

ARTIKEL

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

*Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan**

1 Inleiding

De afgelopen decennia zijn veel rekenkamers onderzoek naar de resultaten van overheidsbeleid als één van hun hoofdtaken gaan beschouwen. Waar in eerdere perioden vooral het oordelen over de rechtmatigheid van overheidsinkomsten en -uitgaven de kern van rekenkamerwerk uitmaakte, zijn rekenkamers zich onder invloed van New Public Management-denken in toenemende mate gaan richten op onderzoek naar de doelmatigheid en doeltreffendheid van overheidsbeleid. Met het doen van dit type onderzoek, internationaal aangemerkt als *performance audits*, beogen rekenkamers een bijdrage te leveren aan het verbeteren van overheidsbeleid (bijv. English & Skærbæk, 2007; Leeuw, 1996; Pollitt & Summa, 1999;). Doelmatigheids- en doeltreffendheidsonderzoek (vanaf hier DDO) wordt daarnaast internationaal gezien als een effectief instrument om de transparantie en verantwoording over het gevoerde overheidsbeleid te vergroten (Bracci et al., 2015; Johnsen et al., 2019). Volgens sommige critici heeft DDO echter niet aan deze verwachtingen voldaan en is veel rekenkameronderzoek verworden tot een 'hol ritueel' (Kells, 2011, p. 385). Dit roept de vraag op wat in de praktijk de daadwerkelijke impact is van DDO, gedaan door rekenkamers (Bawole & Ibrahim, 2016; Johnsen et al., 2019).

Er is echter weinig systematisch empirisch onderzoek beschikbaar over dit belangrijke vraagstuk (Bonollo, 2019; Cordery & Hay, 2022). Bestaande empirische studies richten zich bovendien vooral op impact van rekenkamers in de context van Angelsaksische en Napoleontische (vaak francofone) ambtelijke systemen (Rana et

* Sjoerd Keulen is strategisch adviseur bij de Algemene Rekenkamer en is bijzonder hoogleraar publieke audit, beleidsevaluatie en verantwoording vanwege de Algemene Rekenkamer aan de Universiteit Leiden. Valérie Pattyn is universitair hoofddocent aan het Instituut Bestuurskunde van de Universiteit Leiden en gastprofessor aan KU Leuven Instituut voor de Overheid. Ze doet fundamenteel en toegepast onderzoek naar het gebruik van kennis (evidence) door overheden, beleidswerk, beleidscapaciteit en beleidsanalyse, en specialiseerde zich in beleidsevaluatie als een specifieke vorm van kennis. Ruud van Druenen is methodologisch adviseur en senior onderzoeker bij de Algemene Rekenkamer en is daarnaast universitair docent bestuurskunde aan de Radboud Universiteit. Hij is gespecialiseerd in onderzoeksmethoden, evaluatie- en rekenkameronderzoek en promoveerde in 2022 op handhaving van staatssteunregels in de Europese Unie. Minya Chan studeerde bestuurskunde aan de Universiteit Leiden en schreef een masterscriptie over de impact van doelmatigheids- en doeltreffendheidsonderzoek bij de Algemene Rekenkamer.

al., 2022). Onze studie is een eerste poging tot systematisch onderzoek in de Nederlandse context.¹

In dit artikel presenteren we de resultaten van een enquête onder beleidsambtenaren van Nederlandse ministeries over de impact van het door de Algemene Rekenkamer uitgevoerde DDO tussen 2018 en 2021. Aanvullend identificeren we ook factoren die volgens de ‘gecontroleerden’ (d.w.z. beleidsambtenaren die onderwerp van onderzoek waren) bijdragen aan de impact van onderzoek. We gebruikten hiervoor een wetenschappelijk gevalideerde enquête die eerder werd toegepast in Canada (Morin, 2014) en België (Desmedt et al., 2017). Gebruik van deze enquête maakt het mogelijk om de impact van DDO in de Nederlandse context te vergelijken met die in België en Canada.

Uit de resultaten van de enquête blijkt dat DDO een beperkt positief effect heeft op de rijksoverheid. We constateren dat het onderzoek vooral een conceptuele impact heeft en beperkter instrumentele impact, althans in de perceptie. Als we kijken naar de verschillende factoren die impact kunnen verklaren, vallen de individuele kwaliteit van de rekenkameronderzoeker, en de wil van medewerkers om te ageren op basis van het onderzoek of de daaruit volgende aanbevelingen op als voorspelers. Bijzonder genoeg spelen het parlement en de media slechts een beperkte rol.

In de volgende paragraaf zetten we uiteen dat impact uit verschillende vormen kan bestaan en presenteren we, op basis van het eerdere werk van Van Loocke en Put (2011), een model voor het verklaren van impact. Vervolgens komt in paragraaf 3 de opzet van de gebruikte enquête aan bod. In paragraaf 4 en 5 presenteren we de belangrijkste resultaten. Afsluitend vergelijken we in paragraaf 6 de onderzoeksuitkomsten van deze replicatiestudie met de uitkomsten van de studies over België en Canada en maken we de link naar de bredere internationale literatuur over de impact van DDO.

2 Het meten van impact en verklarende factoren

Impact genereren vormt de kern van het werk van rekenkamers. De *International Organization of Supreme Audit Institutions* (INTOSAI) stelt dat rekenkamers de opdracht hebben om ‘het verschil te maken in het leven van burgers’ (INTOSAI, 2019). Ook de Algemene Rekenkamer stelt zich als doel om ‘het functioneren en presteren van de rijksoverheid en de daarmee verbonden organen te verbeteren’ (Algemene Rekenkamer, 2021). Impact is echter een zeer algemene term die nadere invulling behoeft. Startpunt voor deze nadere invulling is de definitie van impact van DDO zoals deze door Lonsdale (1999, p. 171) is geformuleerd: ‘het directe of indirecte effect dat of de invloed die een rekenkamer kan hebben als gevolg van

1 Er is wel een handvol studies gedaan naar performance audits in Nederland, maar deze studies richten zich niet op het systematisch meten van impact en zijn verouderd en beschrijvend (Van Loocke & Put, 2011; Summa, 1999) of richten zich op lokale rekenkamers (Weets, 2011; Keulen, 2021).

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

haar *performance audits* op de praktijken, prestaties en cultuur van de gecontroleerde entiteit’.

De conceptualisering van het begrip impact behoeft echter meer uitwerking. Een belangrijke kritiek op de literatuur gericht op het meten van de impact van rekenkameronderzoek is immers ook dat deze geen gebruik maakt van een meer genuanceerde benadering op invloed, zoals ook gebruikelijk is in de belendende evaluatieliteratuur (bijv. Weiss, 1979; Johnson et al., 2009; Alkin & King, 2017). Om deze reden volgen we in onze studie de meer verfijnde conceptualisering zoals deze door Desmedt et al. (2017) is voorgesteld in de Belgische studie. Deze auteurs maken een onderscheid tussen instrumentele, conceptuele en strategische impact van DDO.

Onder instrumentele impact verstaan we de meest zichtbare, directe vorm van verandering bij ‘de gecontroleerde’ tijdens of direct na afronding van rekenkameronderzoek. Men kan klassiek denken aan de acceptatie en implementatie van rekenkameraanbevelingen als belangrijkste indicator. Dit is ook de vorm van impact die in de meeste kwantitatieve studies naar impact van rekenkameronderzoek wordt gebruikt (Van Looke & Put, 2011). Daarbij blijven echter belangrijke aspecten van de impact van DDO buiten beeld en wordt geen recht gedaan aan de complexiteit en soms lange duur van de follow-up (Reichborn-Kjennerud & Vabo, 2017). Om die reden kijken we in deze studie ook naar conceptuele en strategische impact.

Onder conceptuele impact verstaan we een minder tastbare vorm van invloed die vooral het denken over beleid geleidelijk verandert op de middellange tot lange termijn (Desmedt et al., 2017). De tijdsperiode en de subtiliteit van de verandering maken het evenwel lastiger om deze vorm van impact te meten.

We spreken van strategische impact wanneer rekenkameronderzoek wordt aangewend als instrument om de besluitvorming te beïnvloeden (Desmedt et al., 2017). In zo’n geval wordt onderzoek bijvoorbeeld aangewend als munitie om een reeds ingenomen standpunt kracht bij te zetten, bijvoorbeeld bij pogingen om bezuinigingen te voorkomen. Door deze drie vormen van impact te onderscheiden zijn we in staat om een rijker beeld te schetsen van de impact van DDO door rekenkamers. Hierbij moet worden opgemerkt dat DDO in sommige gevallen ook gelijktijdig meerdere vormen van impact kan hebben; de hier gepresenteerde vormen zijn dus niet wederzijds exclusief. Vervolgens is het mogelijk om verschillende categorieën van factoren te onderscheiden, en beschouwen we zowel factoren op micro-, meso- en macroniveau, conform het werk van Van Looke en Put (2011), en zoals toegepast door Morin (2014) en Desmedt et al. (2017) in de studies die als basis dienen voor deze replicatiestudie.

Het microniveau heeft hierbij betrekking op factoren die betrekking hebben op het specifieke onderzoek of onderzoeksproces. Hierbij kijken we allereerst naar kenmerken van de relatie tussen de rekenkameronderzoeker en de ‘gecontroleerde’ (beleidsambtenaar), zoals de kwaliteit van de communicatie en de mate van ver-

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

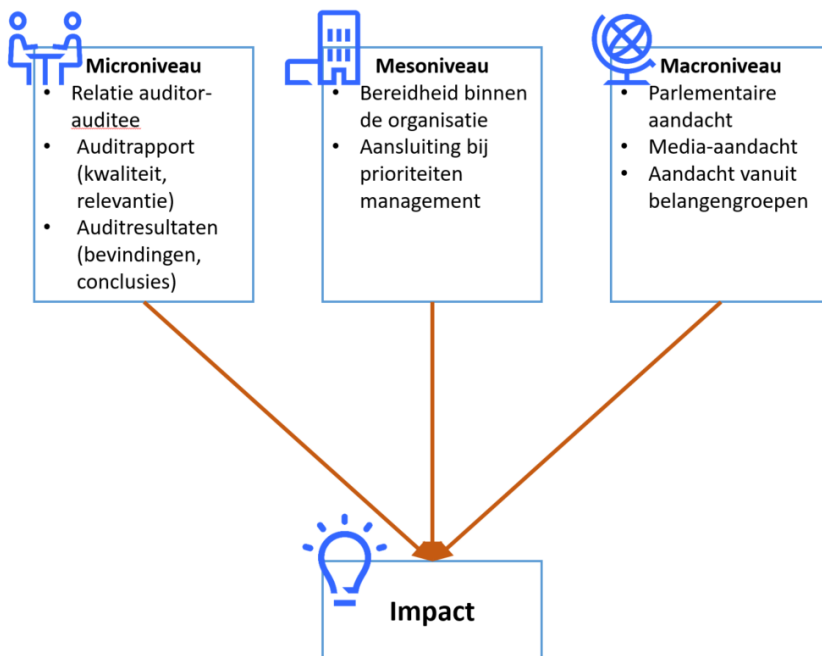
trouwen tussen beiden. Ook de door de gecontroleerde gepercipieerde relevantie en kwaliteit van het onderzoeksrapport zijn relevante factoren op microniveau. Al deze factoren kunnen bijdragen aan een hogere mate van impact (Van Loocke & Put, 2011).

Het mesoniveau heeft betrekking op het organisatieniveau en betreft kenmerken van de rekenkamer of van de gecontroleerde overheidsorganisatie. Hierbij kan men meer bepaald denken aan een positieve houding van de overheidsorganisatie ten opzichte van rekenkameronderzoek, of het feit dat een geplande reorganisatie aansluit bij de uitkomsten van rekenkameronderzoek. Voor de rekenkamer is het hebben van een goede reputatie potentieel van belang voor het genereren van impact (Van Loocke & Put, 2011).

Afsluitend zijn er een aantal factoren op macroniveau te identificeren. Hierbij kijken we naar kenmerken van het politieke systeem waarin de rekenkamer opereert. Belangrijke factoren die kunnen bijdragen aan een hogere impact zijn de mate van druk die door het parlement, media en belangenorganisaties worden uitgeoefend om op basis van het rekenkameronderzoek veranderingen door te voeren (Van Loocke & Put, 2011).

In figuur 1 worden deze factoren overzichtelijk weergegeven.

Figuur 1 Verklarend model voor impact van doelmatigheids- en doeltreffendheidsonderzoek (gebaseerd op Van Loocke & Put, 2011)



De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

3 Methode

Om de impact van DDO in de Nederlandse context te achterhalen hebben we deze replicatiestudie gedaan op basis van een bestaande enquête die ontwikkeld is door Daniëlle Morin voor het onderzoeken van de impact van DDO door de Canadese rekenkamer (Morin, 2008; 2014). Deze enquête stelt ons in staat om zowel een genuanceerd beeld te vormen van de verschillende elementen van impact, als om te achterhalen welke van de hierboven besproken factoren hieraan een bijdrage hebben geleverd. Inmiddels is deze enquête ook toegepast voor onderzoek naar de impact van DDO in de Belgische context (Desmedt et al., 2017). Door zo dicht mogelijk bij de vraagstellingen te blijven van deze andere surveys zijn we in staat om de resultaten in de Nederlandse context (op een indirecte manier) te vergelijken met de eerdere toepassingen in Canada en België.

De afhankelijke variabele (impact van DDO) voor de resultaten uit de enquête, die worden besproken in paragraaf 4, is aan de hand van de volgende acht (sub)variabelen gemeten:

- 1 bijdrage van DDO (instrumenteel/conceptueel);
- 2 relevantie van de aanbevelingen;
- 3 invloed van DDO op de managementpraktijken van de gecontroleerde (instrumenteel);
- 4 bruikbaarheid van de rekenkamerrapporten (instrumenteel/conceptueel/strategisch);
- 5 concrete acties die door de gecontroleerde worden ondernomen naar aanleiding van DDO (instrumenteel);
- 6 organisatorische veranderingen als gevolg van DDO;
- 7 persoonlijke veranderingen als gevolg van DDO;
- 8 gepercipieerde algemene impact van DDO op het management van de gecontroleerde.

Voor de resultaten in paragraaf 5, waarin we ingaan op verklarende factoren voor impact en de correlatie met het 'globaal effect van het AR-onderzoek', gebruikten we dezelfde verklarende factoren uit Van Loocke en Put (2011) als die in de Canadese en Belgische context werden gebruikt (Morin, 2008; 2014; Desmedt et al., 2017) (zie Figuur 1).

De respondenten werd verzocht om te antwoorden op de vragen uit de enquête op een zevenpunts Likertschaal. Op basis van deze scores werd voor elk van de bevraagde variabelen een gemiddelde score berekend door de som van de scores te delen door het aantal respondenten dat de betreffende vraag had beantwoord. Ook hierin hebben we de werkwijze van de eerdere enquêtes in Canada en België zo nauwgezet mogelijk gevolgd.

In paragraaf 4 bespreken we de resultaten ten aanzien van de verschillende vormen van impact aan de hand van rekenkundige gemiddelden. Wij bespreken expliciet de resultaten voor vormen van impact die een 4,5 of hoger scoren op de Likertschaal. Voor deze waarden is een daadwerkelijk effect aannemelijk. Hoewel we hierin de

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

werkwijze van de eerdere studies volgen, hanteren we een hogere grenswaarde voor het vaststellen van een effect. Ter controle presenteren we daarnaast de correlatie tussen de specifieke vorm van impact en 'globaal effect van het AR-onderzoek'.

In paragraaf 5 gaan we in op de correlatie tussen verklarende factoren voor impact en 'globaal effect van het AR-onderzoek'. Ook op dit vlak volgen we in deze replicatiestudie de aanpak van Morin (2014) en Desmedt et al. (2017). We kozen voor deze specifieke correlatiemaat – eerder dan voor de vaak gebruikte Pearsons R – omwille van de niet normaal verdeelde spreiding van meetwaarden (Hauke & Kosowski, 2011). Deze volgt uit het meer ordinale karakter van de meting van onze variabelen op basis van Likertschalen. Waarden $> 0,6$ of $< -0,6$ beschouwen we daarbij als een sterke correlatie. Waarden van $0,4$ tot $0,6$ en $-0,4$ tot $-0,6$ behandelen we als een gematigde correlatie (Schober et al., 2018). Lagere waarden laten we buiten beschouwing, omdat de correlatie in deze gevallen te zwak is. Spearman's rho is daarmee een goede graadmeter voor de betrouwbaarheid van de samenhang tussen verschillende specifieke verklarende factoren en 'globaal effect van het AR-onderzoek' in paragraaf 5. De rekenkundige gemiddeldes en de standaarddeviaties laten in deze paragraaf de spreiding binnen de antwoorden zien.

Voor het invullen van de enquête werden 311 medewerkers van 22 rijksorganisaties die in de periode 2018-2021 betrokken waren bij DDO, aangeschreven. Het gaat hierbij om medewerkers van ministeries en daaraan verbonden agentschappen. Dit heeft uiteindelijk geresulteerd in een responsratio op basis van volledig ingevulde enquêtes van 22,19% ($n = 69$). Meer dan de helft van deze respondenten (55%) was betrokken bij drie of meer rekenkameronderzoeken. De betrokkenheid van de respondenten varieerde van het aanleveren van informatie en de implementatie van aanbevelingen, tot het schrijven van de bestuurlijke reactie op het onderzoeksrapport of het voorbereiden van de beantwoording van parlementaire vragen over het onderzoek.

4 Belangrijkste resultaten voor impact van DDO

De algehele impact van DDO op de rijksoverheid hebben we vastgesteld met de vraag: 'Hoe zou u het globale effect van het AR-onderzoek op uw organisatie beoordelen?' Hierbij gaven de geënquêteerden een gemiddelde score van 4,48/7 ($SD = 1,29$), wat wijst op een beperkt positief effect.

Een belangrijk onderdeel van rekenkamerrapporten zijn de aanbevelingen. Deze worden over het algemeen als enigszins gepast, realistisch en toepasbaar gezien. De respondenten zijn daarnaast overwegend van mening dat de aanbevelingen inderdaad over kernproblemen gaan en tijdig komen, dus vóórdat het beleid gewijzigd is.

We onderzochten eveneens op welke deelterreinen DDO invloed heeft gehad binnen de organisatie en op de individuele ambtenaren die betrokken waren bij het rekenkameronderzoek. Wat opvalt, is dat de invloed van rekenkameronderzoek

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

overwegend conceptueel is. DDO zet aan tot verandering (mean 4,91) en brengt de noodzaak om veranderingen onder de aandacht door te wijzen op zaken die juist wel of nog niet geïmplementeerd zijn (mean 5,07). Net iets minder sterk is dat DDO een nieuw licht kan werpen op de situatie (mean 4,49).

Maar ook strategisch wordt DDO benut. We stellen bijvoorbeeld vast dat rekenkameronderzoeken gebruikt worden om opnieuw aandacht of bevestiging te vragen voor een gedane evaluatie (mean 4,53). In het verlengde hiervan is het wat minder sterke effect (mean 4,44) dat rekenkameronderzoek gebruikt wordt om te wijzen op inconsistenties binnen een programma of organisatie. DDO zet volgens de respondenten daarnaast aan tot verandering, bijvoorbeeld door van de minister, het parlement of het publiek aandacht te vragen voor een bepaald issue (respectievelijk mean 4,66 en 4,67).

Tabel 1 Bijdrage van DDO van de Algemene Rekenkamer

Bijdrage van het AR-onderzoek	Mean (SD)	P
Een nieuw licht werpen op de situatie	4.49 (1.16)	.669**
Aanzetten tot verandering	4.91 (1.09)	.653**
De noodzaak voor verandering onder de aandacht brengen / Geïmplementeerde veranderingen onder de aandacht brengen	5.07 (1.09)	.652**
Een evaluatie van een programma/beleidsmaatregel bevestigen/ onder de aandacht brengen	4.53 (1.08)	.599
Het team/management steunen in de uitvoering van projecten	3.93 (1.30)	.704**
Het wijzen op inconsistenties in sommige programma's of activiteiten van de organisatie	4.44 (1.18)	.447
De aandacht van de minister vestigen op een bepaald probleem	4.66 (1.16)	.491
De aandacht van het parlement en publiek vestigen op een bepaald probleem.	4.67 (1.34)	.445
Feedback krijgen op een onderwerp dat de politiek niet wilde evalueren	3.96 (0.91)	.357
Als voorbeeld dienen voor andere programma's of activiteiten van de organisatie	4.03 (1.18)	.584
De uitvoeringskosten terugdringen	3.46 (1.14)	.412
Verbetering van informatie die gebruikt wordt bij de besluitvorming	4.70 (1.23)	.548
Betrouwbaarder controles implementeren die de rechtvaardigheid van de beslissingen van de organisatie moeten garanderen	4.37 (0.97)	.408
Bestaande controles stroomlijnen om enkel controles te behouden die operationele efficiëntie verzekeren	3.67 (1.00)	.436

** Sterke Spearman's rho-correlatie (> 0,6 of < -0,6)

Tot slot vinden we ook uitingen van instrumentele impact. Deze lijken eerder in de beheersing en de bedrijfsvoering aanwezig te zijn dan in aanpassingen van beleid, maar het effect is beperkt. We zien dit wel terug bij het verbeteren van informatie voor besluitvormingsprocessen (mean 4,70).

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

Dit beeld wordt versterkt als we kijken in welke mate en hoe rekenkameronderzoek wordt gebruikt als instrument om te veranderen, als referentie-instrument, of als onderhandelingsinstrument (Tabel 2). Ook hierbij zien we met name conceptueel gebruik.

Wanneer DDO wordt gebruikt als veranderingsinstrument, wordt dit voornamelijk aangewend om een situatie te verhelderen (mean 4,57) en minder sterk om na te denken over beleidswijzigingen of om van discussie over een onderwerp over te gaan naar actie. Als referentie-instrument geven respondenten aan dat DDO vooral helpt als reflectie-instrument en om te leren van fouten en als verrijking van het organisatiegeheugen (respectievelijk mean 4,63, 4,65 en 4,54). Als onderhandelingsmiddel zien de respondenten vooral heil in DDO als valide basis voor een discussie (mean 4.66). Daarbij wordt het rapport ook gebruikt als argument om in te brengen in discussies met andere partijen (mean 4,51).

Tabel 2 *Bruikbaarheid van rekenkamerrapporten voor verandering, als referentie-instrument en als instrument voor onderhandeling*

Bruikbaarheid van AR-onderzoek om te veranderen	Mean (SD)	P
Een situatie te verhelderen	4.57 (1.45)	.845**
Over te gaan van discussie naar actie	4.36 (1.38)	.774**
Beleid te heroriënteren	4.27 (1.02)	.632**
Overeenkomsten of protocols te ondertekenen	3.88 (0.73)	.398
Bruikbaarheid van AR-onderzoek als referentie-instrument		
Reflectie te stimuleren	4.63 (1.29)	.577
Op een meer objectieve manier te evalueren	4.36 (1.28)	.607**
Standpunten of observaties te valideren	4.44 (1.13)	.632**
Goede managementprincipes te bekrachtigen	4.18 (1.09)	.476
Te leren van fouten	4.65 (1.25)	.512
Bruikbare informatie aan te leveren om bepaalde projecten uit te voeren	4.29 (1.11)	.545
Het geheugen van de organisatie te verrijken	4.54 (1.26)	.552
Bruikbaarheid van AR-onderzoek als onderhandelings-instrument		
Een sterk argument in discussies tussen verschillende partijen	4.51 (1.35)	.636**
Een goede basis voor discussie	4.66 (1.23)	.671**
Een manier om belangrijke overheden of andere partijen onder druk te zetten	4.19 (1.30)	.667**
Een gelegenheid om teams en beleidsmaatregelen/programma's uit te dagen	4.34 (1.09)	.574
Een middel voor de oppositie of belangengroepen	4.30 (0.89)	.201

** Sterke Spearman's rho-correlatie (> 0,6 of < -0,6)

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

De invloed van DDO beperkt zich grotendeels tot beleid. Alhoewel veel onderzoek van de Algemene Rekenkamer ook over bedrijfsvoering gaat, hebben deze rapporten volgens de geënquêteerden geen aanzienlijk effect op het management van de organisatie. Evenmin leidt rekenkameronderzoek tot abrupte wijzigingen binnen de organisatie, zoals reorganisaties of de invoering van nieuwe werkprocessen. De geënquêteerden rapporteren daarom ook nauwelijks negatieve organisatorische implicaties als gevolg van het rekenkameronderzoek. Toename van bureaucratistische procedures, verhoging van uitvoeringskosten en zelfs verlamming van de organisatie werden ook nauwelijks genoemd. De enige organisatieveranderingen waarvan sprake is – zij het dat deze relatie gematigd is – zijn weer gericht op beleid: namelijk beleids- of programmawijzigingen.

Op individueel niveau trad nauwelijks effect van DDO op: de geënquêteerden rapporteren een iets verhoogde motivatie en een beperkt effect op het verhogen van de eigen prestaties. Daarnaast is er een geringe verhoging van het zelfvertrouwen van zowel de geënquêteerde als zijn management.

5 Factoren die van invloed zijn op impact

Uit de resultaten van de enquête komen een aantal factoren, op zowel micro-, meso- als macroniveau, naar voren die in de ogen van de respondenten invloed hebben gehad op de mate van impact van DDO.

5.1 Microfactoren

Het microniveau, waar het draait om de invloed van de individuele rekenkameronderzoeker en om de kwaliteit van het individuele rapport, blijkt in de Nederlandse context de belangrijkste voorspeller van de mate van impact. Alle items die betrekking hebben op de relatie tussen de individuele onderzoeker en de gecontroleerde organisatie, hebben een grote invloed. Hierbij geldt dat als rekenkameronderzoekers als geloofwaardig en legitiem worden beschouwd, de kans op impact hoger is. Ook als de onderzoekers zich open opstellen en prettig communiceren, is dat een belangrijke voorspeller voor een grotere doorwerking van het onderzoek.

Tabel 3 Impact verhogende factoren op microniveau

Relatie met de onderzoekers van de AR	Mean (SD)	P
De geloofwaardigheid van de onderzoekers	4.58 (1.48)	.826**
De legitimiteit van de interventie van de onderzoekers	4.49 (1.32)	.798**
Communicatiestijl tussen de onderzoekers en de gecontroleerde organisatie	4.61 (1.45)	.713**
De openheid van rekenkameronderzoekers	4.61 (1.25)	.735**

** Sterke Spearmans rho-correlatie (> 0,6 of < -0,6)

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

Dat deze individuele karakteristieken van onderzoekers belangrijk zijn voor de doorwerking, wordt verder onderstreept door de factoren die impact verkleinen. Dit gebeurt wanneer de kwaliteit en de empathie van individuele rekenkameronderzoekers als laag worden gezien. Van de elf onderzochte deelfactoren laten er maar liefst negen een matige tot sterke correlatie zien.

De bevindingen die de onderzoekers rapporteren, worden door de geënquêteerden eerder als objectief gezien en de woorden en het gedrag als consequent. Ook zien de ondervraagden dat er een balans is in negatieve en positieve zaken in het rapport en dat de aanbevelingen niet onrealistisch zijn. Verder zagen zij dat de rekenkamer niet uit was op sensatie.

Ook merken we op dat de geënquêteerden vinden dat rekenkameronderzoekers de neiging hebben om meer tijd aan details te besteden dan aan de essentie. Daarnaast hebben rekenkameronderzoekers volgens de respondenten de neiging om niet gevoelig te zijn voor de context, minder subtiel en minder compromisbereid te zijn en was er in een aantal gevallen reden om te twijfelen aan expertise en begrip van het onderzoeksterrein. Al deze genoemde factoren hangen samen met een lagere mate van impact.

Tabel 4 *Impact verlagende factoren op microniveau*

Bronnen van onvrede van de relatie met onderzoekers van de AR	Mean (SD)	P
Aanwezigheid van de onderzoekers verzwaarde de gewone werklust	4.74 (1.30)	-.370
De neiging van de onderzoekers om meer tijd te besteden aan details dan aan essentiële zaken	4.54 (1.49)	-.630**
De neiging van de onderzoekers om een eerder confronterende dan een samenwerkende houding aan te nemen ten opzichte van u [de geënquêteerde]	3.46 (1.54)	-.549**
Aanbevelingen van de Algemene Rekenkamer waren niet realistisch	3.88 (1.42)	-.653**
Een gebrek aan subtiliteit en gevoeligheid voor de context door de onderzoekers, of geen bereidheid tot het sluiten van compromissen	4.08 (1.61)	-.694**
De bevindingen waren niet objectief	3.48 (1.50)	-.725**
De Algemene Rekenkamer zocht sensatie	3.66 (2.04)	-.645**
De expertise van de onderzoekers en hun begrip van het onderzochte domein bleken ontoereikend	4.06 (1.65)	-.734**
Woorden en daden van de onderzoekers waren niet consequent	3.39 (1.55)	-.632**
Onderzoekers hadden een zwakke conceptuele visie	3.43 (1.36)	-.679**
Het rapport vermeldde enkel negatieve aspecten, de positieve zaken werden zelden of nooit vermeld	3.48 (1.46)	-.415

** Sterke Spearmans rho-correlatie (> 0,6 of < -0,6)

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

5.2 Mesofactoren

Het mesoniveau van impact kijkt naar de mate waarin factoren op het niveau van de onderzochte organisatie invloed hebben op de mate van impact. Niet geheel onverwacht heeft DDO in Nederland de meeste impact als de aanbevelingen overeenkomen met de prioriteiten van de onderzochte organisaties. Dit is in lijn met uitkomsten in bijvoorbeeld Noorwegen (Reichborn-Kjennerud & Johnsen, 2015) of Groot-Brittannië (Alwardat, 2010). De juiste timing en de politieke wil om met de uitkomsten van een rekenkameronderzoek iets te doen hangen samen met meer impact.

Verder zien we dat de wil van de medewerkers om te handelen op basis van het rapport en de aanbevelingen een sterke positieve samenhang hebben met impact. Tezamen met het belang dat wordt gehecht aan de geloofwaardigheid, legitimiteit en de openheid en communicatie van de rekenkameronderzoekers laten deze mesofactoren zien dat de impact van DDO sterk afhangt van relationele factoren.

Tabel 5 *Organisatorische factoren die impact verhogen*

Omgevingsfactoren	Mean (SD)	P
Aanbevelingen van de Algemene Rekenkamer ten opzichte van de prioriteiten van het management	4.39 (1.01)	.670**
Bereidheid van de medewerkers	4.21 (1.07)	.650**
Politieke wil	4.51 (0.87)	.413
Bereidheid van de belangrijke overheden	4.27 (0.61)	.271
Timing	4.20 (1.05)	.444
Vertrek van sleutelfiguren	3.89 (0.55)	.054
Hervormingen op overheidsniveau	4.00 (0.57)	.287

** Sterke Spearmans rho-correlatie (> 0,6 of < -0,6)

5.3 Macrofactoren

Op macroniveau wordt in de internationale literatuur doorgaans onderstreept dat parlementaire betrokkenheid (bijv. Desmedt et al., 2017; Johnsen et al., 2019; Morin, 2008; Noussi, 2012; Pollitt & Summa, 1999) en media-aandacht (Desmedt et al., 2017; Morin, 2008; Raudla et al., 2016; Reichborn-Kjennerud, 2013) belangrijke factoren zijn die de impact van DDO kunnen versterken.

Opvallend genoeg hebben we geen duidelijke samenhang gevonden tussen zowel parlementaire betrokkenheid als media-aandacht en de mate van impact die het rekenkameronderzoek binnen de organisatie heeft. Dit is opmerkelijk omdat de Algemene Rekenkamer aan beide zaken veel aandacht schenkt.

Elke vaste Kamercommissie krijgt standaard een briefing aangeboden wanneer de rekenkamer een rapport uitbrengt. 42 respondenten rapporteerden dat een DDO door de Tweede Kamer was besproken. De enige zwakke samenhang die hierbij gemeten is, is het feit dat de behandeling door het parlement leidde tot het implementeren van concrete maatregelen om problemen te corrigeren. Daarnaast hiel-

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

pen DDO's ook als vehikel om een versnelde overgang van discussie naar actie te maken om de problemen aan te pakken die in het rapport werden beschreven.

Hoewel 42 van de 69 respondenten meldden dat een van de DDO's aandacht kreeg in de media, was er geen verband met de mate van impact van de onderzoeken. De enige gematigde samenhang die werd gerapporteerd, is dat de media-aandacht leidde tot een herleving van het politieke debat over het onderzoeksonderwerp.

6 Conclusie en discussie

Dit onderzoek is een van de eerste studies waarin de impact van DDO van rekenkamers op systematische wijze wordt geanalyseerd. Doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer heeft een gematigde impact, met een gemiddelde waarde van 4,48/7 voor de algehele waargenomen impact op de gecontroleerde organisaties. De impact van Nederlandse rekenkamerrapporten is daarmee iets hoger dan die in Canada: 4,2/7 (Morin, 2014), en duidelijk hoger dan de rapporten van het Belgische Rekenhof: 3,2/7 (Desmedt et al., 2017). Al met al blijft de impact evenwel subtiel en vinden er – volgens de respondenten – geen radicale veranderingen plaats als gevolg van de onderzoeken. Wat betreft type impact stellen we in de drie landen wel dezelfde trend vast: overal is de invloed vooral van conceptuele aard, en pas in volgende instantie strategisch en instrumenteel. Ook ten aanzien van deze onderliggende vormen zijn de gemiddelden in Nederland hoger dan in Canada en België.

Als we kijken welke factoren de impact het beste verklaren in Nederland, valt met name de invloed van de individuele rekenkameronderzoeker op. Als rekenkameronderzoekers zich open opstellen en prettig communiceren, is dat een belangrijke voorspeller voor een hogere mate van impact van DDO, zo bleek ook al uit eerder onderzoek (Alwardat, 2010; Raudla et al., 2016; Reichborn-Kjennerud, 2013). Dit effect werd ook gevonden in België (Desmedt et al., 2017). Daar stelde men echter niet vast dat de impact van rekenkameronderzoek minder groot is als de kwaliteit en de empathie van individuele rekenkameronderzoekers lager zijn. Die relatie was in Nederland duidelijker zichtbaar. Verdiepend onderzoek naar de manier waarop de inzet van specifieke gesprekstechnieken of bepaalde omgangsvormen kan bijdragen aan een hogere mate van impact, zou in dit kader van toegevoegde waarde kunnen zijn (zie ook Smith, 2005).

Op mesoniveau zijn de bereidheid van medewerkers in het algemeen en de prioriteiten van het management de belangrijkste verklaringen voor impact. Dit komt sterk overeen met de bevindingen in Canada en België. De Nederlandse casus is daarmee consistent met andere empirische studies, die wezen op de rol van consensus tussen de rekenkamer en de gecontroleerde en het afstemmen van aanbevelingen op managementprioriteiten om de impact te vergroten (voor Noorwegen: Reichborn-Kjennerud & Johnsen, 2015; Reichborn-Kjennerud, 2014; voor Saoedi-Arabië: Alwardat & Basheikh, 2017; voor Roemenië: Ștefănescu & Trincu-Drăgușin, 2020; voor het Verenigd Koninkrijk: Alwardat, 2010).

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

Bijzonder was dat factoren op macroniveau in Nederland geen samenhang vertonen met de mate van impact van rekenkameronderzoek. Hoewel we aan zowel het aantal briefings als aan het aantal mediamomenten zien dat parlement en media interesse hebben in het werk van de Algemene Rekenkamer, heeft dit voor ambtenaren geen invloed op de mate waarin DDO van de rekenkamer tot veranderingen bij de overheid leidt. Wij hadden verwacht dat deze samenhang aanwezig zou zijn, omdat de Algemene Rekenkamer hier zelf sterk op inzet in haar strategie om impact te vergroten. Ook uit eerder onderzoek wisten we dat commissies in de Tweede Kamer rekenkamerrapporten als een van de belangrijkste informatiebronnen zien en dat de jaarlijkse Verantwoordingsdag bijdraagt bij aan de parlementaire belangstelling voor DDO (Keulen, 2020). In België maakt het parlement veel minder gebruik van rekenkameronderzoek en besteedt de pers ook minder aandacht aan rekenkamerrapporten (Desmedt et al., 2017). Toch lijkt de invloed van media en parlement daar groter in het verklaren van impact dan in de Nederlandse casus. Verder bleken ook in Canada, Engeland, de Scandinavische landen en ten dele in Estland de media en/of het parlement een significante invloed te hebben op impact (Brown & Craft, 1980; Johnsen et al., 2019; Raudla et al., 2016). Waarom dat in Nederland niet het geval is, valt moeilijk te verklaren, en maakt verder onderzoek noodzakelijk.

We weten tot nog toe bijvoorbeeld zeer weinig over hoe parlementariërs kennis – en dan vooral in de vorm van evaluaties en rekenkamerrapporten – selecteren en gebruiken (Mavrot & Pattyn, 2022). Dergelijk onderzoek kan op verschillende manieren worden vormgegeven. Hierbij valt bijvoorbeeld te denken aan casestudies naar hoe parlementariërs en hun (ambtelijke) ondersteuning kennis selecteren en omzetten naar bijvoorbeeld Kamervragen (Jorge et al., 2019). Een andere mogelijkheid – gezien de grote hoeveelheden gesproken woord dat in het parlement wordt vastgelegd – is *text mining* van parlementair (niet-)gebruik van extern advies (Sanders et al., 2017).

Verder pleiten we voor vervolgonderzoek waarin op een meer objectieve manier wordt gemeten wat de invloed is van rekenkameronderzoek. Waar onze survey als voordeel had een breed beeld te schetsen van de impact van DDO, zou toekomstig onderzoek zich idealiter kunnen toespitsen op een selectie van bepaalde vormen van impact, waarbij dan kwalitatief zou kunnen worden ingegaan op de causale mechanismen die de link tussen deze factoren en een type van impact zouden kunnen verklaren. Te denken valt bijvoorbeeld aan *process tracing* als methode.

Tot slot vertalen we de conclusies van dit onderzoek naar aanbevelingen voor de praktijk van rekenkameronderzoek, dat niet alleen op nationaal niveau, maar ook in alle waterschappen, provincies en gemeenten wordt uitgevoerd. De individuele rekenkameronderzoeker kan de impact van zijn onderzoek vergroten door geloofwaardig te zijn, open te communiceren en zich empathisch op te stellen. Aan dit aspect wordt nog nauwelijks aandacht gegeven in het opleiden van onderzoekers. Training, bijvoorbeeld via de Vereniging van Rekenkamers, met aandacht voor de persoonlijke vaardigheden van onderzoekers kan hierbij helpen. Daarnaast kunnen rekenkamers in gemeenten – die onderzoek vaak uitbesteden – hier aandacht

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

aan besteden bij het selecteren van onderzoeksbureaus. Verder zijn timing en kennis van de agenda en prioriteiten van de onderzochte organisatie cruciaal. Op mesoniveau zijn prioriteiten van de onderzochte organisatie die parallel lopen met het onderwerp van rekenkameronderzoek, immers de beste voorspeller voor een hogere doorwerking.

Literatuur

- Algemene Rekenkamer. (2021). *Vertrouwen in verantwoording: strategie 2021-2025*. Algemene Rekenkamer.
- Alkin, M. C., & King, J. A. (2017) Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, 38(3), 434-450.
- Alwardat, Y., & Basheikh, A. (2017). The impact of performance audit on public administrations in Saudi Arabia: An exploratory study. *International Journal of Business and Management*, 12(12), 144-157.
- Alwardat, Y. A. (2010). *External auditors and clients: An investigation of perceptions of value for money (VfM) audit practices in the UK public sector* [PhD thesis University of Westminster Harrow Business School, London].
- Bawole, J. N., & Ibrahim, M. (2016). Contesting claims on measuring performance in the public sector using performance audits: Evidence from the literature. *Public Organization Review*, 16(3), 285-299.
- Bonollo, E. (2019). Measuring supreme audit institutions' outcomes: Current literature and future insights. *Public Money & Management*, 39(7), 468-477.
- Bracci, E., Humphrey, C., Moll, J., & Steccolini, I. (2015). Public sector accounting, accountability and austerity: More than balancing the books? *Accounting, Auditing & Accountability Journal*, 28(6), 878-908.
- Brown R. E., & Craft, R. (1980). Auditing and public administration: The unrealized partnership. *Public Administration Review*, 40(3), 259-265.
- Cordery, C. J., & Hay, D. C. (2022). Public sector audit in uncertain times. *Financial Accountability & Management*, 38(3), 426-446.
- Desmedt, E., Morin, D., Pattyn, V., & Brans, M. (2017). Impact of performance audit on the administration: A Belgian study (2005-2010). *Managerial Auditing Journal*, 32(3), 251-275.
- English, L., & Skærbæk, P. (2007). Performance auditing and the modernisation of the public sector. *Financial Accountability & Management*, 23(3), 239-241.
- Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 30(2), 87-93.
- INTOSAI. (2019). *INTOSAI P-12: The value and benefits of supreme audit institutions – making a difference to the lives of citizens*. INTOSAI.
- Johnsen, Å., Reichborn-Kjennerud, K., Carrington, T., Jeppesen, K. K., Taro, K., & Vakkuri, J. (2019). Supreme audit institutions in a high-impact context: A comparative analysis of performance audit in four Nordic countries. *Financial Accountability & Management*, 35(2), 158-181.
- Johnson, K., Greenesid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.

De impact van doelmatigheids- en doeltreffendheidsonderzoek van de Algemene Rekenkamer op de rijksoverheid

- Jorge, S., Jesus, M. A. J. D., & Nogueira, S. P. (2019). The use of budgetary and financial information by politicians in parliament: A case study. *Journal of Public Budgeting, Accounting & Financial Management*, 31(4), 539-557.
- Kells, S. (2011). The seven deadly sins of performance auditing: Implications for monitoring public audit institutions. *Australian Accounting Review*, 21(4), 383-396.
- Keulen, S. J. (2020). 20 jaar Verantwoordingsdag: Inzicht voor Kamercommissies. *Bestuurskunde*, 20(1), 64-75.
- Keulen, S. J. (2021). De doorwerking van lokaal rekenkameronderzoek. *Bestuurswetenschappen*, 75(2), 82-100.
- Leeuw, F. L. (1996). Performance auditing, new public management and performance improvement: questions and answers. *Accounting, Auditing & Accountability Journal*, 9(2), 92-102.
- Lonsdale, J. (1999). *Impacts*. In C. Pollitt & H. Summa (Eds.), *Performance or compliance? Performance audit and public management in five countries* (pp. 171-193). Oxford University Press.
- Mavrot, C., & Pattyn, V. E. (2022). The politics of evaluation. In A. Ladner & F. Sager (Eds.), *Handbook on the politics of public administration* (pp. 244-255). Edward Elgar.
- Morin, D. (2008). Auditors general's universe revisited. *Managerial Auditing Journal*, 23(7), 697-720.
- Morin, D. (2014). Auditors general's impact on administrations: A pan-Canadian study (2001-2011). *Managerial Auditing Journal*, 29(5), 395-426.
- Noussi, K. (2012). *How public accountability is institutionalized: The case of external public auditing in global perspective applying a mixed-methods approach* [Dissertation Universität Wien, Vienna].
- Pollitt, C., & Summa, H. (1999). *Performance audit and public management reform*. In C. Pollitt & H. Summa (Eds.), *Performance or compliance? Performance audit and public management in five countries* (pp. 1-10). Oxford University Press.
- Rana, T., Steccolini, I., Bracci, E., & Mihret, D. G. (2022). Performance auditing in the public sector: A systematic literature review and future research avenues. *Financial Accountability & Management*, 38(3), 337-359.
- Raudla, R., Taro, K., Agu, C., & Douglas, J. W. (2016). The impact of performance audit on public sector organizations: The case of Estonia. *Public Organization Review*, 16(2), 217-233.
- Reichborn-Kjennerud, K. (2013). Political accountability and performance audit: The case of the auditor general in Norway. *Public Administration*, 91(3), 680-695.
- Reichborn-Kjennerud, K. (2014). Performance audit and the importance of the public debate. *Evaluation*, 20(3), 368-385.
- Reichborn-Kjennerud, K., & Johnsen, Å. (2015). Performance audits and supreme audit institutions' impact on public administration: The case of the office of the auditor general in Norway. *Administration & Society*, 50(10), 1422-1446.
- Reichborn-Kjennerud, K., & Vabo, S. I. (2017). Performance audit as a contributor to change and improvement in public administration. *Evaluation*, 23(1), 6-23.
- Sanders, J., Lisi, G., & Schonhardt-Bailey, C. (2017). Themes and topics in parliamentary oversight hearings: A new direction in textual data analysis. *Statistics, Politics and Policy*, 8(2), 153-194.
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768.
- Smith, G. (2005). Communication skills are critical for internal auditors. *Managerial Auditing Journal*, 20(5), 513-519.

Sjoerd Keulen, Valérie Pattyn, Ruud van Druenen & Minya Chan

- Ștefănescu, A., & Trincu-Drăgușin, C. P. (2020). Performance audit in the vision of public sector management: The case of Romania. *Accounting and Management Information Systems*, 19(4), 778-798.
- Summa, H. (1999). *The Subjects of Study*. In C. Pollitt & H. Summa (Eds.), *Performance or compliance? Performance audit and public management in five countries* (pp. 66-78). Oxford University Press.
- Van Loocke, E., & Put, V. (2011). The impact of performance audits: A review of the existing evidence. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing contributing to accountability in democratic government* (pp. 175-209). Edward Elgar.
- Weets, K. (2011). The impact of performance at local government level: A multiple case study within the city audit office of Rotterdam. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance auditing contributing to accountability in democratic government* (pp. 248-267). Edward Elgar.
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426-431.