infrastructuurprojecten enkel rooskleurige evaluatierapporten naar de kamer werden doorgespeeld.

### 3. De relatie tussen beleid en evaluatie

- Het rapport "Kennis voor beleid - beleid voor kennis" van de Adviesraad voor Wetenschaps- en Technologiebeleid (mei 2005) waarin wordt gesteld dat departementen in hun beleidsontwikkeling nog onvoldoende gebruik maken van a) verschillende perspectieven en inzichten uit onafhankelijke kennisbronnen en van b) betrouwbare en valide kennis op een transparante manier.

## 1.2 Richting een beleidsevaluatiecultuur in Nederland

Ondanks de historie die Nederland rijk is aan beleidsevaluatie kan er nog steeds niet worden gesproken van een evaluatiecultuur. Indien de Nederlandse overheid zich serieus wil richten op evidence-based besluitvorming – een beweging die bijvoorbeeld in het Verenigd Koninkrijk door de Labourregering van Tony Blair enkele jaren geleden onder het motto *'what matters is what works'* in gang is gezet – dan zal er een cultuuromslag noodzakelijk zijn. Hiervoor is het eenvoudigweg verbeteren van evaluatiemethoden of het volgen van een checklist niet voldoende. Het vergt een holistische omslag van gedragsnormen en -waarden door alle actoren die betrokken zijn bij beleidsevaluatie. Beleidsmakers moeten duidelijke beleidsdoelen formuleren; verantwoordelijke beleidsambtenaren dienen bereid te zijn om hun programma's kritisch tegen het licht te (laten) houden; professionele evaluatie-experts moeten hun werk vormgeven op een wijze die zowel wetenschappelijk correct als beleidsrelevant is; en stakeholders dienen betrokken te zijn bij de evaluatie en zouden argumenten aangedragen moeten krijgen die vertrouwen scheppen in de evaluatie.

Om meer inzicht te krijgen in de problematiek en mogelijkheden ter verbetering van de huidige evaluatiepraktijk in kaart te brengen, heeft RAND Europe een verkennende studie uitgevoerd, waarbij een analyse van relevante literatuur en opvattingen van betrokkenen over de evaluatiepraktijk in Nederland centraal stonden (zie Bijlage 1 voor een overzicht

van geïnterviewde actoren). In deze analyse hebben wij ons beperkt tot ex-post beleidsevaluatie. Echter, gezien onze opvatting dat er cultuuromslag nodig is die leidt tot een integratie van de beleids- en evaluatiecyclus, hebben de conclusies ook betrekking op ex-ante evaluatie en periodieke monitoring van evaluaties. Tevens pretendeert dit rapport geen overzicht te geven van hoe een ex-post beleidsevaluatie dient te worden uitgevoerd. Hoewel we er hier en daar niet aan ontkomen om enkele methoden te noemen, wordt in deze gevallen naar beschikbare wetenschappelijke literatuur verwezen. Wij hebben hierbij - gezien het doel van het rapport- geen volledigheid nagestreefd.

Dit rapport begint met een korte schets van de achtergrond en geschiedenis van beleidsevaluatie in Nederland. Vervolgens wordt in een overzicht van de wijze waarop de verschillende Nederlandse ministeries evaluatie uitvoeren duidelijk dat ondanks de vele initiatieven en goede bedoelingen er nog geen sprake is van een gecoördineerde evaluatiecultuur in Nederland. Vervolgens is een aantal typen knelpunten in de evaluatiepraktijk te identificeren, welke aan de hand van enkele voorbeelden zal worden toegelicht. Tenslotte wordt bediscussieerd welke stappen dienen te worden genomen om te komen tot betere beleidsverantwoording en kwalitatief betere beleidspraktijk.

## HOOFDSTUK 2 **Welke normatieve eisen worden aan evaluatie gesteld?**

---

Er kunnen verschillende typen evaluatie worden onderscheiden. Ex ante evaluatie is onderzoek naar de verwachte kosten en baten van mogelijke beleidsalternatieven. Hierbij wordt een inschatting gedaan van de mogelijke effectiviteit en efficiëntie van de onderzochte maatregelen. Ex post evaluatie daarentegen maakt juist een inschatting van de effectiviteit en de doelmatigheid van reeds uitgevoerd of in werking zijnd beleid (tussentijdse evaluatie). Tenslotte kan er ook sprake zijn van tussentijdse evaluatie, waarbij de tussenstand wordt opgemaakt of monitoring van de beleidseffecten. De verkennende studie is gericht op de praktijk van ex-post evaluatie.

### 2.1 **Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid**

In 2001 heeft het Ministerie van Financiën richtlijnen vastgelegd voor de uitvoering van evaluaties. Kernbegrippen bij deze Regeling Prestatiegegevens en Evaluatieonderzoek Rijksoverheid (RPE) zijn *doelbereiking, doeltreffendheid en doelmatigheid*. De Regeling onderscheidt verder twee belangrijke evaluatie-instrumenten: systemen van reguliere prestatiegegevens (monitoringsystemen) en periodiek evaluatieonderzoek, waaronder ex ante en ex post evaluaties (Grimmius en Oostdijk, 2004). De RPE volgt uit het eerder door de overheid geïnitieerde traject VBTB (Van Beleidsbegroting Tot Beleidsverantwoording, 1999). De kern van het VBTB proces is gericht op de begroting en verantwoording van uitgaven, waarbij een expliciete koppeling wordt gelegd tussen beleidsdoelstellingen, de te realiseren prestaties en het benodigde budget. Ook in het Regeerakkoord van het Kabinet Balkenende II 'Meedoen, meer werk, minder regels' zijn efficiënt en effectief beleid als speerpunten opgenomen. Dit heeft geleid tot een grote toename van het ex-ante en ex-post evaluatieonderzoek in Nederland. Op dit moment zijn in het Evaluatie Overzicht Rijksoverheid (EOR) ongeveer 1.500 evaluatieonderzoeken opgenomen.

### 2.2 **Eisen aan dekking en kwaliteit van evaluatie**

De RPE stelt eisen aan evaluatiestudies die de effectiviteit van gevoerd beleid moeten beoordelen en maakt daarbij onderscheid tussen de *dekkingseisen* en *kwaliteitseisen*. Om te beginnen met dekkingseisen - de RPE schrijft voor dat alle beleidsdoelstellingen in de regel eens per vijf jaar dienen te worden geëvalueerd. Het is de vraag of aan deze eis wordt voldaan. Om te kunnen nagaan of alle beleidsterreinen daadwerkelijk periodiek worden



afgedekt door beleidsevaluaties zijn namelijk voorzieningen nodig. Uit een rapport van de Algemene Rekenkamer (2000) blijkt echter dat geen enkel Ministerie een sluitend systeem heeft om de afdekking van beleid te monitoren. Volgens de deelnemers van een recent symposium<sup>1</sup> over de RPE is deze dekkingseis te zwaar. Het aantal onderzoeken stijgt wel, maar de kwaliteit blijft van de evaluaties achter en er wordt nauwelijks effectenonderzoek uitgevoerd. De deelnemers waren van mening dat “de huidige dekkingseis niet zinvol is en kan leiden tot (nog) meer evaluaties van onvoldoende kwaliteit.”

Wat betreft de kwaliteit van evaluatie stelt de RPE eisen aan de *validiteit, de betrouwbaarheid en de bruikbaarheid*. Ondanks het feit dat er met introductie van VBTB en de RPE een grote stap is gezet richting het verantwoorden van overheidsbeleid, lijken veel evaluatiestudies de eisen niet te kunnen inwilligen (zie Hoofdstuk 4). Bovendien vergt elk type (evaluatie)onderzoek zijn eigen kwaliteitseisen. Daarom wordt op de verschillende ministeries op uiteenlopende wijzen invulling aan de RPE gegeven. Hoe en in welke mate deze invulling verschilt, wordt in het volgende hoofdstuk beschreven.

---

<sup>1</sup> VIDE (2004) discussiebijeenkomst inzake Kwaliteitsverbetering van evaluaties. 22 september. Samenvatting beschikbaar op: <http://www.videnet.nl/?page=26642>).



## HOOFDSTUK 3 **Hoe wordt evaluatie bij verschillende ministeries vormgegeven?**

---

Beleidsvaluatie is van groot belang om inzicht te krijgen in de resultaten van het beleid dat door de ministeries is gevoerd, de inspanningen die daarvoor geleverd zijn en de middelen die daarmee gemoeid gaan. Kwalitatief goede beleidsvaluatie vormt derhalve een belangrijk onderdeel van het beleidsproces.

In de Comptabiliteitswet en het Handboek Financiële Informatie en Administratie is vastgelegd dat beleidsvaluatie bij de directies Financieel-Economische Zaken (FEZ) zijn ingebed (Algemene Rekenkamer, 2000), terwijl verantwoordelijkheid voor de kwaliteit en systemen voor prestatiegegevens bij de afdeling FEZ, (meerdere) beleidsdirecties of bij beleidsdirectie Beleidsvaluatie, dan wel een combinatie hiervan ligt (Algemene Rekenkamer, 2004). Dergelijke fragmentatie bemoeilijkt een interdepartementale coördinatie van beleidsvaluatie. Bovendien kan dit leiden tot onduidelijkheden over de eindverantwoordelijkheid van het evaluatietraject en de onafhankelijkheid van de afdeling die verantwoordelijk is voor de uitvoering. Hierdoor ontstaan er potentiële risico's voor de kwaliteit van de beleidsinformatie van de departementen in een meer beleidsmatige verantwoording.

In aanvulling op de analyse van de Algemene Rekenkamer (2005; 2004; 2003a), hebben wij op basis van interviews (zie Bijlage 1) geïnventariseerd hoe bij de verschillende ministeries vorm wordt gegeven aan de uitvoering van beleidsvaluaties. In Tabel 1 (zie bijlage 2) zijn de resultaten van deze inventarisatie globaal weergegeven.

Het is niet mogelijk om op basis van deze verkenning een rijksbrede uitspraak te doen.<sup>2</sup> Wel is op basis van deze inventarisatie een aantal dingen duidelijk geworden. Slechts enkele ministeries (Justitie, Defensie, LNV en Buitenlandse Zaken) hebben een aparte directie of een ander organisatieonderdeel voor beleidsvaluatie. Daarnaast blijkt dat evaluaties vaak 'ad hoc' plaatsvinden en onvoldoende invulling wordt gegeven aan de RPE. Uit het rapport van de Algemene Rekenkamer (2004) blijkt bovendien dat alle ministeries met uitzondering van Economische Zaken op onderdelen geen uitwerking hebben van de RPE. Beleidsmedewerkers zijn zelfs niet of nauwelijks op de hoogte van de RPE, hetgeen een risico voor kwaliteitseisen betekent. Tevens is duidelijk geworden dat evaluaties regelmatig

---

<sup>2</sup> De rapporten van de Algemene Rekenkamer en de interviews hebben namelijk betrekking op specifieke beleidsdirecties van de ministeries.

– door het ontbreken van een systematische en integrale analyse – een beperkte scope hebben.

Recent onderzoek van de Algemene Rekenkamer (2005) laat bijvoorbeeld zien dat slechts 41% van alle beoogde maatschappelijke effecten concreet geformuleerd is in de begrotingen van de ministeries<sup>3</sup>. Daarnaast is het onduidelijk op welke manier en op basis van welke informatie ministeries structureel evaluaties uitvoeren. Wel is duidelijk geworden dat in 66% van de gevallen<sup>4</sup> (een deel van het) evaluatieonderzoek door externe partijen wordt uitgevoerd, waarbij ex ante evaluatie onderzoek in beperkte mate wordt uitgevoerd (Algemene Rekenkamer, 2004). De uitkomsten van evaluatieonderzoek worden, met uitzondering van het ministerie van Algemene Zaken, gerapporteerd aan de ambtelijke en politieke top volgens een globale standaardprocedure. De verantwoordelijke afdeling voor de kwaliteit van evaluatieonderzoek verschilt per ministerie, zo is bij het ministerie van Algemene Zaken de afdeling FEZ hiervoor verantwoordelijk, terwijl bij het ministerie van VROM de directoraten-generaal volledige verantwoordelijkheid hebben. Bij de meerderheid van de ministeries ligt echter de verantwoordelijkheid bij zowel de afdeling FEZ als beleidsdirecties, waarbij de rol van de FEZ met name toezichhoudend is (Algemene Rekenkamer, 2004).

Op basis van onze bevindingen en die van de Algemene Rekenkamer (2004) kan worden geconcludeerd dat de ministeries op de goede weg zijn om kwalitatief goede beleidsinformatie aan te leveren volgens de richtlijnen van de RPE. In vergelijking met 2002 - toen de RPE van kracht werd - hebben verschillende ministeries vorderingen gemaakt met de uitwerking van de RPE (Algemene Rekenkamer, 2004). Er zijn echter ook tekortkomingen te identificeren ten aanzien van de wijze waarop de RPE door de ministeries wordt uitgewerkt en toegepast. De resultaten in Tabel 1 laten zien dat de vertaling van richtlijnen voor evaluatie bij de Rijksoverheid een gefragmenteerd beeld vertoont. In het volgende hoofdstuk wordt duidelijk dat naast een beperkte coördinatie in de normatieve aanpak ook de dagelijkse praktijk van beleidsevaluatie te wensen over laat.

---

<sup>3</sup> Het percentage weerspiegelt de mate waarin de beleidsinformatie over de gerealiseerde effecten in de departementele jaarverslagen van 2004 aansluit op de beleidsinformatie over beoogde effecten in de begroting van 2004 (Algemene Rekenkamer, 2005).

<sup>4</sup> In het betreffende onderzoek van de Algemene Rekenkamer (2004) zijn 51 evaluatieonderzoeken van één of twee beleidsdirecties per ministerie betrokken.

## HOOFDSTUK 4 **Welke knelpunten bestaan er in de evaluatiepraktijk?**

---

Ondanks de introductie van verscheidene beoordelingskaders wordt de kwaliteit van dergelijke evaluaties door verschillende betrokkenen vaak in twijfel getrokken, zoals bijvoorbeeld in de evaluatie van de RPE (Ministerie van Financiën, 2004a). Ten aanzien van de eisen die de RPE stelt aan de *validiteit, de betrouwbaarheid en de bruikbaarheid* van evaluaties constateert het Ministerie van Financiën dat:

- De effecten van beleid meestal niet goed worden ingeschat (validiteit en betrouwbaarheid);
- Beleidsdoelstellingen vaak niet goed (concreet) worden geformuleerd (bruikbaarheid);
- Dat verplichte evaluatie elke vijf jaar leidt tot irrelevant onderzoek (bruikbaarheid);
- De rapportage veelal van onvoldoende kwaliteit zijn (validiteit en betrouwbaarheid); en
- Onafhankelijkheid van de evaluatie onvoldoende wordt geborgen (validiteit en betrouwbaarheid).

Op basis van onze analyse stellen wij dat, behalve deze drie, een vierde criterium van cruciaal belang is voor een kwalitatief goede evaluatie, namelijk *geloofwaardigheid*. Zonder een geloofwaardige evaluatie zal er immers weinig draagvlak zijn voor beleid (swijzigingen) of juist het uitblijven daarvan als gevolg van de evaluatie. Momenteel wordt er slechts zijdelings rekening gehouden met geloofwaardigheid als onderdeel van de "betrouwbaarheid" van evaluaties. De ontvanger van beleidsinformatie bepaalt echter in eerste instantie of informatie geloofwaardig is. Geloofwaardigheid kan met andere woorden niet als een intrinsieke kwaliteit van de informatie zelf worden beschouwd.<sup>5</sup>

Met het doel inzicht te verschaffen in de praktijk van evaluatie in Nederland wordt de uitwerking van de criteria bruikbaarheid, validiteit, betrouwbaarheid en geloofwaardigheid van beleidsevaluaties toegelicht aan de hand van enkele praktijkvoorbeelden.<sup>6</sup>

---

<sup>5</sup> Bijdrage Peter van der Knaap, (2005). e-mail correspondentie. 22 Juli 2005

<sup>6</sup> Hierbij is gebruik gemaakt van een drietal evaluatiestudies, waarbij is gekozen voor een studie die onlangs is afgerond (Evaluatie Wijziging Taxiwet 2001), een studie momenteel wordt uitgevoerd (Evaluatie Wet

#### 4.1 Bruikbaarheid

De bruikbaarheid van evaluatie heeft betrekking op de mate waarin de evaluatie gebruikt kan worden voor het (verbeteren van) beleid. De bruikbaarheid van evaluatiestudies hangt af van een aantal aspecten. Om te beginnen is ex-post evaluatie pas zinvol als er reeds in de beleidsvoorbereiding aandacht aan is besteed (Dekker, 2004). Immers, men kan de effectiviteit van beleid pas inschatten als de doelstellingen vooraf duidelijk zijn geformuleerd. Om een goede beleidsevaluatie te doen, moet de onderzoeker niet alleen weten wat de doelstelling van het beleid is, maar ook hoe de beleidsmakers deze doelstelling dachten te bereiken met behulp van het betreffende beleidsinstrument. Een stroming in de beleidsevaluatie *'Realistic Evaluation'*, (o.a. Pawson en Tilley, 1997) richt zich op deze thematiek, waarbij de wijze waarop de effecten worden gegenereerd door specifieke mechanismen in een bepaalde context centraal staat.

Uit de VBTB-evaluatie (Ministerie van Financiën, 2004a) blijkt dat dit veelal niet goed gebeurt. Een voorbeeld is de evaluatie van het Schipholbeleid, waaruit blijkt dat zonder een goed geformuleerde doelstelling het niet eenvoudig is om relevante en irrelevante effecten te onderscheiden.

De evaluatie van het Schipholbeleid kampt met onzorgvuldig geformuleerde doelstellingen ten tijde van de beleidsvoorbereiding. Aan de regio was beloofd dat met de komst van de Polderbaan de geluidshinder zou verminderen, maar er bestaat discussie over de betekenis van deze doelstelling. Sinds de ingebruikname van de polderbaan zijn de vliegbewegingen in bestaande overlastgebieden weliswaar afgenomen, maar in een aantal gemeenten is nieuwe overlast ontstaan. Betekent een afname van de geluidshinder dat de gemiddelde overlast moet afnemen of dat er nergens een toename mag optreden?

Vervolgens is het beoordelen van de effecten pas zinvol als duidelijkheid bestaat over hoe de meetresultaten er uit hadden gezien indien de beleidsmaatregel niet was ingevoerd (*counterfactual*). Indien het bepalen van de counterfactual, zoals vaak wordt aangegeven, onmogelijk of zeer moeilijk is, is het minstens van belang te weten hoe de relevante effecten zich verhouden tot de oorspronkelijk situatie voor invoering van de maatregel (nulmeting). In de meta-evaluatie door het Ministerie van Financiën (2004b) ontbreekt in alle gevallen een counterfactual, hetgeen niet toegelicht noch gerechtvaardigd wordt. Bovendien ontbreekt in tweederde van de wetevaluaties, die door het WODC (Klein Haarhuis et al, 2005) onder de loep zijn genomen, de nulmeting.

---

Afbreking Zwangerschap), en tenslotte een studie die binnenkort zal aanvangen (Evaluatie Schipholbeleid; ten tijde van de analyse (Dec.'04 -Jan. '05) moest nog met de Schipholevaluatie worden begonnen). Voor deze aanpak is gekozen omdat het idee bestond dat problemen geïdentificeerd konden worden en lessen worden geleerd in verschillende fasen van het proces. De cases zijn onderzocht door middel van een analyse van relevante documenten en interviews met een aantal betrokken actoren. Er wordt uiteraard rekening gehouden met het feit dat dit slechts een beperkte steekproef is van het geheel aan evaluatiestudies. Hoewel veel lessen kunnen worden gegeneraliseerd, is het bewijs te beperkt om deze resultaten te vertalen naar andere beleidsgebieden. Daarom is eveneens een aantal meta-evaluaties bekeken met een grotere steekproef aan evaluaties. Bijvoorbeeld: Ministerie van Financiën (2004a, 2004b), Klein Haarhuis et al. (2005), Olsthoorn-Heim (2003) en Wittebrood en van Heemst (2004).

Het ontbreken van nulmetingen is deels te verklaren door het feit dat beleid over het algemeen incrementeel tot stand komt. Dit pleit voor een integratie van de beleidscyclus en de beleidsevaluatiecyclus. Tot op heden zijn deze cycli grotendeels gescheiden doorlopen. Ministeries en parlement vestigen de meeste aandacht op het maken van nieuw beleid, terwijl de belangstelling voor evaluatie ontbreekt. Volgens de VBTB-evaluatie (Ministerie van Financiën, 2004a) kenmerken veel ex-post evaluaties zich door zogenaamd 'vinkgedrag' - zonder toegevoegde waarde - om te voldoen aan de richtlijnen van de RPE.

Het zou voor de hand liggen om bij een ex-post evaluatie gebruik te maken van ex-ante evaluatie, die beiden verplicht worden gesteld in de RPE. Van een koppeling tussen ex-ante en ex-post evaluatie is echter nauwelijks sprake (Dekker, 2004; Grimmus en Oostdijk, 2004). Ex-ante evaluatie wordt vaak overgeslagen en vindt zelfs, afgaande op de gegevens in het Evaluatie Overzicht Rijksoverheid (EOR) – niet of nauwelijks plaats, zoals ook blijkt uit de evaluatie van de Taxiwet.

In het geval van de gewijzigde Taxiwet werd tijdens de beleidsvoorbereiding te weinig aandacht besteed aan 1) de manier waarop transacties op de taximarkt tot stand komen en 2) de voorwaarden waaronder klanten tot een geschikte keuze kunnen komen. Vanwege een transparantieprobleem kon er destijds geen goede inschatting worden gemaakt van mogelijke gevolgen van de liberalisering van de taximarkt.

Een ander aspect dat betrekking heeft op de kwaliteitseis bruikbaarheid betreft de afbakening van de evaluatie. Het is niet eenvoudig, om niet te zeggen onmogelijk, om alle directe, indirecte, bedoelde en onbedoelde effecten te meten. Echter wil de evaluatie bruikbaar zijn, is het minstens noodzakelijk om alle *relevante* beleidseffecten in kaart te brengen. Zo werden bij de evaluatie van het prostitutiebeleid de neveneffecten onvoldoende onderzocht volgens Klein Haarhuis et al. (2005), terwijl deze juist op dit beleidsterrein van cruciaal belang zijn voor het succes of falen van overheidsinterventies. Er bestaan verscheidene oorzaken voor een te beperkte afbakening van de evaluatie. Zo werden in het kader van de evaluatie van de Taxiwet omwille van de consistentie in metingen belangrijke concessies gedaan. De taxistandplaatsen bij Schiphol werden hierdoor buiten beschouwing gelaten, terwijl hier de grootste problemen bestonden.

Aan een evaluatie vanuit een beperkt perspectief kunnen ook politieke en budgettaire motieven ten grondslag liggen. Uit de meta-evaluatie van het Programma Regelgeving van ZonMw (Olsthoorn-Heim, 2003) blijkt dat het ministerie een duidelijke rol heeft in de afbakening van de opdracht en het bepalen van welke aspecten benadrukt worden in de evaluatie. Daarbij kan de betrokken beleidsambtenaar pragmatisch te werk te gaan, waarbij deze zich soms lijkt te laten beïnvloeden door de mate waarin een onderwerp controversieel is. Bovendien vormen gebrek aan financiële middelen en daarmee analytische instrumenten een belangrijke oorzaak voor het onvolledig inschatten van de beleidseffecten. "De aard en inhoud van de [evaluatie]opdracht [wordt] in belangrijke mate bepaald door het beschikbare budget" (Olsthoorn-Heim, 2003), hetgeen ook blijkt uit het rapport van de Algemene Rekenkamer (2005). Een gerelateerd aspect betreft de waarneming dat het in een groot aantal gevallen onmogelijk is om een richtlijn of wet afzonderlijk te evalueren omdat ze vaak onderdeel zijn van een instrumentenmix op een bepaald beleidsterrein. Dit geldt voornamelijk wanneer er geen counterfactual meting beschikbaar is. In de huidige

praktijk wordt er weinig gekeken naar dergelijke aanpalende wetgeving of het overkoepelend beleid als geheel (Olsthoorn-Heim, 2003). De evaluatie van de Taxiwet is hiervan een goed voorbeeld.

In de evaluatie van de Taxiwet is niet onderkend dat veel taxibedrijven - door de verplichte openbare aanbesteding van contracten boven de 225.000 euro - in financiële moeilijkheden waren gekomen. Velen van hen moesten daardoor diensten onder kostprijs aanbieden. De maatregelen als gevolg van de gewijzigde taxiwet kwamen daardoor extra hard aan bij deze stakeholders.

Tenslotte blijkt uit verscheidene bronnen dat een onevenwichtige tijdsplanning ten koste gaat van de bruikbaarheid van de evaluatie. Zo komt uit de VBTB-evaluatie naar voren dat het eens in de vijf jaar verplicht stellen van evaluaties leidt tot een vloed aan irrelevant onderzoek (Ministerie van Financiën, 2004b). De meta-evaluatie van negen evaluaties van wet- en regelgeving door het WODC (Klein Haarhuis et al, 2005) geeft aan dat de wetten of wetswijzigingen nog maar kort voor het evaluatieonderzoek zijn ingevoerd. Daarom kan er van effecten nog geen sprake zijn. In geval van de Taxiwet eiste de Tweede kamer van het ministerie van Verkeer en Waterstaat dat deze reeds 19 maanden na wijziging zou worden geëvalueerd. Volgens betrokkenen is dit irreëel omdat een wet enkele jaren (5-10 jaar) dient te zijn geïmplementeerd alvorens deze geëvalueerd kan worden. De ontwikkelingen in de sector kunnen uiteraard wel al van meet af aan worden gemeten (monitoring).

## 4.2 Validiteit en betrouwbaarheid

Voor het waarborgen van de validiteit en betrouwbaarheid van evaluatiestudies dient aan een aantal eisen te worden voldaan. Simpel gezegd verwijst de validiteit van evaluatie naar de mate waarin het gemeten kenmerk overeenkomt met het bedoelde concept, dat wil zeggen meet het instrument daadwerkelijk datgene wat de onderzoeker beweert te meten.<sup>7</sup> Daarnaast kan een evaluatie als *betrouwbaar* worden beschouwd wanneer bij herhaalde toepassing van het meetinstrument bij dezelfde onderzoekssubjecten leidt tot dezelfde uitkomsten. Een vorm van betrouwbaarheid is bijvoorbeeld dat een andere onderzoeker met dezelfde methode tot gelijke resultaten moet kunnen komen (interbeoordelaars betrouwbaarheid). Hieruit volgt dat objectiviteit een belangrijke voorwaarde is voor een betrouwbaar onderzoek.

De RPE heeft validiteit en betrouwbaarheid geoperationaliseerd in de vorm van een aantal - voor de hand liggende - kwaliteitscriteria, die bij de uitvoering van evaluatieonderzoeken in acht moeten worden genomen. Voorbeelden van validiteiteisen betreffen: valide afbakening van het systeemobject, heldere definities van gehanteerde begrippen, meerjarige vergelijkbaarheid, valide analyse en opstellen van conclusies en/of aanbevelingen, etc. Voorbeelden van eisen met betrekking tot de betrouwbaarheid zijn onafhankelijke gegevensbronnen, waarborging van de onafhankelijkheid en professionaliteit van de

---

<sup>7</sup> Er bestaat uitgebreide literatuur over de verschillende typen validiteit en de voorwaarden om deze te garanderen in een onderzoek. Bijvoorbeeld: Cook and Campbell (1979); Swanborn (1999), Rossi; Freeman and Lipsey (1999).



uitvoerende partij, de opgeleverde informatie is in overeenstemming met de werkelijkheid, etc. (zie ook Algemene Rekenkamer, 2003b).

In de recente meta-evaluatie “De staat van beleidsinformatie” door de Algemene Rekenkamer (2005) zijn criteria voor een selectie van evaluaties geïnventariseerd. Uit deze publicatie blijkt dat er veel aandacht is voor kwaliteitscriteria. Ook in -de door ons onderzochte - case studies zijn weinig validiteitproblemen geïdentificeerd. Het is echter moeilijk de kwaliteit van het onderzoek uit de beschikbare rapportages te achterhalen. In de meta-evaluatie van de VBTB (Ministerie van Financiën, 2004a; 2004b) worden echter vraagtekens geplaatst bij de kwaliteit van 20 onderzochte evaluatiestudies. Ook Wittebrood en van Beem (2004) concludeerden dat de door hen geïnventariseerde maatregelen nauwelijks systematisch zijn geëvalueerd en dat de meeste effectevaluaties niet voldeden aan de minimale kwaliteitseisen om uitspraken te kunnen doen over de effecten van een bepaalde maatregel. Desalniettemin lijkt de validiteit van de analyses over het algemeen niet het grootste knelpunt.

Er is echter wel één probleem en dat betreft objectiviteit. Een gedegen analyse die zich laat leiden door dubbele agenda's en strategisch gedrag heeft weinig waarde. Uit de interviews met verschillende actoren en publicaties over de evaluatiepraktijk in Nederland (bijv. Ministerie van Financiën, 2004a; 2004b) komt het beeld naar voren dat de aandacht voor een onafhankelijke uitvoering van evaluaties momenteel tekort schiet. De voornaamste oorzaak hiervan ligt in de positie van het verantwoordelijke orgaan, te weten het betreffende ministerie. Het Ministerie heeft een dubbelrol - ten aanzien van de voorbereiding en implementatie van de wetgeving enerzijds en de evaluatie ervan anderzijds. Verscheidene malen is naar deze situatie verwezen als “de slager keurt zijn eigen vlees”.

Het valt moeilijk te bewijzen, maar het ligt niet voor de hand dat ambtenaren heel kritisch zullen zijn ten opzichte van het door henzelf geïnitieerde beleid. Om te beginnen hebben beleidsmakers vaak geen onderzoekservaring. Daarnaast hangt er veel van de uitkomsten van evaluaties af. Er worden immers carrières mee gemaakt en gebroken. Bovendien heeft het ministerie het beleid van de Minister en Staatssecretaris te verdedigen, die op hun beurt weer zijn gebonden aan het regeerakkoord. Tenslotte kunnen andere ministeries druk uitoefenen om evaluaties te sturen.

Zo wordt gesuggereerd dat de negatieve geluiden over de liberalisering van de taximarkt het ministerie van Economische Zaken in haar privatiseringsdrift zeer slecht uit kwamen. Uiteraard bestaat er voor het ministerie van Economische Zaken voldoende onderhandelingsruimte met het ministerie van Verkeer en Waterstaat om elkaar mogelijk van dienst te zijn. Zonder dat er daadwerkelijk sprake is van strategisch handelen of dubbele agenda's, leent de positie van het ministerie zich zeer moeilijk voor de rol van objectieve evaluator.

De bijzondere positie van het ministerie heeft vervolgens onmiskenbare gevolgen voor de kwaliteit van de analyse. Uit een meta-evaluatie van 42 evaluaties op het gebied van sociaal veiligheidsbeleid blijkt dat evaluatieonderzoeken, uitgevoerd door medewerkers van de organisatie die de maatregel heeft ingevoerd, relatief de minste kwaliteit hebben (Wittebrood en Beem, 2004). Deze constatering geldt volgens de auteurs voor verslaglegging, maar ook voldeed geen van de uitgevoerde evaluaties aan de minimale eisen

van interne validiteit.<sup>8</sup> Er bestaan natuurlijk controlemechanismen om objectiviteit van het proces te waarborgen, maar deze worden slechts sporadisch ingezet. Soms worden begeleidingscommissies geïnstalleerd.

Een voorbeeld van een onafhankelijke commissie is de commissie voor de procesbegeleiding die in het kader van de evaluatie van het Schipholbeleid door het projectbureau is ingesteld. De commissieleden hebben geen van allen inhoudelijke betrokkenheid om vooringenomenheid te voorkomen. Om neutraliteit en een internationaal perspectief te waarborgen wordt een buitenlandse expert om het terrein van luchtvaart aan de commissie toegevoegd.

Een alternatief is het uitbesteden van het evaluatieonderzoek aan externe bureaus, hetgeen in 66% van de evaluaties gebeurt (Algemene Rekenkamer, 2005, zie ook voetnoot 4). Uitbesteden wil echter niet per definitie zeggen dat objectiviteit is gewaarborgd.

De resultaten van de evaluatie van de Taxiwet, dat door een extern bureau is uitgevoerd, geven aan dat de kwaliteit van taxidiensten is gedaald, de tarieven zijn gestegen en de omzet in de sector als geheel is gedaald. Op basis van deze gegevens zou men kunnen concluderen dat de privatisering voorlopig heeft gefaald. Het ministerie van Verkeer en Waterstaat concludeerde echter dat men op de goede weg is, maar dat de positieve effecten wat langer op zich laten wachten dan verwacht. Ondanks dat het wellicht te vroeg zou zijn om een definitief oordeel te vellen, had het ministerie aanzienlijk kritischer mogen zijn over het succes van de wijziging van de Taxiwet in 2001.

In kwantitatief evaluatieonderzoek ligt de nadruk vooral op de validiteit van gehanteerde meetinstrumenten. Veel evaluaties zijn echter gebaseerd op kwalitatieve inschattingen, waarbij de aandacht vooral uitgaat naar de validiteit van interpretaties (King, 1994; Bergsma, 2003). Hierdoor is de betrokkenheid van actoren van cruciaal belang. Reason en Rowan (1981) stellen hierover het volgende: *“the only criterion for the ‘rightness’ of an interpretation is inter-subjective – that is to say, that it is right for a group of people who share a similar world.”* Zoals hierboven beschreven, stelt ook de betrouwbaarheidseis strenge voorwaarden aan de geloofwaardigheid van het onderzoek.

### 4.3 Geloofwaardigheid

Daar waar de RPE richtlijn eisen stelt aan de validiteit, betrouwbaarheid en bruikbaarheid van evaluatieonderzoeken, constateren wij dat er aan een additionele eis moet worden voldaan om van waarde te zijn voor beleidsmakers: geloofwaardigheid. Geloofwaardigheid heeft in wezen betrekking op de perceptie die stakeholders en andere betrokken actoren hebben van de validiteit en de betrouwbaarheid van de evaluatie.<sup>9</sup> Geloofwaardigheid is

<sup>8</sup> De term interne validiteit in deze studie verwijst naar: “de mate waarin de gebruikte onderzoeksopzet geschikt is om na te kunnen gaan of een bepaalde maatregel de gewenste verandering heeft veroorzaakt in de sociale veiligheid.” De termen interne en externe validiteit zijn nader uitgewerkt in o.a. Campbell (1978) en Cook en Campbell (1979).

<sup>9</sup> Guba en Lincoln (1981) geven vier criteria voor de geloofwaardigheid van kwalitatieve evaluaties: geloofwaardigheid van de interne validiteit, generaliseerbaarheid van de resultaten, afhankelijkheid van de

nodig voor acceptatie van een evaluatieproces en gerelateerde uitkomsten door de stakeholders. Zonder acceptatie zal er geen draagvlak zijn voor eventueel wijziging of voortzetting van het geëvalueerde beleid.

Een evaluatie kan met de beste intenties en volgens de geldende validiteits- en betrouwbaarheidseisen zijn uitgevoerd, maar als stakeholders hiervoor niet voldoende garanties wordt geboden, kan acceptatie alsnog uitblijven. In zowel de evaluatie van de Taxiwet als van het Schipholbeleid was dit het geval. De voornaamste oorzaak van het negatieve beeld dat stakeholders van de validiteit en betrouwbaarheid van deze evaluaties hebben is de perceptuele objectiviteit. Het is echter niet altijd duidelijk of deze weerstand bij stakeholders een gevolg is van de wijze waarop het evaluatieproces is vormgegeven. De oppositie van belanghebbenden wordt immers vaak reeds gevoed door de (potentieel) negatieve gevolgen die zij van het gevoerde beleid (zullen) ondervinden.

Het ministerie van Verkeer en Waterstaat heeft een lange voorgeschiedenis ten aanzien van het Schipholbeleid, waarbij belangen en doelstellingen politiek gekleurd waren. De verschillende websites<sup>10</sup> van bewonersplatforms tegen geluidshinder van Schiphol geeft een beeld van structureel wantrouwen jegens het ministerie. Desalniettemin wil het ministerie van Verkeer en Waterstaat zelf de regie voeren over de evaluatie. Dit lijkt onverstandig omdat het ministerie zelf partij is en bovendien niet onomstreden is, zoals bijvoorbeeld in de affaire Berkhout.<sup>11</sup> Ook al heeft het ministerie het beste voor met de evaluatie van het onderhavige beleid, de geloofwaardigheid is in het geding.

Ook in de evaluatie van de Taxiwet was het ministerie van Verkeer en Waterstaat 'besmet' door een te grote betrokkenheid bij het dossier. Bovendien kregen enkele stakeholders de indruk dat het kabinet onder aanvoering van het ministerie van Economische Zaken de dereguleringsstrend wilde doorzetten, ongeacht de uitkomsten van de evaluatie. Feitelijk hebben zij hierin gelijk gekregen - de liberalisering is doorgezet, terwijl de evaluatieresultaten lieten zien dat de kwaliteit van dienstverlening was verslechterd, de tarieven waren gestegen en de omzet was gedaald.

Voor een geloofwaardige evaluatie is het voor de uitvoerende instantie derhalve belangrijk aan te tonen dat het onderzoek valide en betrouwbaar is uitgevoerd. Hiervoor zijn enerzijds controlemechanismen nodig, waarvan er reeds enkele zijn besproken. Anderzijds vereist geloofwaardigheid ook transparantie. Dit lijkt geen grote opgave, aangezien men zou verwachten dat het geen grote concessie is om transparantie te verschaffen als de analyse ook daadwerkelijk valide is uitgevoerd. De praktijk blijkt echter anders. Uit een aantal

---

onderzoeksopzet en de controleerbaarheid (zie ook: Altheide en Johnson, 1994). Twee van deze aspecten worden in deze paragraaf besproken: geloofwaardigheid van de interne validiteit en de controleerbaarheid.

<sup>10</sup> Bijvoorbeeld: Platform Vlieg hinder Regio Castricum: [www.vlieghinder.nl](http://www.vlieghinder.nl); Platform Vliegoverlast Amsterdam (PVA): [www.vliegoverlast.nl](http://www.vliegoverlast.nl); Wijkbelangen Velzen-Noord: <http://www.wijkbelangen-vn.nl/geluidshinder.htm>; Platform Overlast Schiphol Uithoorn: [www.platformuithoorn.nl](http://www.platformuithoorn.nl).

<sup>11</sup> De commissie Berkhout die het vliegtuiggeluid rondom Schiphol zou meten, is in 2002 opgestapt. In een artikel in het Parool van 14 december 2002 zegt de heer Berkhout: "We moeten onafhankelijk ons werk kunnen doen en daarvan is geen sprake als de staatssecretaris haar zin krijgt."

interviews en uit de evaluatie van de WAZ blijkt dat het evaluatieproces veelal als een ‘black box’ wordt ervaren.

Het voortraject – de (openbare) aanbesteding van de opdracht, van de WAZ – evaluatie was redelijk transparant. Daarna volgde echter een betrekkelijk ondoorzichtig proces. Externe onderzoekers voeren het onderzoek uit en rapporteren aan de begeleidingscommissie (bestaande uit experts). De onderzoekers en het ministerie van Volksgezondheid, Welzijn en Sport waren echter niet bereid mededelingen te doen over de inhoud noch het proces van de evaluatie. De evaluatie wordt pas openbaar gemaakt wanneer het eindrapport aan het ministerie van Volksgezondheid, Welzijn en Sport wordt aangeboden.

In de Taxiwet- en Schipholevaluatie daarentegen is getracht enige transparantie aan te brengen door websites<sup>12</sup> te openen en de resultaten voor iedereen beschikbaar te maken.

Voorts vergt een transparant evaluatieproces naast beschikbare ook een begrijpelijke en volledige rapportage. Volledig wil zeggen dat ook inzicht wordt verschaft in de voorgestelde aanpak, de aannames, randvoorwaarden, gebruikte bronnen en data en gehanteerde methodieken.

Tenslotte wordt de geloofwaardigheid van evaluatie vergroot wanneer stakeholders worden betrokken in het evaluatieproces. Als beleidsmakers actoren in het beleidsveld negeren, riskeren ze gebrek aan draagvlak of zelfs oppositie. Deze oppositie kan leiden tot het vertragen of zelfs vastlopen van beleid (Edelenbos, 2000). Er bestaat uitgebreide beleidsanalytische literatuur over de wenselijkheid tot stakeholderconsultatie, waarnaar wordt verwezen voor een meer theoretische benadering.<sup>13</sup> Niettemin bestaan er varianten van consultatieve evaluatie waarbij weinig aandacht bestaat voor de validiteits. Indien verkeerd toegepast, kan een dergelijke evaluatie niet meer zijn dan een serie meningen, waarbij ieders ‘waarheid’ -onafhankelijk van het bewijs- even zwaar weegt. Pawson en Tilley (1997), bijvoorbeeld, bekritisieren de “vierde generatie evaluatie“ van Guba en Lincoln (1989) zwaar. Zij stellen dat als alle observaties van de sociale leefomgeving zijn gebaseerd op individuele aannames en percepties, dan bestaat er geen neutraal, feitelijk of definitief beeld van een systeem. Dat wil zeggen dat elke evaluatiecyclus duizenden afzonderlijke gedachten en acties genereert, die onderhevig zijn aan een oneindige reeks beschrijvingen. Dit impliceert dat de onderzoeker gedwongen wordt om subjectief te zijn in het registreren daarvan, hetgeen indruist tegen de essentie van de benadering. Het is aldus van belang om deze valkuilen te voorkomen door vooraf duidelijke randvoorwaarden te stellen aan het betrekken van belanghebbenden.

Aan het betrekken van stakeholders in het evaluatieproces worden momenteel geen eisen gesteld door de RPE, waardoor het waarschijnlijk nauwelijks wordt toegepast in de huidige evaluatiepraktijk. Echter in beleidsvelden waar de oppositie tegen het geldende beleid zeer

<sup>12</sup> Zie respectievelijk [www.taxiwet.nl](http://www.taxiwet.nl) en <http://www.evaluatieschipholbeleid.nl/>

<sup>13</sup> Zie bijvoorbeeld: Van de Riet (2003); Klijn et al. (2000); Wildavsky (1993); Thissen en Twaalfhoven (2001); Rosenhead (1989); De Vries en Van den Heuvel (1998); en Edelenbosch (2000).

groot is, zoals in het Schipholdossier, wordt het betrekken van stakeholders haast noodzakelijk.

Via de speciale evaluatie website van het projectbureau verantwoordelijk voor de Schiphol-evaluatie wordt de mogelijkheid geboden om verbetervoorstellen in te dienen. Iedereen die een voorstel tot verbetering van het stelsel heeft, kan bij de evaluatie worden betrokken, mits wordt voldaan aan enkele voorwaarden.



### 5.1 **Conclusies**

Op basis van onze verkenning kan worden geconcludeerd dat er een stijgende aandacht voor beleidsverantwoording is, maar dat de coördinatie van beleidsevaluatie in Nederland en de kwaliteit ervan hier en daar te wensen over laat.

Het is duidelijk dat het genereren, gebruiken en rapporteren van beleidsinformatie maatwerk vereist. Wellicht is dit de reden dat beleidsevaluatie veelal als een ad hoc proces wordt ervaren, zonder eenduidige aanpak. Desalniettemin is het van belang dat de overheid een aantal voorwaarden stelt aan de kwaliteit van beleidsinformatie. Kwaliteitseisen betreffen bruikbaarheid, validiteit en betrouwbaarheid. Voorwaarden voor een bruikbare, valide en betrouwbare evaluatie kunnen als volgt worden samengevat:

#### **Voorwaarden voor bruikbare evaluatie<sup>14</sup>**

##### *Ontwerp evaluatieproces:*

- Goed geformuleerde beleidsdoelstellingen
- Counterfactual en/of nulmeting
- Geen eenzijdige evaluatie, maar effectschatting in den brede

##### *Geïntegreerde beleids- en evaluatiecyclus:*

- Afstemmen ex-post en ex-ante evaluatie
- Evenwichtig tijdsspanne
- Tijdig beschikbare evaluatie-uitkomsten
- Informatie over gebruiksdoel
- Leesbare rapportage met beknopte, representatieve samenvatting

---

<sup>14</sup> Deels gebaseerd op Ministerie van Financiën (2004b)

**Voorwaarden voor valide en betrouwbare evaluatie:<sup>15</sup>***Ontwerp evaluatieproces:*

- Aanwezigheid onderzoeksplan
- Adequate formulering doelstelling
- Heldere probleemstelling en onderzoeksvragen die aansluiten op de doelstelling
- Consistente en representatieve operationalisatie van de te onderzoeken begrippen
- Een adequate selectie van de te onderzoeken eenheden
- Waarborging van de onafhankelijkheid en objectiviteit

*Geïntegreerde beleids- en evaluatiecyclus:*

- Juiste analyse onderzoeksgegevens
- De onderzoeksvragen, bevindingen, conclusies en aanbevelingen sluiten op elkaar aan

Gebleken is dat de ministeries over het algemeen voldoende aandacht schenken aan validiteit en betrouwbaarheid van evaluaties (Tweede Kamer, 2004). Echter, objectiviteit en onafhankelijkheid van evaluaties worden als knelpunten ervaren.

Het voornaamste knelpunt in de huidige praktijk blijkt echter de geloofwaardigheid van evaluaties. Er wordt veelal geen transparante procedure gevolgd en stakeholders hebben het idee dat ze worden buitengesloten. Een geloofwaardige evaluatie kan worden bereikt door te voldoen aan de volgende voorwaarden:

**Voorwaarden voor geloofwaardige evaluatie:**

- De resultaten (incl. aannames, randvoorwaarden, uitgangspunten) zijn voor iedereen beschikbaar
- Heldere argumentatie van methoden van dataverzameling
- Periodieke voortgangsrapportage
- De documentatie is vindbaar
- Begrijpelijke en volledige inhoud
- Alle relevante stakeholders zijn betrokken
- Gebruik maken van de beschikbare expertise en ervaring bij stakeholders

Hoewel er met introductie van VBTB en de RPE een grote stap is gezet richting het verantwoorden van overheidsbeleid, blijkt dat de beleidsevaluatiepraktijk nog niet zonder problemen is. Indien Nederland wil voldoen aan de noodzaak tot betere

---

<sup>15</sup> Deels gebaseerd op Ministerie van Financiën (2004b)



beleidsverantwoording en kwalitatief betere wetgeving, lijken generieke doch strikte richtlijnen zoals de RPE niet altijd effectief te zijn. De Algemene Rekenkamer en het kabinet onderkennen dat regels niet vanzelf leiden tot goede beleidsevaluatie (Algemene Rekenkamer, 2004). Wij onderstrepen deze bevinding en zijn van mening dat het voor een betere evaluatiepraktijk nodig is dat evaluatie wordt geïntegreerd in de beleidscyclus: er moet een evaluatiecultuur ontstaan, die zich kenmerkt door interdepartementale afstemming, transparantie en geloofwaardigheid, maar waarin wel ruimte wordt gelaten voor flexibiliteit en maatwerk. Hierbij kan de verantwoordelijk niet op één partij worden afgewenteld: de betrokkenheid van stakeholders, politici, topambtenaren, uitvoerende ambtenaren, evaluatie-experts en overige actoren is van essentieel belang in dit traject.

De vraag is vervolgens of dit ideaalbeeld een utopie is. Bestaan er mogelijkheden voor de overheid om de dekking enerzijds en de validiteit, de betrouwbaarheid, de bruikbaarheid en de geloofwaardigheid van evaluatie anderzijds te verbeteren? Zijn er bijvoorbeeld in het buitenland mogelijke aangrijpingspunten?

## 5.2 Vervolgstappen

Een stap in de goede richting zou daarom een internationale review zijn, waarbij 'leerzame ervaringen' uit andere landen in kaart worden gebracht. Dit sluit aan bij de aandacht die het kabinet heeft voor kennisbevordering en -uitwisseling op het terrein van beleidsevaluatie (Algemene Rekenkamer, 2004). Wij zijn van mening dat zowel succesvolle als minder succesvolle ervaringen leerpunten kunnen zijn. Met name de achterliggende oorzaken voor succes of falen kunnen interessante lessen bieden voor de Nederlandse context.

Voor een internationale review dient een analytisch raamwerk ontwikkeld te worden, waarbij de relevante aspecten van beleidsevaluatie - zoals verwoord in deze notitie - worden meegenomen. Tevens dient een selectie van landen worden gemaakt, waarbij ons inziens zeker een vergelijking met de volgende landen zou moeten plaatsvinden:

- Het Verenigd Koninkrijk, dat binnen Europa een koppositie vervult op het gebied van beleidsevaluatie;
- Zweden, Denemarken en/of Finland, waar veel aandacht is voor de integriteit in beleidsverantwoording;
- Canada, dat op wereldniveau een vooraanstaande positie inneemt;
- De Verenigde Staten, de pioniers op het gebied van beleidsanalyse en -evaluatie.

De uitkomst van de internationale review zal, met name door de vertaling van de lessen uit het buitenland, een praktisch handvat zijn om te komen tot een betere evaluatiecultuur in Nederland.



## Literatuur

---

- Adviesraad voor Wetenschaps- en Technologiebeleid (2005). Kennis voor beleid - beleid voor kennis. Rijswijk: Quantes.
- Altheide, D., en Johnson, J. (1994). Criteria for assessing interpretive validity in qualitative research. In N. Denzin and Y. Lincoln (eds.) *Handbook of Qualitative Research*, Thousand Oaks, CA: Sage, 485-499.
- Algemene Rekenkamer (2000) Organisatie van beleidsevaluatie. Tweede Kamer der Staten-Generaal. Vergaderjaar 1999-2000, 27 065, nrs. 1-2. 's-Gravenhage: Sdu Uitgevers.
- Algemene Rekenkamer (2003a). Staat van de beleidsevaluatie 2002. Tweede Kamer der Staten-Generaal. Vergaderjaar 2002-2003, 28 656, nrs 1-2. 's-Gravenhage: Sdu Uitgevers.
- Algemene Rekenkamer (2003b). Tussen beleid en uitvoering: lessen uit recent onderzoek. Tweede Kamer der Staten-Generaal. Vergaderjaar 2002-2003, 28 831, nr. 2. 's-Gravenhage: Sdu Uitgevers.
- Algemene Rekenkamer (2004) Staat van de beleidsinformatie 2004. Tweede Kamer der Staten-Generaal. Vergaderjaar 2003-2004, 29 550, nrs 1-2. 's-Gravenhage: Sdu Uitgevers.
- Algemene Rekenkamer (2005) Staat van de beleidsinformatie 2005. Tweede Kamer der Staten-Generaal. Vergaderjaar 2004-2005, 30 116, nrs 1-2. 's-Gravenhage: Sdu Uitgevers.
- Bergsma, M. (2003) *Betrouwbaarheid en Validiteit van Kwalitatief georiënteerde Operational Audits. – Methoden en technieken die de betrouwbaarheid en validiteit van kwalitatief georiënteerde audits waarborgen*. Erasmus Universiteit Rotterdam, Economische Faculteit.
- Duivesteyn, A., A. Slob, G.P.J. Koopmans, Ch.B. Aptroot en M.J.L.M. Hermans (2004) *Onderzoek naar Infrastructuurprojecten*. Tweede Kamer, vergaderjaar 2004-2005, 29 283, nrs. 5-6. Sdu Uitgevers, 's-Gravenhage.
- Campbell, D. T. (1978) *Experimental design: Quasi-experimental design*. In: W. H. Kruskal and J. M. Tanur, eds., *International Encyclopedia of Statistics*, pp. 299–304. New York: The Free Press.
- Commission of the European Communities (2002). *Communication from the Commission on Impact Assessment*. Brussels: COM(2002) 276 Final.
- Cook, T. D., en Campbell, D. T. (1979) *Quasi-experimentation*. Chicago, Ill.: Rand McNally.
- Dekker B (2004). *Paddestoelen en valkuilen: Faalfactoren in evaluatieonderzoek*. Basis (Mei, Themanummer Beleidsevaluatie). via [http://www.researchvoorbeleid.nl/basis/artikelen\\_beleidsevaluatie-2004/faalfactoren.htm](http://www.researchvoorbeleid.nl/basis/artikelen_beleidsevaluatie-2004/faalfactoren.htm).
- Edelenbos, J. (2000). *Proces in Vorm: Procesbegeleiding van Interactieve Beleidsvorming over Lokale Ruimtelijke Projecten*. Utrecht, Lemna.
- Grimmius, T. K. and A. Oostdijk (2004). "Kans of valkuil? De regeling prestatiegegevens en evaluatieonderzoek rijksoverheid." Basis (Mei, Themanummer Beleidsevaluatie).
- Guba, E.G., en Lincoln, Y. S. (1981) *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.
- Guba, E.G., en Lincoln, Y. S. (1989) *Fourth Generation Evaluation*. Beverly Hills, CA: Sage

- Hoogerwerf, A., Effecten van overheidsbeleid, in: *Beleid en Maatschappij*, jrg. 4, 1977, nr. 11, pp. 302-315.
- King, G. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, N.J. : Princeton University Press.
- Klein Haarhuis, C. M., M. M. J. v. Ooyen-Houben, et al. (2005). *Kennis voor Beleid - Een synthese van 58 (evaluatie)onderzoeken op het gebied van rechtshandhaving*. Den Haag, Wetenschappelijk Onderzoek- en Documentatiecentrum: 68.
- Klijn, E.H., E. van Bueren en J. Koppenjan (2000) *Spelen met Onzekerheid: over Diffuse Besluitvorming in Beleidsnetwerken en Mogelijkheden voor Management*. Delft, Eburon.
- Leeuw, F.L. de (1993) "Institutionalisering van evaluatie-onderzoek bij de overheid" in Donker, M.; Derks, J., (eds.), *Rekenschap. Evaluatie-onderzoek in Nederland, de stand van zaken*. Utrecht: Nederlands centrum Geestelijke volksgezondheid, pp. 35-54.
- Ministerie van Financiën (2001) *Regeling prestatiegegevens en evaluatieonderzoek rijksoverheid*. Den Haag.
- Ministerie van Financiën (2004a) *Eindrapport VBTB-evaluatie - Lessen uit de praktijk*. Den Haag, Interdepartementaal Overlegorgaan Financieel Economische Zaken: 57.
- Ministerie van Financiën (2004b) *Evaluatie ex post onderzoek; Bijlage bij het eindrapport VBTB-evaluatie*. Den Haag, Interdepartementaal Overlegorgaan Financieel Economische Zaken.
- Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Bestuurlijke Vernieuwing en Koninkrijksrelaties (2005) *Persbericht ministerraad. Kabinet: Minder regels en meer ruimte voor dienstverleners*.
- Olsthoorn-Heim, E. T. M. (2003). *Vijf jaar evaluatie regelgeving via ZonMw - Verslag van een analyse van het programma Evaluatie Regelgeving in opdracht van het Bestuur van ZonMw*. Den Haag, ZonMw.
- Pawson, R. en Tilley, N. (1997) *Realistic Evaluation*, Sage, London.
- Reason P, J. Rowan. (1981). *Issues of validity in new paradigm research*. In P. Reason and John Rowan (Eds.), *Human inquiry : A sourcebook of new paradigm research*. New York: John Wiley & Sons: 239-250.
- Riet, O.A.W.T. Van de (2003) *Policy Analysis in Multi-Actor Policy Settings – Navigating Between Negotiated Nonsense & Superfluous Knowledge*. Eburon Publishers, Delft.
- Rist, R. C. (1990). *Program evaluation and the management of government. Patterns and prospects across eight nations*. New Brunswick and London, Transaction Publishers.
- Swanborn, P.G. (1999) *Evalueren*. Amsterdam, Boom: 364.
- Rosenhead, J. (1989) *Old and new paradigms of analysis*. In: J. Rosenhead (ed.), *Rational analysis for a Problematic World: Problem Structuring Methods for Complexity, Uncertainty and Conflict*. West-Sussex, John Wiley & Sons: 341-358.
- Rossi, P.H., Freeman, H.E. en Lipsey, M.W. (1999) *Evaluation. A systematic approach*. 6th edition. Thousand Oaks. California. Sage: 500.
- Thissen, W.A.H. en P.G.J. Twaalfhoven (2001) *Towards a conceptual structure for evaluating policy activities*. *European Journal of Operations Research*, (129): 627-649.
- Vries, M.S. de en H.J. van den Heuvel (1998) *De Kracht van het Besluit*. *Beleidswetenschappen*, 12(4): 335-35.
- Wildavsky, A. (1993) *Speaking Truth to Power: The Art and Craft of Policy Analysis*. Transaction Publishers, New Brunswick and London.
- Wittebrood, K. and M. van Beem (2004). *Sociale veiligheid vergroten door gelegenheidsbeperking: wat werkt en wat niet?* Den Haag, Sociaal Cultureel Planbureau: 102.

## Bijlage 1: Interviews

---

Hieronder volgt een overzicht<sup>16</sup> van de organisaties waarvan wij één of meerdere vertegenwoordigers gesproken hebben in het kader van de verkennende studie:

- Dr. K. Putters (Kamerlid), Eerste Kamer der Staten Generaal. 6 april 2005
- Drs. S. Oostlander (Hoofd) en Mw. E.M. Naborn (onderzoekscoördinator), Onderzoeks- en Verificatiebureau van de Tweede Kamer. 13 Januari 2005 en 30 mei 2005
- Mr. E.P. Nas (plaatsvervangend directeur/coördinator wetgeving) en dhr. E. van Viersen, Ministerie van Economische Zaken. 26 Januari 2005 en 27 April 2005
- Dhr. G. Schwarz (Directeur IRAL) en Mw. D. Bergh (Hoofddirectie Juridische Zaken), Ministerie van Verkeer en Waterstaat. 26 Januari 2005 en 28 April 2005.
- Ir. L.H. Keijts (Directeur-Generaal Water), Ministerie van Verkeer en Waterstaat. 26 Januari 2005
- Mr. J. Demmink (Secretaris-Generaal) Ministerie van Justitie. 17 Januari 2005
- Prof. Dr. F. Leeuw (Directeur) Wetenschappelijk Onderzoeks- en Documentatiecentrum van het Ministerie van Justitie (WODC). 17 Januari 2005 en 30 mei 2005
- Drs. D. Kabel (Hoofd Bureau Beleidsonderzoek), Ministerie van Financiën. 30 mei 2005
- Drs. P.A.J. Oomens (hoofd Directie FEZ, VBTB en Beleidsinformatie), Ministerie van Sociale Zaken en Werkgelegenheid. 3 juni 2005
- Mr. R. Bekker (Secretaris Generaal), Ministerie van Volksgezondheid, Welzijn en Sport. 24 Januari 2005
- Dr. H.E.J. Jorritsma (Plaatsvervangend Directeur), Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie (IOB), Ministerie van Buitenlandse Zaken (IOB). 19 april 2005
- Dr. K. Wittebrood (Senior Onderzoeker), Sociaal Cultureel Planbureau. 3 Mei 2005
- Dr. P. van der Knaap (Directeur) en drs. F. van der Broek (Bureauhoofd), Algemene Rekenkamer. 13 Januari 2005 en 26 mei 2005
- Dr. D. Webbink (Hoofd Afdeling Onderwijs en Wetenschap), Centraal Plan Bureau. 19 april 2005

---

<sup>16</sup> In willekeurige volgorde.

Met betrekking tot de drie case studies hebben we de volgende personen gesproken<sup>17</sup>:

- Th. Van Eck, ROVER – Amsterdam. 17 Januari 2005
- M. Van Leeuwen (Districtssecretaris), KNV Taxi. 17 Januari 2005
- Paul. Kok (Coördinator Taxizaken), Gemeente Amsterdam. 21 Januari 2005
- Th. Van Schaik, DG Personenvervoer - Ministerie van Verkeer en Waterstaat. 17 Januari 2005
- Jan Hofsteenge (Plv. Projectleider Project Monitoring Evaluatie Implementatie Schipholbeleid), DG Luchtvaart - Ministerie van Verkeer en Waterstaat. 26 Januari

---

<sup>17</sup> Geen van de betrokkenen bij de evaluatie van de WAZ heeft ingestemd met een interview. Met enkele vertegenwoordigers van bewonersplatforms tegen geluidsoverlast van Schiphol hebben wij wel gesproken. Zij stemden echter niet in met een semi-gestructureerd interview.

## Bijlage 2: Beleidsevaluatie bij ministeries

---

- *Centrale afdeling*: Heeft het betreffende ministerie een aparte afdeling voor beleidsevaluatie/wetevaluatie?
- *Onafhankelijkheid*: Indien een centrale afdeling aanwezig is - wat is de mate van onafhankelijkheid van de afdeling?
- *Externe experts*: Wordt gebruik gemaakt van externe inhoudelijke experts of experts op het terrein van beleidsevaluatie?
- *Aanpak*: Welk stramien wordt gebruikt voor het uitvoeren van evaluaties?
- *Intensiteit*: Hoeveel evaluaties worden gemiddeld per jaar van elk type afgerond (ex-ante/ex-post en monitoring)?
- *Frequentie*: Hoe vaak worden evaluaties uitgevoerd?
- *Verantwoordelijkheid*: Onder wiens verantwoordelijkheid worden evaluaties in de meeste gevallen uitgevoerd?

Tabel 1. Normatieve vormgeving van beleids evaluatie bij Minsiteries.

Ministerie	Centrale afdeling	Onafhankelijkheid	Externe experts	Aanpak	Intensiteit <sup>18</sup>	Frequentie	Verantwoordelijkheid
Buza	IOB	Apart organisatie-onderdeel	+, bijv. universiteiten, nationale/ internationale consultants	'Logic framework' (input, output, outcome, impact); Heeft RPE systematiek verder geoperationaliseerd	Ex-post: 5	Onderwerpen voor 4-5 jaar vastgelegd; tenminste eens per vijf jaar ex post evaluatie voorgeschreven	Beleidsdirectie (decentraal), IOB (centraal)
EZ	Nee	Onderdeel lijnmanagement	+, bijv. CBS	Onderzoeksmethodologie vaak ontwikkelt door afdeling strategie; VBTB aangrijpingspunt voor een allesomvattend evaluatiesysteem; Heeft RPE systematiek verder geoperationaliseerd (Handboek beleidsevaluatie en prestatiegegevens); Hoofdpijn evaluatiebeleid EZ	Ex-post: 3	Niet altijd duidelijk - wel worden er periodieke beleidsnota opgesteld zoals t.a.v. energiebeleid en telecommunicatie	FEZ, Beleidsdirectie
VenW	Nee, wel programma 'herijking wetgeving'	Onderdeel lijnmanagement	+, bijv. NS	Heeft RPE systematiek verder geoperationaliseerd (departementale richtlijn)	Ex-post: 4 Ex-ante: 2	-	FEZ, Beleidsdirectie
Financiën	Nee	Onderdeel lijnmanagement	+, bijv. universiteiten	Heeft RPE systematiek verder geoperationaliseerd (procedurebeschrijving)	Ex-post: 5	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie
Justitie	WODC	Apart organisatie-onderdeel	+, bijv. universiteiten	Heeft RPE systematiek verder geoperationaliseerd (gericht op ministerie); Ex ante uitvoeringstoets	Ex-post: 5	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie, WODC
VWS	Nee, wel Inspectie	Onderdeel lijnmanagement	+, bijv. universiteiten	Heeft RPE systematiek verder geoperationaliseerd (Handboek RPE VWS)	Ex-post: 4	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie
SZW	Nee	Onderdeel lijnmanagement	+, bijv. CWI	Heeft RPE systematiek verder geoperationaliseerd (nadere regeling) SZW toets	Ex-post: 3 Ex-ante: 2	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie
BZK	Nee	Onderdeel lijnmanagement	+, bijv. Horeca Nederland	Heeft RPE systematiek verder geoperationaliseerd (notitie)	Ex-post: 3 Ex-ante: 2	Tenminste eens per vijf jaar ex	FEZ, Beleidsdirectie

<sup>18</sup> Bron: Algemene Rekenkamer, 2004. Tabel 1.1: betreft aantal beleidsevaluaties bij geselecteerde directies in de periode 2002-2003



Ministerie	Centrale afdeling	Onafhankelijkheid	Externe experts	Aanpak	Intensiteit <sup>18</sup>	Frequentie	Verantwoordelijkheid
				Self-audit voor evaluatieonderzoek		post evaluatie voorgeschreven	
AZ	Nee	Onderdeel lijnmanagement		-	Ex-post: 1	-	FEZ
OCW	Nee	Onderdeel lijnmanagement	+, bijv. adviesorganisaties	RPE	Ex-post: 1 Ex-ante: 2	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie
Defensie	Directie Beleidsevaluatie	Onderdeel lijnmanagement	+, bijv. adviesorganisaties	Heeft RPE systematiek verder geoperationaliseerd (Evaluatiefunctie Defensie)	Ex-post: 3 Ex-ante: 1	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	Directie Beleidsevaluatie
VROM	Nee	Onderdeel lijnmanagement	+, bijv. belastingdienst	-	Ex-post: 4	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	Beleidsdirectie en Directoraten-generaal
LNV	Expertisecentrum	Apart organisatie-onderdeel	+, bijv. Stichting SKAL	Heeft RPE systematiek verder geoperationaliseerd (RPE LNV) Normen beleidsinformatiesysteem	Ex-post: 2	Tenminste eens per vijf jaar ex post evaluatie voorgeschreven	FEZ, Beleidsdirectie, Expertisecentrum

Bronnen: Tweede Kamer (2004). Staat van de beleidsinformatie 2004; interviews met vertegenwoordigers van de ministeries BuZa, VenW, EZ, SZW, VWS, Justitie en Financiën, alsmede met vertegenwoordigers van de Eerste Kamer, Onderzoeks- en Verificatiebureau Tweede Kamer, Algemene Rekenkamer, CPB en SCP