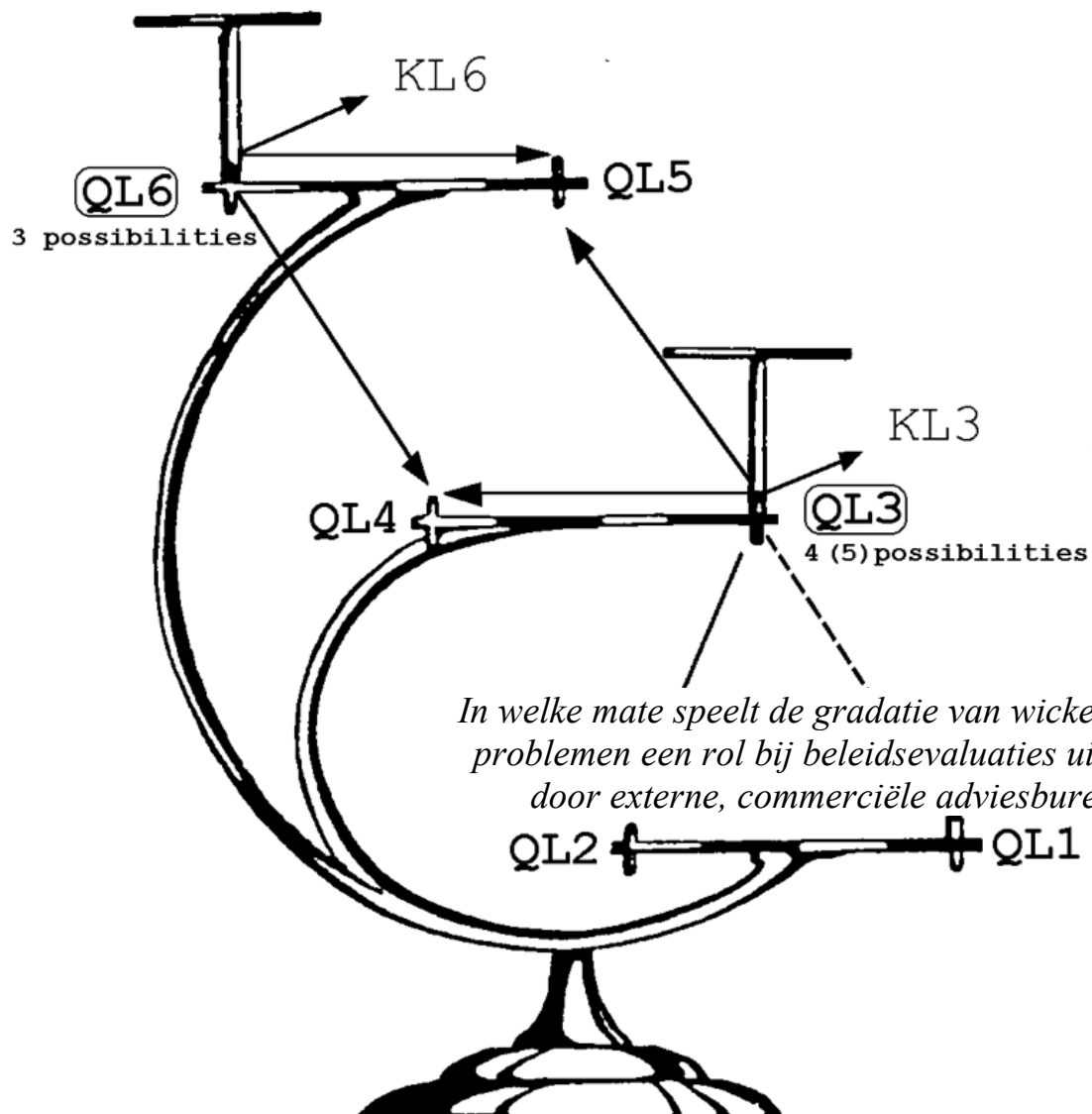


# Wickedness & Beleidsvaluaties



Erasmus Universiteit Rotterdam

Master Bestuurskunde – Publiek Management

Afstudeerscriptie

Pauline Swart

555734

Begeleider: prof.dr. Martijn van der Steen

Tweede beoordelaar: prof.dr. Sandra van Thiel

Augustus 2022

19.924 woorden

## Voorwoord

Beste lezer,

Laat ik, deze toch wel eenzijdige vorm van communicatie, beginnen met een feitje over mezelf: ik ben nogal fan van The Big Bang Theory. Een Amerikaanse tv comedy over Sheldon Cooper, een briljante fysicus met minimale sociale vaardigheden. Sheldon heeft nogal de neiging om dingen ingewikkeld te maken. Het welbekende *rock-paper-scissors* (dat tweevoudig Formule 1 wereldkampioen Fernando Alonso voor het eerst speelde op de respectabele leeftijd van 40 jaar) wordt enkel gespeeld in de vorm van *rock-paper-scissors-lizard-Spock*. En ook schaken gebeurt niet op een gewoon schaakbord, maar wordt enkel gespeeld in 3-D.

Dit onderzoek gaat over wickedness, dat wordt bepaald aan de hand van drie dimensies: *kennis*, *consensus*, en *betrokken actoren*. Het bepalen van wickedness is een constant spel tussen deze drie dimensies. Wat er gebeurt in de ene dimensie heeft invloed op de anderen. Het is een spel waarbij je continu op de hoogte moet zijn van alles wat er in alle dimensies gebeurt. Niet bepaald een opgave waar één enkele strategie je uit gaat redden. Het 3-D schaakbord op de voorpagina is beeldend voor het spel dat continu wordt gespeeld bij wickedness, en zo ook gedurende dit onderzoek. De drie bepalende dimensies van wickedness - *kennis*, *consensus*, en *betrokken actoren* – zijn voor alle beleidsevaluaties hetzelfde, maar de gekozen zetten, en daarmee de uitkomst, varieert.

De spelregels van 3-D chess zijn te complex om zo één, twee, drie uit te leggen, *therefore*:

*"Scissors cuts paper, paper covers rock, rock crushes lizard, lizard poisons Spock, Spock smashes scissors, scissors decapitates lizard, lizard eats paper, paper disproves Spock, Spock vaporizes rock, and as it always has, rock crushes scissors."*

Veel leesplezier,

Pauline

## Samenvatting

In dit rapport staat de volgende vraag centraal: *In welke mate speelt de gradatie van wickedness van problemen een rol bij beleidsevaluaties uitgevoerd door externe, commerciële adviesbureaus?* Aan de hand van een kwalitatieve documentenanalyse van twaalf beleidsevaluaties uitgevoerd door vijf externe, commerciële adviesbureaus ben ik tot de conclusie gekomen dat de mate van wickedness een rol speelt bij de inrichting van het rapport, het gehanteerde evaluatieproces, de gehanteerde instrumenten, en de mate waarin een uitspraak wordt gedaan over de doelmatig- en doeltreffendheid van het geëvalueerde beleid. Deze conclusie is gebaseerd op de analyses van onderstaande deelvragen:

*1. Hoe kunnen twaalf beleidsevaluaties uitgevoerd door vijf externe, commerciële adviesbureaus worden ingedeeld op een schaal van wickedness?*

Aan de hand van de mate waarin kennis beschikbaar is, er consensus aanwezig is, en de betrokken actoren bekend zijn, zijn de evaluaties ingedeeld in de gestructureerde, slecht, matig, of ongestructureerde categorie. Een overzicht van de categorisering is afgebeeld in figuur 2. Deze categorisering is de basis voor de analyse van de overige deelvragen.

### Figuur 2

#### *Beleids-evaluaties op een schaal van wickedness*

Gestructureerd	Slecht gestructureerd	Matig gestructureerd	Ongestructureerd
<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Beleids-evaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>
<i>Evaluatie van de Aanwijzingen inzake de rijksinspecties</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Sturing op duurzaamheid</i>
<i>Evaluatie Convenant onbedwelmd slachten volgens religieuze riten</i>		<i>Evaluatie Radioactief afval</i>	<i>Preventie gehoorschade.</i>
<i>Evaluatie Convenant e-lending</i>			

*2. In welke mate speelt de gradatie van wickedness een rol bij de inrichting van de rapportage van de geselecteerde beleidsevaluaties?*

In de inrichting van de rapportage speelt wickedness een rol bij de plek die betrokken actoren krijgen: in de inleiding of in het resultatenonderdeel. De mate van wickedness speelt ook een rol bij het wel of niet bespreken van het leerproces van de beleidsvoerende organisatie: dit gebeurt enkel bij een hoge mate van wickedness. Ten derde speelt de mate van wickedness een rol bij de plaats en vormgeving van de conclusies en aanbevelingen: hoe meer wicked, hoe meer hiermee wordt gevarieerd.

*3. In welke mate speelt de gradatie van wickedness een rol bij het evaluatieproces gehanteerd in de geselecteerde beleidsevaluaties?*

Betreffende het evaluatieproces speelt de mate van wickedness een rol bij het moment waarop de dataverzameling, data-analyse, en de beoordeling plaatsvinden. Hoe hoger de mate van wickedness, hoe later in het evaluatieproces deze stappen plaatsvinden. Ten tweede speelt de mate van wickedness een rol bij het al dan niet lineaire verloop van dit proces. Hoe hoger de mate van wickedness, hoe minder lineair dit proces verloopt. Ten derde speelt wickedness een rol bij het gebruik van kwalitatieve data en het gebruik van evaluatievragen. Hoe meer wicked, des te vaker deze elementen opzichzelfstaand worden gebruikt, zonder kwantitatieve data of evaluatiekader. Ten vierde speelt de mate van wickedness een rol bij het betrekken van actoren bij de analyse: dit gebeurt alleen in de ongestructureerde categorie. Tot slot kan er een fase worden geïdentificeerd die niet terugkomt in de literatuur: de toetsfase. De mate van wickedness speelt een rol bij de frequentie waarin deze fase wordt toegepast: hoe meer wicked het probleem, hoe vaker een toetsfase plaatsvindt.

*4. In welke mate speelt de gradatie van wickedness een rol bij de instrumenten gehanteerd in de geselecteerde beleidsevaluaties?*

Bij de gehanteerde instrumenten speelt de mate van wickedness een rol bij het gebruik van praktijkbezoeken, dit gebeurt enkel bij een lage mate van wickedness. Ten tweede worden klankbordgroepen en validatie sessies alleen toegepast bij een gemiddelde mate van wickedness. De mate van wickedness speelt ook een rol bij het gebruik van de feiten scan, dit wordt enkel gebruikt in de categorie met de hoogste mate van wickedness. Tot slot speelt de mate van wickedness een rol bij reflectiesessies en het expertpanel, deze worden enkel toegepast in de twee categorieën met een hogere mate van wickedness.

*5. In welke mate speelt de gradatie van wickedness en rol bij de mate waarin de geselecteerde beleidsevaluaties een uitspraak doen over de doelmatig- en doeltreffendheid van het beleid?*

Tot slot speelt de mate van wickedness een rol bij het vaststellen van de doelmatig- en doeltreffendheid. Bij een lage mate van wickedness wordt altijd een oordeel geveld over de doelmatig- en doeltreffendheid wanneer dit als doel wordt gesteld. Terwijl bij een gemiddelde mate van wickedness geen uitspraak wordt gedaan over de doelmatig- en doeltreffendheid, ondanks dat dit wel als doel van de evaluatie wordt gesteld. Bij een hoge mate van wickedness wordt geconcludeerd dat het niet mogelijk is om een uitspraak te doen over de doelmatig- en doeltreffendheid.

Op basis van bovenstaande conclusies doe ik de volgende aanbevelingen.

Ten eerste, aan externe, commerciële adviesbureaus, adviseer ik om de mate van wickedness vast te stellen zodat de evaluatie hierop kan worden aangepast. Ten tweede is het gebruik van kwalitatief onderzoek in toenemende mate van belang naarmate wickedness toeneemt. Adviesbureaus zouden daarom moeten investeren in de kwaliteit en het gebruik hiervan. Ten derde, bespreek bij beleidsevaluaties met een hogere mate van wickedness de verwachtingen en perspectieven van de betrokken actoren. Ten vierde, houdt rekening met een wederkerig evaluatieproces inclusief toetsfase, naarmate wickedness toeneemt. Ten vijfde, bij problemen met een hogere mate van wickedness is het evalueren van de doelmatig- en doeltreffendheid nagenoeg onmogelijk. Het leerproces en het proces waarmee het beleid tot stand is gekomen, zouden onderwerp moeten zijn van de evaluatie. Tot slot, een managementsamenvatting is essentieel, neem dit onderdeel altijd op in rapportages van beleidsevaluaties.

Aan opdrachtgevers van beleidsevaluaties adviseer ik om bij de evaluatievraag meer nadruk te leggen op het proces waarmee het beleid tot stand is gekomen, in plaats van enkel te focussen op doelmatig- en doeltreffendheid.

Tot slot adviseer ik om in toekomstig wetenschappelijk onderzoek de bevindingen uit dit onderzoek op een grotere schaal en met andere methoden te toetsen. Daarbij kunnen externe en interne evaluatoren worden vergeleken en kan worden gekeken naar andere elementen van beleidsevaluaties, zoals het evaluandum of het evaluatiemotief. Ook kan toekomstig onderzoek focussen op de toegevoegde waarde van praktijkbezoeken bij problemen met een hoge mate van wickedness. Aangezien deze problemen zeer contextafhankelijk en uniek zijn, zou dit instrument bij kunnen dragen aan het begrip van het probleem.

# Inhoudsopgave

<b>Voorwoord .....</b>	<b>1</b>
<b>Samenvatting .....</b>	<b>2</b>
<b>1. Inleiding .....</b>	<b>8</b>
1.1 Aanleiding .....	8
1.2 Probleemstelling .....	9
1.3 Doelstelling .....	10
1.4 Wetenschappelijke relevantie.....	10
1.5 Bestuurlijke relevantie .....	11
1.6 Maatschappelijke relevantie.....	11
1.7 Leeswijzer .....	11
<b>2. Theoretisch Kader .....</b>	<b>12</b>
2.1 Wickedness .....	12
2.1.1 Achtergrond van wickedness .....	12
2.1.2 Karakteristieken van wickedness .....	13
2.1.3 Schaal van wickedness .....	14
2.2 Beleid .....	15
2.2.1 Perspectieven op beleid.....	16
2.2.2 Perspectieven op het beleidsproces .....	17
2.3 Beleidsevaluaties .....	18
2.3.1 Perspectieven op beleidsevaluaties.....	18
2.3.2 Inhoud van beleidsevaluaties .....	19
2.3.3 Het evaluatieproces .....	20
2.3.4 Instrumenten voor evaluatieonderzoek .....	21
2.3.5 Rapporteren over beleidsevaluaties .....	21
2.3.6 Kritiek op beleidsevaluaties .....	22
2.4 Externe, commerciële adviesbureaus .....	23
2.4.1 Definitie .....	23
2.4.2 Historische ontwikkeling.....	24
2.5 Conclusie.....	25
<b>3. Methodologie .....</b>	<b>26</b>
3.1 Casestudy.....	26
3.2 Documentenanalyse .....	27
3.3 Betrouwbaarheid & validiteit .....	27
3.4 Biases & overige limitaties.....	28
<b>4. Beleidsevaluaties op een schaal van wickedness .....</b>	<b>29</b>
4.1 Evaluatie van de Aanwijzingen inzake de rijksinspecties .....	29
4.2 Evaluatie Transitie Autoriteit Jeugd.....	30
4.3 Sturing op duurzaamheid. Collegeonderzoek naar de sturing van gemeente Apeldoorn op klimaatadaptatie, circulaire economie en de energietransitie .....	31

4.4 Evaluatie Radioactief afval. Een evaluatie van de Nota Radioactief afval en het Nationaal Programma Radioactief afval.....	32
4.5 Preventie gehoorschade. Aanpak problematische blootstelling versterkte muziek .....	32
4.6 Evaluatie Wet taken meteorologie en seismologie.....	33
4.7 Evaluatie Convenant onbedweld slachten volgens religieuze riten .....	34
4.8 Periodieke evaluatie Gezondheidsraad 2017-2021 .....	35
4.9 Evaluatie Convenant e-lending .....	36
4.10 Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant 2016-2021 (onderdeel Milieu).....	36
4.11 Beleidsvaluatieonderzoek Streekomroepvorming .....	37
4.12 Ministerie van BZK Evaluatie Nationaal Energie Bespaarfonds 2014-2019.....	38
4.13 Conclusie.....	39
<b>5. De Inrichting .....</b>	<b>41</b>
5.1 Gestructureerde categorie.....	41
5.2 Slecht gestructureerde categorie.....	43
5.3 Matig gestructureerde categorie.....	44
5.4 Ongestruceerde categorie.....	45
5.5 Analyse .....	47
5.6 Conclusie.....	48
<b>6. Het Evaluatieproces .....</b>	<b>49</b>
6.1 Gestructureerde categorie.....	49
6.2 Slecht gestructureerde categorie.....	50
6.3 Matig gestructureerde categorie.....	52
6.4 Ongestruceerde categorie.....	53
6.5 Analyse .....	55
6.6 Conclusie.....	57
<b>7. De Instrumenten .....</b>	<b>59</b>
7.1 Gestructureerde categorie.....	59
7.2 Slecht gestructureerde categorie.....	60
7.3 Matig gestructureerde categorie.....	61
7.4 Ongestruceerde categorie.....	62
7.5 Analyse .....	63
7.6 Conclusie.....	63
<b>8. Doelmatig- &amp; Doeltreffendheid .....</b>	<b>65</b>
8.1 Gestructureerde categorie.....	65
8.2 Slecht gestructureerde categorie.....	66
8.3 Matig gestructureerde categorie.....	67
8.4 Ongestruceerde categorie.....	68

8.5 Analyse .....	69
8.6 Conclusie .....	70
<b>9. Conclusie .....</b>	<b>72</b>
<b>10. Discussie .....</b>	<b>74</b>
<b>11. Aanbevelingen .....</b>	<b>75</b>
11.1 Aanbevelingen voor externe, commerciële adviesbureaus.....	75
11.2 Aanbevelingen voor opdrachtgevers van beleidsevaluaties.....	75
11.3 Aanbevelingen voor toekomstig wetenschappelijk onderzoek.....	75
<b>12. Literatuurlijst .....</b>	<b>76</b>
<b>13. Geraadpleegde Casussen.....</b>	<b>82</b>
<b>Bijlage 1: Methodologie .....</b>	<b>84</b>
<b>Bijlage 2: De Inrichting .....</b>	<b>86</b>
<b>Bijlage 3: Het Evaluatieproces .....</b>	<b>89</b>
<b>Bijlage 4: De Instrumenten.....</b>	<b>94</b>
<b>Bijlage 5: Doelmatig- &amp; Doeltreffendheid.....</b>	<b>101</b>



# 1. Inleiding

## 1.1 Aanleiding

De huidige samenleving heeft te maken met grote problemen, zoals klimaatverandering en ondermijnende criminaliteit (Korsten, 2016b). Deze problemen hebben veel invloed op de maatschappij, onder andere door de hoge mate van onzekerheid en complexiteit.

Tegelijkertijd kiezen mensen er vaak voor om complexiteit te vermijden (Van Berkel & Manickam, 2019). Als we er echt niet omheen kunnen, delen we lastige vraagstukken graag op in kleine behapbare stukjes en doen we ons best om eenduidige oorzaken aan te wijzen die we makkelijk kunnen wegnemen of oplossen. Echter, deze drang naar versimpeling werkt niet bij de vraagstukken waar we momenteel als maatschappij voor aan de lat staan.

Problemen bevinden zich op een schaal van wickedness: zij zijn niet per definitie simpel of complex, maar in meer of mindere mate wicked (Stolk et al., 2021). Rittel en Webber (1973) hebben een aantal karakteristieken vastgelegd waaraan wickedness kan worden herkend. Aan de hand van deze karakteristieken kunnen drie overkoepelende dimensies van wickedness worden geïdentificeerd: kennis, consensus, en betrokken actoren (Korsten, 2019).

Aan de ene kant van de schaal van wickedness staan problemen met een lage mate van wickedness: er is veel kennis beschikbaar over het probleem en het is eenvoudig om met de betrokken actoren tot een overeenstemming te komen (Korsten, 2019). Aan de andere kant staan problemen met een hoge mate van wickedness: over deze problemen is weinig tot geen kennis beschikbaar en het is moeilijk tot onmogelijk om tot een overeenstemming te komen, mede doordat het niet bekend is wie de betrokken actoren precies zijn. Tussen deze twee uitersten staan nog twee mengvormen: er is wel kennis beschikbaar, maar geen overeenstemming tussen partijen of er is geen kennis beschikbaar, maar het is wel mogelijk om met de betrokken actoren tot een overeenstemming te komen.

Over alle problemen op de schaal maakt de overheid beleid, met als doel het probleem aan te pakken. Dit beleid komt tot stand middels het beleidsproces (Bekkers, 2017). Als onderdeel van dit proces wordt beleid geëvalueerd. De evaluaties vinden met name achteraf (*ex post*) plaats, maar kunnen ook vooraf (*ex ante*) of gedurende het beleidsproces (*ex durante*) plaatsvinden (Terry, 2016). Daarbij kunnen evaluaties verschillende doelen nastreven: het ondersteunen van de beleidsplanning, het verbeteren van de beleidsuitvoering en bijdragen aan institutionele ontwikkeling van de betrokken organisaties, het afleggen van verantwoording, en beleidsleren (De Peuter et al., 2007; Terry, 2016).

Er zijn verschillende partijen die een beleidsevaluatie kunnen uitvoeren. Deze partijen kunnen intern zijn, bijvoorbeeld via een interdepartementaal beleidsonderzoek (Ministerie van Financiën, z.d.), maar ook extern, door partijen buiten de eigen organisatie (De Peuter et al., 2007). Voor veel beleidsevaluaties wordt een externe partij ingeschakeld (Haarhuis & Parapuf, 2016). Dit kunnen planbureaus zijn, zoals het Sociaal Cultureel Planbureau of het Planbureau voor de Leefomgeving, maar ook de Sociaal-Economische Raad, het WODC, of de Algemene Rekenkamer (Bovens et al., 2017). Daarnaast kunnen commerciële adviesbureaus worden ingezet als externe partij (Houppermans, 2018). Voorbeelden van enkele grotere bureaus in Nederland die veel beleidsevaluaties uitvoeren zijn Andersson Elffers Felix, Berenschot, Deloitte, KWINK Groep, en PWC.

Zoals gezegd staat de samenleving voor een groot aantal problemen met verschillende mate van wickedness, bepaald door de beschikbare kennis en consensus, en de betrokken actoren. De overheid probeert problemen in alle gradaties van wickedness aan te pakken door middel van beleid. Dit beleid tot komt stand via het beleidsproces. Dit proces behelst ook een evaluatie van het beleid, vaak uitgevoerd door een externe, commerciële partij. Echter, het is onduidelijk in welke mate de gradatie van wickedness een rol speelt in de evaluatie door dergelijke externe partijen. Wegen zij de mate van wickedness mee in evaluaties, en zo ja, hoe dan? Deze vragen staan in dit onderzoek centraal.

## 1.2 Probleemstelling

Op basis van het voorgaande formuleer ik de hoofdvraag van dit onderzoek als volgt: *In welke mate speelt de gradatie van wickedness van problemen een rol bij beleidsevaluaties uitgevoerd door externe, commerciële adviesbureaus?*

Om de hoofdvraag te beantwoorden, is dit onderzoek opgedeeld in een theoretisch en een empirisch deel.

Het theoretische deel wordt gevormd door middel van een literatuurstudie en zet het kader uiteen met betrekking tot wickedness, beleidsevaluaties, en externe, commerciële adviesbureaus. Dit kader vormt de basis van dit onderzoek.

In het empirische deel wordt, via een documentenonderzoek, antwoord gegeven op de volgende vijf deelvragen:

1. Hoe kunnen twaalf beleidsevaluaties uitgevoerd door vijf externe, commerciële adviesbureaus worden ingedeeld op een schaal van wickedness?
2. In welke mate speelt de gradatie van wickedness een rol bij de inrichting van de rapportage van de geselecteerde beleidsevaluaties?
3. In welke mate speelt de gradatie van wickedness een rol bij het evaluatieproces gehanteerd in de geselecteerde beleidsevaluaties?
4. In welke mate speelt de gradatie van wickedness een rol bij de instrumenten gehanteerd in de geselecteerde beleidsevaluaties?
5. In welke mate speelt de gradatie van wickedness een rol bij de mate waarin de geselecteerde beleidsevaluaties een uitspraak doen over de doelmatig- en doeltreffendheid van het beleid?

### 1.3 Doelstelling

Het doel van dit onderzoek is om inzicht te krijgen in welke mate de gradatie van wickedness een rol speelt bij beleidsevaluaties uitgevoerd door externe, commerciële adviesbureaus. Aan de hand van wetenschappelijke literatuur en empirische inzichten worden aanbevelingen gedaan betreffende het proces en de instrumenten van beleidsevaluaties, alsmede over de inrichting van de rapportage over de evaluatie. Op basis van de inzichten en aanbevelingen uit dit onderzoek kunnen externe, commerciële adviesbureaus, die jaarlijks een groot aantal beleidsevaluaties uitvoeren, hun beleidsevaluaties beter aansluiten op de gradatie van wickedness en een grotere bijdrage leveren het beleidsproces. Voor de maatschappij betekent dit mogelijk een slag naar een betere aanpak van maatschappelijke problemen met een hoge mate van wickedness, betere sturing, en een meer effectieve en efficiënte inzet van publieke middelen.

### 1.4 Wetenschappelijke relevantie

Over het algemeen streven overheden naar beleid dat effectief, betrouwbaar, en geloofwaardig is (Knottnerus et al., 2016). Om dit te beoordelen wordt beleid geëvalueerd (Korsten, 2013) door onder andere externe, commerciële adviesbureaus (Haarhuis & Parapuf, 2016). Dit onderzoek beoogt een bijdrage te leveren door inzicht te bieden in de mate waarin wickedness een rol speelt bij beleidsevaluaties uitgevoerd door externe, commerciële adviesbureaus. Deze bijdrage is gestoeld op het toetsen van wetenschappelijke kennis over wicked problems, alsmede de invloed die wickedness heeft op de manier waarop beleidsevaluaties door externe, commerciële adviesbureaus worden uitgevoerd en de rapportage hierover wordt ingericht.

## 1.5 Bestuurlijke relevantie

Beleidsbeoordelingen uitgevoerd door externe, commerciële adviesbureaus worden dikwijls aangeboden aan de Tweede Kamer ter verantwoording van beleid of als basis voor vervolgbestuur (Houppermans, 2018). Met bestuur wordt geprobeerd problemen aan te pakken die - op basis van de beschikbaarheid van kennis en consensus - een bepaalde mate van wickedness kennen (Korsten, 2019). Het is de vraag in hoeverre de mate wickedness een rol speelt bij het evalueren van bestuur. Aangezien een groot deel van de beleidsbeoordelingen wordt uitgevoerd door externe, commerciële adviesbureaus (Houppermans, 2018), focust dit onderzoek specifiek op beleidsbeoordelingen uitgevoerd door deze partijen.

## 1.6 Maatschappelijke relevantie

De samenleving zal steeds vaker te maken krijgen met problemen met een hoge mate van wickedness (Peters, 2017). Om de aanpak van deze problemen te evalueren is het van belang dat er rekening wordt gehouden met de gradatie van wickedness van het probleem (Duckett et al., 2016). Voor het evalueren van bestuur wordt met publieke middelen vaak een extern, commercieel adviesbureau ingeschakeld (Bouterse & Pattyn, 2018). Een vereiste bij het gebruik van publieke middelen is dat dit efficiënt en effectief gebeurt. Voor beleidsbeoordelingen betekent dit dat de evaluatie bruikbare, valide, en betrouwbare inzichten geeft om de kwaliteit van het bestuur te verbeteren (Haarhuis & Parapuf, 2016). Dit onderzoek beoogt inzicht te bieden in de mate waarin de gradatie van wickedness een rol speelt bij beleidsbeoordelingen uitgevoerd door externe, commerciële adviesbureaus.

## 1.7 Leeswijzer

Allereerst zal ik het theoretisch kader van dit onderzoek uiteenzetten. Waarna ik in hoofdstuk 3 de gehanteerde methoden bespreek. In hoofdstuk 4 tot en met 8 beantwoord ik de vijf deelvragen. Ten eerste deel ik de geselecteerde beleidsbeoordelingen in op een schaal van wickedness. Deze categorisering is de basis voor de analyse van de inrichting van de rapportages, het gehanteerde proces, de gehanteerde instrumenten, en de mate waarin een uitspraak wordt gedaan over de doelmatig- en doeltreffendheid. Ik sluit het rapport af met een conclusie en een aantal aanbevelingen voor externe, commerciële adviesbureaus, opdrachtgevers van beleidsbeoordelingen, en toekomstig wetenschappelijk onderzoek.

## 2. Theoretisch Kader

### 2.1 Wickedness

#### 2.1.1 Achtergrond van wickedness

Dit onderzoek deelt beleidsevaluaties in op een schaal van wickedness, gebaseerd op de term ‘wicked problem’. Deze term werd in 1973 uiteengezet door Rittel en Webber als kritiek op de dominante rationeel-technische aanpak van problemen. Deze aanpak focust op het efficiënt en effectief behalen van doelen op basis van adequate informatie en zorgvuldig geformuleerde doelstellingen en methoden (Head, 2008). Volgens Rittel en Webber (1973) is deze aanpak geschikt voor problemen met een lage mate van wickedness. Denk hierbij aan de uitdagingen ten tijde van de industriële revolutie: het bestraten van steden en het aanleggen van sanitaire voorzieningen, bijvoorbeeld. Deze aanpak is minder geschikt voor problemen met een hoge mate van wickedness. De kritiek bestaat uit vier punten:

Het eerste kritiekpunt betreft het systeem waarbinnen maatschappelijke problemen worden begrepen (Head & Alford, 2015). Maatschappelijke problemen die wicked zijn, kunnen niet worden begrepen en aangepakt in isolatie, maar zijn onderdeel van een systeem waarbinnen zij gekoppeld zijn aan en afhankelijk zijn van andere problemen. Dit systeem maakt het moeilijk om problemen te definiëren en om de oorzaak, juiste aanpak, en mogelijke gevolgen te identificeren (Rittel & Webber, 1973).

Het tweede kritiekpunt focust op de onmogelijkheid alle benodigde informatie te verkrijgen, doelen duidelijk vast te leggen, en het benodigde niveau van coördinatie te behalen bij problemen met een hoge mate van wickedness (Pressman & Wildavsky, 1973).

Het derde kritiekpunt focust op het gebrek aan aandacht vanuit de rationeel-technische aanpak voor de waarden, perspectieven, en ervaringen van betrokken actoren (Head & Alford, 2015). Dit kritiekpunt is gekoppeld aan het vierde kritiekpunt: de hedendaagse maatschappij is te pluralistisch voor de rationeel-technische aanpak. De subgroepen binnen de samenleving vertonen in steeds grotere mate verschillen in waarden en perspectieven waardoor het lastiger wordt om deze te verenigen en tot overeenstemming te komen over de definitie en aanpak van maatschappelijke problemen. Daarbij proberen alle subgroepen via de politiek hun eigen doelen na te streven, maar wat bevredigend is voor de één, kan verwerpelijk zijn voor de ander, en wat een oplossing is voor de een, kan een probleem zijn voor de ander (Rittel & Webber, 1973). Het is onmogelijk om te zeggen welke groep gelijk heeft en het verdient om zijn doel te streven.

### 2.1.2 Karakteristieken van wickedness

Rittel en Webber (1973) hebben tien karakteristieken geïdentificeerd waaraan wickedness valt te herkennen.

De eerste karakteristiek focust op het gebrek aan definitie van een beleidsprobleem (Rittel & Webber, 1973). Dit komt, ten eerste, doordat de betrokken actoren het niet eens kunnen worden over de definitie. Daarbij is de informatie die nodig is om een definitie vast te stellen afhankelijk van de mogelijke oplossing (Ritchey, 2013). Om een wicked probleem te definiëren, is een inventarisatie nodig van alle mogelijke oplossingen.

Echter, problemen met een hoge mate van wickedness hebben geen perfecte oplossing en zijn nooit definitief opgelost (Rittel & Webber, 1973). Dit komt onder andere doordat er geen objectieve criteria zijn om te bepalen wanneer een probleem is opgelost, alsmede dat het probleem voortdurend kan evalueren. Meestal wordt er gewerkt aan een oplossing die goed genoeg wordt gevonden totdat de middelen op zijn (Ritchey, 2013). Ook zijn oplossingen niet waar of onwaar, maar goed of slecht volgens de betrokken actoren. Zij beoordelen mogelijke oplossingen aan de hand van hun eigen interesses, waarden, en ideologische voorkeuren. Daarbij zijn er geen criteria om aan te toetsen of alle mogelijke oplossingen zijn verkend.

Bovendien is het testen van een potentiële oplossing niet mogelijk (Rittel & Webber, 1973). Elke oplossing die wordt toegepast, kan in de praktijk onvoorziene consequenties hebben of leiden tot nieuwe problemen (Ritchey, 2013). Daarbij is er slechts één poging om een mogelijke oplossing in de praktijk toe te passen (Rittel & Webber, 1973). Consequenties van de implementatie kunnen niet gemakkelijk worden teruggedraaid en er is geen mogelijkheid om te leren door middel van *trial-and-error*.

Hier komt nog bij dat de maatschappij geen tolerantie heeft voor oplossingen die falen (Rittel & Webber, 1973). Onderzoekers krijgen geen ruimte om hypothesen te toetsen die achteraf onjuist blijken te zijn (Ritchey, 2013). Het doel is dan ook niet om de waarheid te vinden, maar om enkele aspecten van het probleem succesvol aan te pakken.

De achtste karakteristiek van wicked problemen is dat elk probleem uniek is (Rittel & Webber, 1973). Er kunnen geen oplossingen worden ontwikkeld die van toepassing zijn op een overkoepelende groep problemen (Ritchey, 2013).

Niet alleen is elk wicked probleem uniek, de meeste wicked problemen zijn ontstaan als gevolg van een ander probleem (Rittel & Webber, 1973). Ook binnen een probleem zijn symptomen vaak afhankelijk van andere interne symptomen: er is sprake van veel wederzijdse en circulaire causaliteit (Ritchey, 2013).

Tot slot kunnen wicked problemen op verschillende manieren worden uitgelegd en is er geen manier om vast te stellen welke verklaring correct is (Rittel & Webber, 1973). Mensen zullen de verklaring kiezen die het best past bij hun eigen waarden, perspectieven, en doelen. De manier waarop een probleem wordt uitgelegd is belangrijk omdat het bepaalt op welke manier een probleem wordt aangepakt (Ritchey, 2013).

### 2.1.3 Schaal van wickedness

Dit onderzoek gaat uit van een schaal van wickedness: problemen worden niet geclassificeerd als wel of niet wicked, maar er wordt uitgegaan van gradaties van wickedness. Deze aanpak is in lijn met de typologie van Korsten (2016a). Hij maakt aan de hand van drie elementen: *kennis*, *consensus*, en *betrokken actoren*, een onderscheid tussen gestructureerde, slecht gestructureerde, matig gestructureerde, en ongestructureerde problemen.

De problemen met de laagste mate van wickedness zijn *gestructureerde* problemen (Korsten, 2016a). Over dit type probleem is veel kennis beschikbaar en het is makkelijk om met de betrokken actoren tot een overeenstemming te komen over wat het probleem is, welk doel wordt nagestreefd, en welke kennis belangrijk is (Korsten, 2016a; Stolk et al., 2021). Deze problemen zijn doorgaans goed gedefinieerd en kunnen daardoor worden aangepakt met standaardmethoden en -technieken. De verantwoordelijkheid voor het aanpakken van het probleem ligt bij één actor. Er kunnen meerdere actoren betrokken zijn, maar deze hebben dan vergelijkbare doelen en delen gemakkelijk kennis met elkaar (Alford & Head, 2017).

Ten tweede zijn er *slecht gestructureerde* problemen (Korsten, 2016a). Over dit type is eveneens veel kennis beschikbaar. Echter, er geen consensus over welke aanpak juist is en welke doelen worden nagestreefd (Stolk et al., 2021). Anders gezegd: een probleem wordt beschouwd als slecht gestructureerd wanneer er genoeg kennis voorhanden is over het probleem, maar waarden van betrokken actoren met elkaar botsen.

Het derde type probleem is *matig gestructureerd* (Korsten, 2016a). Dit type omvat problemen waarover de kans op consensus groot is: de betrokken actoren zijn het eens over wat het probleem is en welke doelen worden nagestreefd (Korsten, 2016a). Echter, de kennis die beschikbaar is, is beperkt.

Het vierde en laatste type probleem heeft een hoge mate van wickedness en is *ongestructureerd* (Korsten, 2016a). Bij dit type is er zowel geen kennis beschikbaar als dat er geen consensus is tussen de betrokken actoren, waardoor complexiteit en onzekerheid overheersen. Bij dit type staat het organiseren van ontmoetingen tussen de betrokken actoren centraal (In 't Veld, 2010). Het geleidelijk concretiseren en door middel van overleg tot een

consensus komen, is hierbij het best haalbare. Door nieuwe kennis en een verschuiving van waarden en voorkeuren bij actoren kan er na verloop van tijd een benadering ontstaan, waarvan het onduidelijk is hoe lang deze meegaat. De verantwoordelijkheid voor het oplossen van het probleem is versnipperd en de betrokken actoren moeten hun krachten bundelen om het probleem aan te pakken.

Tot slot: bovenstaande categorisering biedt op zichzelf geen oplossing voor de problemen (Head, 2008). Het is slechts een middel om problemen, en de context waarin deze problemen zich bevinden, beter te begrijpen. In dit rapport wordt deze schaal gebruikt om te onderzoeken of, en in welke mate, de gradatie van wickedness een rol speelt bij het evalueren van beleid.

### **Figuur 1**

*Schaal van wickedness*

Gestructureerd	Slecht gestructureerd	Matig gestructureerd	Ongestructureerd
Kennis beschikbaar over het probleem en de oplossing	Kennis beschikbaar over het probleem en de oplossing	Geen kennis beschikbaar over het probleem en de oplossing	Geen kennis beschikbaar over het probleem en de oplossing
Consensus over de kennis, het probleem, en het doel	Geen consensus over de kennis, het probleem en het doel	Consensus over de kennis, het probleem, en het doel	Geen consensus over de kennis, het probleem, en het doel
Betrokken actoren bekend	Betrokken actoren (on)bekend	Betrokken actoren (on)bekend	Betrokken actoren onbekend

## **2.2 Beleid**

Om problemen uit alle gradaties van wickedness aan te pakken, wordt beleid ontwikkeld. Bekkers (2017) definieert beleid als een afweging tussen verschillende waarden die de samenleving belangrijk vindt. In beleid wordt vastgesteld welke waarden moeten worden gerealiseerd en welke maatregelen en instrumenten worden ingezet om dit te verwezenlijken. De term ‘beleid’ refereert dus aan het realiseren van bepaalde doelstellingen aan de hand van bepaalde middelen in een bepaalde tijdsvolgorde (Hoogerwerf, 1987).



### 2.2.1 Perspectieven op beleid

Er zijn vier perspectieven op beleid: het rationele, politieke, culturele, en institutionele perspectief (Bekkers, 2017). Het rationele en politieke perspectief zijn twee meer klassieke benaderingen, terwijl de culturele en institutionele perspectieven recenter zijn en de afgelopen twintig jaar zijn opgekomen.

Het *rationele perspectief* bekijkt problemen aan de hand van feiten en cijfers, en kiest oplossingen aan de hand van een analyse van welke optie de beste uitkomst genereert tegen de laagste kosten (Simon, 1976; Bekkers, 2017). Beleid wordt gezien als een resultaat van keuzes die bewust zijn gemaakt aan de hand van een rationele afweging tussen de voor- en nadelen van diverse mogelijkheden. Het uitgangspunt van dit perspectief is om doelgericht en weloverwogen te handelen om de effectiviteit, efficiëntie, en legitimiteit van beleid te verhogen. Volgens het rationele perspectief is de maatschappij enkel het object waar de overheid op stuurt en heeft dan ook geen invloed op het proces van beleidsvorming.

Het *politieke perspectief* beschouwt beleid als een proces van macht en verdeling (Bekkers, 2017). Bij dit perspectief gaat het niet zozeer om de vraag of beleid goed doordacht is, maar of er voldoende steun is. Niet zozeer het maken van een rationele afweging, maar het sluiten van compromissen staat centraal. Volgens het politieke perspectief komt beleid tot stand door middel van strategische acties van betrokken actoren, in pogingen steun voor het beleid te verwerven, blokkeren, of ondermijnen.

Het derde perspectief, het *culturele perspectief*, ziet beleid als een manier van betekenisverlening (Bekkers, 2017). Dat wil zeggen dat actoren zelf betekenis geven aan de definitie en oplossingen van problemen door middel van interacties. Hierbij spelen taal en beelden een belangrijke rol. Beleid wordt dan ook gezien als een sociale constructie dat het product is van doorlopende interacties tussen actoren (Termeer, 1993; Stone, 2003). Met andere woorden, beleid komt tot stand via discourscoalities: een samenkomst van actoren die gezamenlijk een interpretatie van de werkelijkheid ontwikkelen, structureren, en in stand houden (Hajer, 1989).

Het laatste perspectief, het *institutionele perspectief*, stelt net als het culturele perspectief betekenisverlening centraal (Bekkers, 2017). Echter, in het institutionele perspectief overheerst een deterministische invulling: zingeving vindt plaats door gegroeide praktijken die zijn vastgelegd in instituties, regels, en procedures. Hierdoor hebben individuele actoren maar beperkte mogelijkheden om zelf invulling te geven aan beleid (Bekkers, 2017) en ontstaat beleid via een afgebakend pad dat vorm heeft gekregen door deze gegroeide praktijken, ofwel pad afhankelijkheid (Pierson, 2000).

### 2.2.2 Perspectieven op het beleidsproces

Beleid komt niet zomaar tot stand, maar is het gevolg van het beleidsproces (Bekkers, 2017). Dit proces kan worden onderverdeeld in zes fasen: agendavorming, beleidsvorming, beleidsbepaling, beleidsuitvoering, beleidsevaluatie, en terugkoppeling (Van den Berg et al., 2021). Agendavorming gaat over welke onderwerpen belangrijk genoeg worden geacht door beleidsmakers om beleid over te maken (Bovens et al., 2017). Nadat een onderwerp op de agenda is gekomen, wordt beleid gevormd via twee stappen: beleidsvoorbereiding en beleidsbepaling. Tijdens de beleidsvoorbereiding stellen beleidsmakers beleidsvoorstellen op waarin het probleem wordt onderzocht, doelstellingen worden vastgelegd, en mogelijke oplossingen worden verkend. Deze beleidsvoorstellen moeten vervolgens door de politiek worden aangenomen: de beleidsbepaling. Nadat beleid is aangenomen worden de beleidsvoorstellen omgezet naar daadwerkelijke handelingen om de doelstellingen te behalen. Om te bepalen of dit gebeurt, wordt beleid geëvalueerd. De resultaten van deze evaluatie worden in de laatste fase van het beleidsproces teruggekoppeld aan de beleidsmakers. Aan de hand van deze terugkoppeling, wordt het proces opnieuw doorlopen, tenzij wordt besloten het beleid te stoppen.

Het beleidsproces kan vanuit twee perspectieven worden beschouwd: het analytische en politieke perspectief (Van den Berg et al., 2021).

Volgens het *analytische perspectief* hangen de fasen uit het beleidsproces op een logische manier met elkaar samen en volgen ze elkaar doorlopend op, tot wordt besloten het beleid stop te zetten (Van den Berg et al., 2021). Dit proces schept orde in de chaos rond het beleidsprobleem.

Volgens het *politieke perspectief* is er geen duidelijk onderscheid tussen de fasen van het beleidsproces (Van den Berg et al., 2021). Waar het analytische perspectief een oplossing zoekt voor het probleem, kan volgens het politieke perspectief net zo goed een probleem worden gezocht voor een oplossing. Het is, volgens het politieke perspectief, wel mogelijk om fasen te identificeren in het beleidsproces, maar deze fasen zijn het gevolg van bewuste politiek strategische acties van actoren ten behoeve van de eigen belangen.

Deze twee perspectieven komen ook als mix voor in het barrièremodel en in het stromenmodel (Van den Berg et al., 2021).

Het *barrièremodel* komt voort uit het analytische perspectief en gaat uit van rationaliteit en een logische opeenvolging van de fasen van het beleidsproces (Van den Berg et al., 2021). Tegelijkertijd gaat dit model uit van de macht van betrokken actoren. Aan de ene kant kan een actor *realisatiemacht* hebben: het vermogen om de eigen belangen te verwezenlijken via het

beleidsproces. Aan de andere kant staan actoren met *hindermacht*: het vermogen om barrières op te werpen en realisatiemacht te verhinderen. Deze barrières kunnen in elke fase en in verschillende vormen voorkomen. Volgens het barrièremodel is beleidsvorming een continue strijd tussen actoren met realisatiemacht en actoren met hindermacht.

Het *stromenmodel* neemt afstand van de fasering van het beleidsproces en gaat ervan uit dat beleid tot stand komt wanneer drie stromen - problemen, partijen, en oplossingen – samenkomen (Van den Berg et al., 2021). Dit kan per toeval gebeuren, maar een samenkomst kan ook worden geforceerd door policy entrepreneurs die stromen actief bij elkaar brengen.

## 2.3 Beleidsevaluaties

Dit onderzoek focust op de evaluatiefase waarin beleid wordt beoordeeld met als doel beleidsmakers, bestuurders, en burgers te informeren over de doeltreffendheid en doelmatigheid van het gevoerde beleid (Van der Knaap et al., 2020).

### 2.3.1 Perspectieven op beleidsevaluaties

Het analytische, politieke, culturele, en institutionele perspectief hebben niet alleen een eigen visie op beleid, maar ook op beleidsevaluaties (Bekkers, 2017).

Volgens het *analytische perspectief* focussen beleidsevaluaties op de doelmatig- en doeltreffendheid van beleid (Bekkers, 2017). Hierbij worden twee evaluatievragen gesteld. Ten eerste “in hoeverre komen de feitelijke resultaten overeen met de beoogde resultaten?” (Bekkers, p.323, 2017). Ten tweede “wat is de bijdrage van een bepaalde maatregel [...] aan de resultaten die feitelijk zijn geboekt?” (Bekkers, p.323, 2017). Om deze vragen te kunnen beantwoorden dient het beleid heldere, meetbare doelen te stellen. Volgens dit perspectief dient het proces van evalueren vooral als leerproces waarbij beleidsmakers bereid zijn beleid aan te passen als daar een feitelijk onderbouwde reden toe is (Bovens et al., 2017).

Bij het *politieke perspectief* draait het evalueren van beleid om het vaststellen in welke mate de eigen positie is verbeterd of verslechterd (Bekkers, 2017). Succesvol beleid heeft draagvlak doordat het rekening houdt met de belangen, wensen, en voorkeuren van betrokken actoren. Volgens het politieke perspectief zijn beleidsmakers zich bewust dat draagvlak precair is en zullen zich daarom inzetten om elementen van beleid die kunnen worden geëvalueerd te beïnvloeden (Bovens et al., 2017). De evaluatie is simpelweg een volgend deel van het politieke spel en de politieke strijd.

Het derde perspectief, het *culturele perspectief*, gaat ervan uit dat beleid succesvol is geweest wanneer het een gemeenschappelijk beeld heeft gecreëerd over het nut en noodzaak

van het beleid via een collectief leerproces (Bekkers, 2017). Kennis die wordt ingebracht in het leerproces heeft voornamelijk een attenderingsfunctie: informatie vormt een signaal waarmee wordt aangegeven wat wel en niet belangrijk is (Feldman & March, 1981).

Tot slot, het *institutionele perspectief* is een combinatie van bovenstaande perspectieven (Bekkers, 2017). Evaluaties stellen niet alleen vast of beleid effectief en efficiënt is geweest, maar kijken ook naar de technische, politieke, en juridische haalbaarheid. Daarbij fungeren beleidsevaluaties als toets van het leervermogen van de beleidsmakers: in hoeverre zijn beleidsmakers in staat pad afhankelijkheid te doorbreken? Omdat deze waarden op gespannen voet met elkaar kunnen staan, dienen evaluaties met enige terughoudendheid conclusies te trekken.

### 2.3.2 Inhoud van beleidsevaluaties

Buiten bovenstaande perspectieven, kunnen beleidsevaluaties ook inhoudelijk van elkaar verschillen. De inhoud heeft betrekking op het evaluatiemotief, de gebruiksvorm, het evaluandum, en het moment in de beleidscyclus (De Peuter et al., 2007; Terryn, 2016).

Het evaluatiemotief heeft betrekking op het doel dat wordt nagestreefd met de evaluatie (De Peuter et al., 2007; Terryn, 2016). Beleidsevaluaties streven over het algemeen één van deze vier doelen na: het ondersteunen van de beleidsplanning, het verbeteren van de beleidsuitvoering en bijdragen aan institutionele ontwikkeling van de organisatie, het afleggen van verantwoording, of beleidsleren.

Het doel dat een evaluatie nastreeft is van belang omdat het in grote mate bepaalt op welke manier een evaluatie wordt gebruikt (De Peuter et al., 2007; Terryn, 2016). Er kunnen vier gebruiksvormen worden onderscheiden: instrumenteel, conceptueel, proces gerelateerd, en symbolisch gebruik (Balthasar, 2006). Bij instrumenteel gebruik verschaffen evaluaties inzichten die direct kunnen worden toegepast om het beleidsproces te informeren (De Peuter et al., 2007). Bij conceptueel gebruik beïnvloeden de evaluatieresultaten de overtuigingen van de beleidsmakers met betrekking tot het voorwerp, ofwel evaluandum, van de evaluatie. Proces gerelateerde evaluaties creëren voordelen voor de betrokken stakeholders door, bijvoorbeeld, gemeenschappelijk begrip te faciliteren. Tot slot kunnen evaluaties symbolisch worden gebruikt: ter legitimatie van een beleidsinitiatief.

Het evaluandum gaat over de vraag: wat wordt geëvalueerd? (De Peuter et al., 2007). Hierbij zijn drie mogelijkheden: de beleidsinhoud, het beleidsproces, en de beleidsimpact (Hoogerwerf, 2003; De Peuter et al., 2007; Terryn, 2016). Bij het evalueren van de beleidsinhoud worden het concept, de structuur, en de instrumenten bekeken (Bressers &

Hoogerwerf, 1995). Bij een evaluatie van het beleidsproces wordt gekeken naar het proces en de acties van het beleid. Wanneer de impact van beleid onderwerp is van een evaluatie wordt gekeken naar de veranderingen die het beleid teweeg heeft gebracht in de maatschappij.

Tot slot kan het moment van de evaluatie variëren. Een beleidsevaluatie kan plaatsvinden op drie momenten: voorafgaand aan het beleidsproces (*ex ante*), gedurende het beleidsproces (*ex durante*), en na het uitvoeren van het beleid (*ex post*) (Terry, 2016). *Ex ante* evaluaties focussen op het ondersteunen van de beleidsvoorbereiding, beoordelen beleidsdoelen en de beleidstheorie, en toetsen alternatieve invoeringsstrategieën (De Peuter et al., 2007). *Ex durante* beleidsevaluaties focussen op de beleidsuitvoerende fase en evalueren het beleidsproces op basis van de gestelde doelen (De Peuter et al., 2007; Terry, 2016). *Ex post* evaluaties vinden plaats na de beleidsuitvoering en evalueren de effecten, effectiviteit, en/of efficiëntie van het gevoerde beleid (De Roo, 2002) en leveren input aan de hand waarvan wordt bepaald of het beleid, al dan niet met wijzigingen, wordt doorgezet of stopgezet (Terry, 2016).

### 2.3.3 Het evaluatieproces

Ondanks dat beleidsevaluaties inhoudelijk van elkaar kunnen verschillen, doorlopen zij over het algemeen hetzelfde evaluatieproces bestaande uit de structurerings-, dataverzamelings-, data-analyse-, en beoordelingsfase (De Peuter et al., 2007).

In de eerste fase, de structurerings-, staat het ontwerpen van het evaluatiedesign centraal (De Peuter et al., 2007). Dit omvat twee kerntaken: het afbakenen van de inhoudelijke scope - wat is het evaluandum - en het plannen van de vervolgstappen van het evaluatieproces - welke bronnen en technieken zullen worden ingezet tijdens de evaluatie.

De vastgestelde technieken zullen worden ingezet ten behoeve van fase twee, het verzamelen van de data die nodig is om het evaluandum te analyseren en beoordelen (De Peuter et al., 2007). Afhankelijk van het evaluandum kan gebruik worden gemaakt van bestaande of nieuw verkregen data. Vervolgens wordt deze data geanalyseerd aan de hand van de evaluatievragen en -criteria.

Tot slot wordt een waardeoordeel geveld over het evaluandum (De Peuter et al., 2007). Dit oordeel komt tot stand door de resultaten van de analyse af te zetten tegen een bepaald referentiekader: normen en waarden die vanuit de politiek of beleid worden aangereikt.

#### 2.3.4 Instrumenten voor evaluatieonderzoek

Om data te verzamelen kunnen externe evaluatoren gebruik maken van verschillende instrumenten: documentenanalyses, interviews, focusgroepen, expertpanels, en workshops (Terry, 2016).

Bij een documentenanalyse wordt onderzoek gedaan met behulp van bestaande data (Verhoeven, 2011) zoals jaarverslagen, rapporten van organisaties, beleidsnota's, en jaarplannen (Van Thiel, 2021).

Bij interviews wordt door middel van een gesprek tussen de evaluator en één of meerdere respondenten informatie verzameld over het evaluatievraagstuk (Van Thiel, 2021). Interviews kunnen verschillende vormen aannemen: open, zonder specifieke richting anders dan het evaluatieonderwerp, of semigestructureerd, waarbij een topiclijst sturing geeft aan het interview.

Het derde instrument is de focusgroep. Bij dit instrument wordt een klein aantal deelnemers geselecteerd om over bepaalde onderwerpen te discussiëren (Terry, 2016). Tijdens de focusgroep verzamelt de evaluator informatie over het evaluandum en over de waarden van de betrokken actoren.

Bij het vierde instrument, het expertpanel, worden onafhankelijke deskundigen ingezet om input te leveren over de functie, relevantie, en kwaliteit van het beleid (Terry, 2016).

Tot slot kunnen externe evaluatoren, volgens Terry (2016), gebruik maken van workshops.

#### 2.3.5 Rapporteren over beleidsevaluaties

Zoals gezegd zijn beleidsevaluaties bedoeld om beleidsmakers, bestuurders, en burgers te informeren (Van der Knaap et al., 2020). Dit gebeurt veelal via een rapport (De Peuter et al., 2007). Dit rapport dient een heldere structuur te hebben, maar er bestaat niet één ideale indeling. Het evaluandum, de context van het evaluandum, en de verwachtingen van de betrokken actoren hebben hier invloed op. Desondanks kunnen enkele basis onderdelen worden onderscheiden die terug dienen te komen in alle rapportages van beleidsevaluaties.

Ten eerste is het gebruik van een managementsamenvatting essentieel (De Peuter et al., 2007). Dit onderdeel wordt geschreven om beleidsmakers en politici te informeren over de resultaten, conclusies, en aanbevelingen van het onderzoek.

Het tweede onderdeel is de inleiding (De Peuter et al., 2007). In dit hoofdstuk worden de aanleiding, betrokken actoren, en afkadering van het onderzoek besproken. De inleiding dient af te sluiten met een leeswijzer.

Na de inleiding volgt het evaluatiedesign (De Peuter et al., 2007). In dit hoofdstuk wordt het evaluatieonderwerp beschreven, wat de basis vormt voor het evaluatieonderzoek en de verdere inrichting van het rapport. Daarbij worden de vraagstelling en evaluatiecriteria besproken. Tot slot lichten de evaluatoren de uitvoering en eventuele knelpunten van het onderzoek toe.

In het vierde onderdeel worden de resultaten van de analyse gepresenteerd (De Peuter et al., 2007). Dit onderdeel fungeert tevens als basis voor de twee laatste onderdelen van de rapportage: de conclusies en aanbevelingen.

In de conclusie geven de evaluatoren een beoordeling en samenvatting van de resultaten (De Peuter et al., 2007). Daarbij dienen de onderzoeksvragen te worden beantwoord.

Na de conclusie kunnen aanbevelingen worden gedaan (De Peuter et al., 2007). Echter, er is in de literatuur geen consensus over de noodzaak van aanbevelingen. Het bepalen van wat er moet gebeuren na de evaluatie kan de onafhankelijkheid van evaluatoren in gevaar brengen.

### 2.3.6 Kritiek op beleidsevaluaties

In dit hoofdstuk zijn verschillende onderdelen van beleidsevaluaties uiteengezet. In de literatuur worden ook enkele kritiekpunten bij deze onderdelen geplaatst.

Met betrekking tot de verschillende perspectieven en evaluatiecriteria kan ten eerste worden gesteld dat beleidsevaluaties nauwelijks een definitief verband vaststellen tussen resultaten en de oorzaak van deze resultaten (Korsten, 2013). Om objectief vast te stellen dat beleid een bepaald effect heeft gehad, is het noodzakelijk om een nulmeting uit te voeren, een controlegroep in te zetten, en randomisering toe te passen (Van Noije, 2017). Echter, het is haast onmogelijk deze methoden toe te passen bij beleidsevaluaties. Hierdoor kan niet alleen de causaliteit worden betwist, maar ook de externe validiteit. Door externe invloeden, die mogelijk effect uitoefenen op de omstandigheden die worden onderzocht, uit te schakelen, wordt een fictieve situatie gecreëerd die niet representatief is voor de werkelijkheid. Ten tweede worden mogelijke oorzaken niet kritisch beoordeeld (Korsten, 2013). Dit komt mogelijk door de manier waarop evaluatievragen worden uitgezet. Evaluatievragen focussen met name op de doelmatigheid van beleid. In combinatie met beperkt budget en tijd is er geen ruimte om breder te evalueren dan deze opdracht. Dit heeft gevolgen voor de relevantie van beleid. Wanneer een beleidstheorie niet kan worden getoetst, zullen beleidsmakers pad afhankelijk blijven.

Met betrekking tot de doorwerking van beleidsevaluaties kunnen ook twee punten van kritiek worden onderscheiden. Ten eerste kan een beleidsevaluatie het beleidsproces op een onvoorziene en oneigenlijke manier beïnvloeden (Verlet & Devos, 2008). Zo kunnen

doelstellingen dusdanig worden geformuleerd dat deze makkelijk geëvalueerd kunnen worden of worden doelstellingen minder ambitieus ingezet zodat bij de evaluatie gesteld kan worden dat de beoogde resultaten zijn behaald. Ten tweede kunnen vraagtekens worden geplaatst bij de invloed van beleidsevaluaties op het beleid. Zo kan het zijn dat evaluaties leiden tot het vervangen, continueren, opschorten, of bijstellen van beleid (Pröpper, 1998), maar het kan ook voorkomen dat er geen enkele invloed wordt uitgeoefend (Verlet & Devos, 2008). De rol van beleidsevaluaties is vaak kleiner dan evaluatoren willen geloven.

Tot slot kunnen er twee kritiekpunten over de inhoud van beleidsevaluaties worden onderscheiden. Ten eerste vinden beleidsevaluaties vaak plaats in een micro context (Mathison, 2018). In plaats van het toetsen van de beleidstheorie zijn evaluaties gevoelig voor de ideeën van anderen over de effectiviteit van het beleid. Als gevolg hiervan heeft een evaluatie gelimiteerde mogelijkheden om de beleidstheorie kritisch te benaderen. Hier komt nog bij dat evaluaties van beleid vaak worden uitgevoerd door evaluatoren wiens referentiekader gelijk is aan dat van de beleidsmakers, waardoor de kans op een serieuze toets van de beleidstheorie verder wordt verkleind. Ten tweede kijken beleidsevaluaties te veel naar het verleden en niet genoeg naar de toekomst (Korsten, 2013). De uitvoering van het beleid wordt beoordeeld zonder dat er aandacht wordt besteed aan toekomstige problemen omtrent het beleid.

## 2.4 Externe, commerciële adviesbureaus

Beleidssevaluaties kunnen worden uitgevoerd door verschillende partijen (Bovens et al., 2017), waaronder externe evaluatoren die, in tegenstelling tot interne evaluatoren, geen onderdeel uitmaken van de organisatie waarvoor de evaluatie wordt uitgevoerd (De Peuter et al., 2007). Onder externe evaluatoren vallen onder andere publieke instanties zoals universiteiten, de Algemene Rekenkamer, en de ombudsman (Bovens et al., 2017). Maar ook organisaties als de Sociaal-Economische Raad en de Wetenschappelijke Raad voor het Regeringsbeleid leveren gevraagd en ongevraagd evaluaties van beleid. Daarnaast kunnen externe evaluaties worden uitgevoerd door private partijen, waaronder adviesbureaus. Dit onderzoek gaat in op evaluaties die zijn uitgevoerd door externe, commerciële adviesbureaus.

### 2.4.1 Definitie

In dit onderzoek hanteer ik hier de volgende definitie als het gaat om externe, commerciële adviesbureaus: een verzameling van individuele adviseurs die expertise verschaffen aan een derde partij tegen een vergoeding (Consultancy.nl, z.d.). De Orde van organisatiekundigen en -



adviseurs (z.d.) stelt dat deze adviseurs zich dienen te houden aan vier kernwaarden: deskundigheid, betrouwbaarheid, zorgvuldigheid, en professionele onafhankelijkheid.

Deskundigheid houdt in dat adviseurs de benodigde kennis, ervaring, competenties, vaardigheden, en tijd inzetten om een opdracht uit te voeren (Orde van organisatiekundigen en -adviseurs, z.d.). Daarbij moeten adviseurs in staat zijn een opgeleverd rapport te onderbouwen.

Echter, Bongers en Den Hertog (2007) stellen dat externe evaluatoren over het algemeen onvoldoende kennis hebben over het betreffende beleidsveld, waardoor er beperkte mogelijkheden zijn om diepgaande aanbevelingen te doen (Van der Zwan & Noordegraaf, 1998).

De tweede kernwaarde, betrouwbaarheid, houdt in dat een adviseur gemaakte afspraken nakomt, en de geheimhoudingsplicht naleeft (Orde van organisatiekundigen en -adviseurs, z.d.).

Vanuit zorgvuldigheid, de derde kernwaarde, wordt van adviseurs verwacht dat zij rekening houden met de wensen, verwachtingen, rechten, en belangen van alle betrokkenen (Orde van organisatiekundigen en -adviseurs, z.d.). Dit houdt ook in dat de opgestelde aanbevelingen rekening houden met mogelijke effecten op derden.

Desondanks constateert Korsten (2013) dat evaluatoren zich onvoldoende focussen op het netwerk rondom beleid en zich met name richten op de opdrachtgever. Evaluatoren zouden de zorgvuldigheidswaarde meer moeten naleven en niet moeten uitgaan van een top-down benadering (Klijn & Koppenjan, 2001).

Tot slot wordt verwacht dat adviseurs de opdracht aanpakken vanuit een deskundig oogpunt, onafhankelijk van persoonlijke belangen of belangen van derden (Orde van organisatiekundigen en -adviseurs, z.d.).

Echter, een volledig onafhankelijke adviseur bestaat niet (De Peuter et al., 2007). Omdat externe adviseurs worden betaald voor hun dienst is het mogelijk dat zij gevoelig zijn voor invloed vanuit de opdrachtgever (Mathison, 2018). Daarbij staan beleidsverantwoordelijken onder druk om de meerwaarde van hun beleidskeuzes aan te tonen (De Peuter et al., 2007). Ondanks dat adviseurs niet verantwoordelijk zijn voor het beleid, speelt deze druk wel een rol in de relatie tussen de opdrachtgever, de beleidsverantwoordelijke, en de adviseur.

#### 2.4.2 Historische ontwikkeling

Het begrip en de rol van externe, commerciële adviesbureaus is door de geschiedenis heen sterk veranderd (Strikwerda, 2004). De oorsprong van de externe adviseur wordt toegekend aan Taylor, die aanraade adviseurs in te zetten om, in lijn met het *Scientific Management*, de

arbeidsproductiviteit te verhogen (Strikwerda, 1998). In deze rol waren adviseurs continu bezig managers ervan te overtuigen hun advies over te nemen ten behoeve van efficiëntie.

Na de Tweede Wereldoorlog kwam de *Human Relations School*, en daarmee de psychologische en sociologische behoeften van de mens, meer centraal te staan (Mullins, 2002; Scott, 2003). Externe adviseurs werden ingehuurd om verandering teweeg te brengen met het doel vooruitgang te boeken. Om dit te realiseren moesten leidinggevenden en medewerkers hun eigen overtuigingen inruilen voor de meer rationele en efficiënte opvattingen van de adviseurs.

In 1982 kwam een tweede verschuiving op gang als gevolg van het boek *In search of Excellence* (Strikwerda, 2004). Dit boek focust op standaardmodellen voor adviseurs. Deze aanpak was een breuk met de rol die adviseurs eerder aannamen en zorgde voor een sterke opkomst van commerciële adviesbureaus en een verstening van methoden. In plaats van organisaties en leidinggevenden uit te dagen, werden deze bureaus een verlengstuk van de organisaties die zij beoordeelden. De nieuwe generatie adviseurs is dan ook voornamelijk getraind op het implementeren van simplistische standaardmodellen en mist kennis van de wetenschappelijke theorieën waar de modellen op zijn gestoeld. Hiermee verloor het vak een deel van de wetenschappelijke grondslag en het vermogen om, om te gaan met fundamentele veranderingen in de samenleving.

De toekomst van adviesbureaus zal liggen in het overbrengen van routinekennis, het oplossen van traditionele problemen, en het uitvoeren van opdrachten voor traditionele aanpassingen (Strikwerda, 2004), terwijl de complexiteit van maatschappelijke problemen alleen maar toe zal nemen (Korsten, 2016b).

## 2.5 Conclusie

Op basis van dit hoofdstuk zal ik een aantal onderdelen specifiek meenemen in het vervolg van dit onderzoek. Ten eerste zal de schaal van wickedness, gebaseerd op de beschikbaarheid van kennis, de aanwezigheid van consensus, en het veld van de betrokken actoren, een centrale rol spelen in dit onderzoek. Ten tweede zullen de basisonderdelen van de rapportage van beleidsevaluaties - managementsamenvatting, inleiding, evaluatiedesign, resultaten, conclusie, en aanbevelingen - worden getoetst. Ook het evaluatieproces, bestaande uit de structurering, dataverzameling, data-analyse, en beoordelingsfase, wordt meegenomen in dit onderzoek. Tot slot zullen de instrumenten waar externe evaluatoren gebruik van kunnen maken bij het verzamelen van data worden onderzocht: documentenanalyses, interviews, focusgroepen, expertpanels, en workshops.

### 3. Methodologie

#### 3.1 Casestudy

Voor dit onderzoek heb ik twaalf rapportages van beleidsevaluaties geselecteerd, opgesteld door vijf commerciële adviesbureaus. De selectie van de casussen vond plaats aan de hand van een oppervlakkige analyse van de rapportages. Bij deze analyse focuste ik op de status van de evaluatie, deze moest volledig afgerond zijn. Daarbij moesten de rapportages openbaar zijn gemaakt, in verband met eventuele vertrouwelijke informatie. De uitvoerende adviesbureaus zijn: Andersson Elffers Felix, Berenschot B.V., Deloitte, KWINK Groep, en PWC. De selectie van bovenstaande bureaus is gemaakt op basis van hun toonaangevende status binnen het veld van beleidsevaluaties en de beschikbaarheid van de rapportages. De evaluaties van Berenschot B.V. zijn verkregen door middel van een uitvraag bij medewerkers. Hierbij werd gevraagd naar evaluaties van beleid waarvan de rapportage openbaar is gemaakt. De evaluaties van de overige bureaus zijn verkregen door middel van Google searches op de trefwoorden “openbare beleidsevaluatie” en “evaluatie overheidsbeleid” in combinatie met de zoekterm “extern”. In tabel 1 zijn de geraadpleegde beleidsevaluaties weergegeven, gecategoriseerd per bureau op alfabetische volgorde en gerangschikt op omgekeerde chronologische volgorde.

**Tabel 1**

*Geselecteerde beleidsevaluaties*

	<i>Bureau</i>	<i>Evaluatie</i>	<i>Datum van oplevering</i>
1	Andersson Elffers Felix	Evaluatie van de Aanwijzingen inzake de rijksinspecties	30 november 2020
2		Evaluatie Transitie Autoriteit Jeugd	18 december 2019
3	Berenschot B.V.	Sturing op duurzaamheid. Collegeonderzoek naar de sturing van gemeente Apeldoorn op klimaatadaptatie, circulaire economie en de energietransitie	4 februari 2022
4		Evaluatie Radioactief afval. Een evaluatie van de Nota Radioactief afval en het Nationaal Programma Radioactief afval	1 februari 2022

5		Preventie gehoorschade. Aanpak problematische blootstelling versterkte muziek	9 december 2021
6		Evaluatie Wet taken meteorologie en seismologie	1 november 2021
7	Deloitte	Evaluatie Convenant onbedwelmd slachten volgens religieuze ritens	15 november 2021
8		Periodieke evaluatie Gezondheidsraad 2017-2021	8 maart 2022
9		Evaluatie Convenant e-lending	30 november 2021
10	KWINK Groep	Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant 2016-2021 (onderdeel Milieu)	29 oktober 2021
11		Beleidsvaluatieonderzoek Streekomroepvorming	7 september 2021
12	PWC	Ministerie van BZK Evaluatie Nationaal Energie Bespaarfonds 2014-2019	Juli 2021

### 3.2 Documentenanalyse

Dit onderzoek is gebaseerd op kwalitatieve data, verzameld aan de hand van twee inhoudsanalyses. In de eerste plaats werd de operationalisering uit het theoretisch kader deductief getoetst via Atlas.ti. Omdat er niet voor elke deelvraag voldoende literatuur beschikbaar was en omdat er in sommige rapportages elementen naar voren kwamen die niet in de literatuur terugkwamen, is er ook inductief gecodeerd. De inductieve coderingen zijn via Atlas.ti getoetst in alle geselecteerde rapportages. Een overzicht van de deductieve coderingslijst (tabel 30) en inductieve coderingslijst (tabel 31) is te vinden in bijlage 1.

Na het toetsen en verzamelen van alle deductieve en inductieve coderingen, is er via Atlas.ti een uitdraai gemaakt van de data. Deze data heb ik vervolgens geordend in overkoepelende thema's, wat de basis vormt voor de analyse in dit onderzoek.

### 3.3 Betrouwbaarheid & validiteit

De betrouwbaarheid van onderzoek heeft twee onderdelen: nauwkeurigheid en consistentie (Van Thiel, 2021). Nauwkeurigheid verwijst naar de gehanteerde meetinstrumenten: meten deze instrumenten de juiste variabelen. Consistentie verwijst naar de herhaalbaarheid van het onderzoek: wordt bij replicatie van het onderzoek, hetzelfde resultaat behaald.

De grootste limitatie met betrekking tot de betrouwbaarheid van dit onderzoek is dat de literatuur niet volledig aansluit bij de complexiteit van de praktijk, waardoor een zekere mate van eigen interpretatie plaats heeft gevonden. Dit is met name van toepassing bij het beantwoorden van de eerste deelvraag. Om dit zoveel mogelijk te ondervangen en de algemene betrouwbaarheid van dit onderzoek te vergroten, heb ik gebruik gemaakt van zowel deductieve als inductieve coderingen. Daarnaast heb ik continu aangetoond op basis van welke gegevens ik tot een bepaalde conclusie ben gekomen. Ten tweede zijn de coderingslijsten consequent bijgehouden in Atlas.ti, getoetst in alle geselecteerde rapportages, en opgenomen in de bijlagen van dit rapport. Tot slot maak ik gebruik van openbare rapportages. Dit maakt exacte replicatie van dit onderzoek gemakkelijker, omdat deze evaluaties toegankelijk zijn voor iedereen. Daarbij ligt de inhoud van de rapportages vast, waardoor de bronnen voor iedereen gelijk zijn.

De validiteit van onderzoek kent ook twee onderdelen: interne en externe validiteit (Van Thiel, 2021). De interne validiteit verwijst naar de geldigheid van het onderzoek: wordt gemeten wat gemeten moet worden. De externe validiteit verwijst naar generaliseerbaarheid: gelden de bevindingen ook buiten de kaders van het onderzoek. In het geval van casestudy's is de externe validiteit vaak klein, terwijl de interne validatie vaak hoog is.

Dit onderzoek maakt gebruik van een relatief klein aantal rapportages, wat de externe validiteit inderdaad beperkt. Om deze beperking enigszins te ondervangen maak ik, ter aanvulling op de tactieken genoemd ter bevordering van de betrouwbaarheid, gebruik van beleidsevaluaties uitgevoerd door verschillende adviesbureaus.

### 3.4 Biases & overige limitaties

Aanvullend op de limitaties die ik hierboven heb genoemd, was er een beperkte hoeveelheid tijd om dit onderzoek uit te voeren. Dit heeft met name gevolgen voor het aantal geanalyseerde casussen en de beperkte gehanteerde methoden. Dit onderzoek is enkel gebaseerd op een inhoudelijke documentenanalyse, met meer tijd waren interviews en een analyse van de visuele vormgeving van de rapportages een waardevolle aanvulling geweest. Een groot voordeel van de documentenanalyse is dat de data vastligt. Hierdoor is er een minimaal risico op sociaal wenselijke antwoorden. De sociale wenselijkheden die eventueel wel aanwezig zijn in de geselecteerde casussen, zullen voor alle onderzoekers van deze data gelijk zijn.

Tot slot zijn de persoonlijke biases van onderzoekers een risico bij het uitvoeren van onderzoek. Om dit zoveel mogelijk te minimaliseren is systematisch gebruik gemaakt van de deductieve en inductieve coderingslijsten, en zijn de namen van de adviesbureaus zoveel mogelijk weggelaten zodat dit geen rol speelt bij de beoordeling van de data.

## 4. Beleidsevaluaties op een schaal van wickedness

Om de hoofdvraag van dit onderzoek te beantwoorden zullen de geselecteerde beleidsevaluaties allereerst worden ingedeeld op een schaal van wickedness (figuur 1). Na een analyse per evaluatie (tabel 2 t/m 13) sluit dit hoofdstuk af met een totaaloverzicht van de geselecteerde beleidsevaluaties op een schaal van wickedness (figuur 2).

### 4.1 Evaluatie van de Aanwijzingen inzake de rijksinspecties

De eerste evaluatie betreft een onderzoek naar de Aanwijzingen voor Rijksinspecties. De Aanwijzingen dienen een bijdrage te leveren aan de balans tussen onafhankelijke oordeelsvorming van rijksinspecties en de ministeriële verantwoordelijkheid voor het opereren van de inspectie (Veldhuisen et al., 2020).

De Aanwijzingen zijn gestoeld op advies van de Wetenschappelijke Raad voor het Regeringsbeleid om meer focus te leggen op onafhankelijk toezicht met oog voor systeemrisico's (Veldhuisen et al., 2020).

Er is consensus tussen de betrokken partijen over het doel dat wordt nagestreefd met de Aanwijzingen: "Er is een brede consensus dat onafhankelijk opererend rijkstoezicht van groot maatschappelijk belang is" (Veldhuisen et al., 2020, p. 5). Daarbij is er consensus over de wijze waarop de Aanwijzingen dit doel nastreven: "De Aanwijzingen worden door alle betrokkenen gezien als een waardevolle en belangrijke set aan regels voor de verhouding tussen inspecties en departementen" (Veldhuisen et al., 2020, p. 43).

Tot slot is bekend welke partijen betrokken zijn bij het beleid: departementen en de rijksinspecties.

**Tabel 2**

*Categorisering casus 1*

Kennis probleem	Beschikbaar	Gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

## 4.2 Evaluatie Transitie Autoriteit Jeugd

De Transitie Autoriteit Jeugd (TAJ) is een onafhankelijke commissie opgericht in 2014 met het doel de continuïteit van zorg voor jeugdigen en gezinnen te helpen borgen tijdens de stelselherziening jeugd (Askari et al., 2019). De TAJ had ook als taak te voorkomen dat onmisbare functies uit het zorglandschap zouden verdwijnen. De evaluatie focust op de doelmatig- en doeltreffendheid van de TAJ, die in 2019 is opgeheven.

De taak van de TAJ, “het voorkomen dat essentiële functie zouden verdwijnen”, was onvoldoende duidelijk gedefinieerd (Askari et al., 2019, p.5). In combinatie met de context waarbinnen de TAJ opereerde, de implementatie van de nieuwe Jeugdwet, geeft dit de indruk dat er beperkte kennis beschikbaar was over de specifieke problemen waarmee de TAJ te maken zou krijgen. Het wordt uit de evaluatie niet duidelijk waarom voor deze specifieke inrichting van de TAJ is gekozen als oplossing van het probleem.

Er was breed draagvlak bij gemeenten, jeugdhulpaanbieders, en de betrokken departementen voor de instelling van de commissie. Er was consensus over de tijdelijke aard van de TAJ. Binnen het stelsel was er frictie over de doelstellingen: “de TAJ heeft bijgedragen aan de schaalvergroting bij jeugdhulpaanbieders, maar deze ontwikkeling was door gemeenten lang niet altijd gewenst, gegeven hun wens om - in lijn met de transformatie - te werken aan zorg dichtbij en aansluiting op de lokale situatie” (Askari et al., 2019, p. 39).

Het veld van direct betrokken actoren is beperkt tot het ministerie van VWS. Het veld van indirect betrokkenen is breder, doch bekend: gemeenten en jeugdhulpaanbieders.

**Tabel 3**

### *Categorisering casus 2*

Kennis probleem	Niet beschikbaar	Matig gestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

### 4.3 Sturing op duurzaamheid. Collegeonderzoek naar de sturing van gemeente Apeldoorn op klimaatadaptatie, circulaire economie en de energietransitie

Duurzaam Apeldoorn, één van de vier strategische doelen uit het Apeldoornse bestuursakkoord, is uitgewerkt in drie thema's: energie(transitie), water en natuur (klimaatadaptatie), en circulaire economie (Kohlmann et al., 2022). Eén van de doelstellingen is dat Apeldoorn in 2040 circulair en klimaatneutraal is. De evaluatie focust onder meer op de doelmatig- en doeltreffendheid van het beleid en welke ervaringen kunnen worden meegenomen naar de ontwikkeling van een nieuw uitvoeringskader.

Er wordt binnen de gemeente geen algemeen geldende definitie voor duurzaamheid en circulaire economie gehanteerd. Daarbij is er niet voor elk thema een doel opgesteld. Bij de doelstellingen die wel bekend zijn, is het niet altijd mogelijk om vast te stellen of deze zijn behaald: “het is onmogelijk om real-time inzicht te hebben in de gerealiseerde reductie van CO2 uitstoot in Apeldoorn” (Kohlmann et al., 2022, p. 19). Er is eveneens beperkte consensus:

“Diverse interne betrokkenen twijfelen aan de haalbaarheid van de ambitie om in 2050 energieneutraal te zijn. Sommige van de betrokkenen geven aan dat het in theorie wel haalbaar is, maar dat daarvoor andere (stevigere) strategische keuzes nodig zijn. Diverse gesprekspartners zien in de praktijk dat de inzet op duurzaamheidsthema's het regelmatig aflegt tegen concurrerende belangen” (Kohlmann et al., 2022, p. 23).

Naast de gemeente zijn diverse andere actoren betrokken. Er zijn samenwerkingen op verschillende bestuursniveaus en met private actoren. Betrokken actoren, zowel direct als indirect, zijn onder meer: waterschappen, de Regionale Energiestrategie, de G40, en burgerinitiatieven. Echter, het is niet bekend welke actoren hier nog meer bij horen.

**Tabel 4**

*Categorisering casus 3*

Kennis probleem	Niet beschikbaar	Ongestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Niet aanwezig	
Consensus probleem	Niet aanwezig	
Consensus oplossing	Niet aanwezig	
Betrokken actoren	Onbekend	



#### 4.4 Evaluatie Radioactief afval. Een evaluatie van de Nota Radioactief afval en het Nationaal Programma Radioactief afval

Het Nationaal Programma is opgesteld in 2016 en bouwt voort op de Nota uit 1984. Het doel van dit beleid is om mensen, dieren, planten, en goederen te beschermen tegen de negatieve effecten van straling (Oostdijk et al., 2022).

Over het achterliggende beleidsprobleem is tot op een bepaalde hoogte kennis beschikbaar. Echter, in verband met de lange periode die het opslaan van radioactief afval betreft, wordt er continu onderzoek gedaan naar alternatieve mogelijkheden om afval op te slaan en de langetermijneffecten hiervan.

Betreffende de consensus over het doel van het beleid: dit is aanwezig tussen de actoren die direct betrokken zijn bij het beleid; zij benadrukken het belang van de vier uitgangspunten waar het beleid op is gebouwd.

Er zijn naast de directe betrokkenen ook indirecte betrokkenen: wetenschappelijke kennisinstututen, burgers, NGO's, en bedrijven, en er wordt op internationale schaal samengewerkt om, onder meer, kennis uit te wisselen. Het is onbekend of dit beeld compleet is.

**Tabel 5**

*Categorisering casus 4*

Kennis probleem	Beschikbaar	Matig gestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Onbekend	

#### 4.5 Preventie gehoorschade. Aanpak problematische blootstelling versterkte muziek

De vijfde geselecteerde evaluatie betreft het Derde Convenant Preventie Gehoorschade, ondertekend door het Ministerie van VWS en acht partners.

Betreffende de relatie tussen gehoorschade en blootstelling aan versterkte muziek is er geen kennis beschikbaar die causaliteit tussen deze twee elementen kan vaststellen. Het is namelijk onmogelijk dit wetenschappelijk te toetsen:

“Het is een uitdaging om vast te stellen bij welk geluidsniveau precies gehoorschade optreedt, te meer omdat het niet alleen om het niveau, maar ook om de duur van blootstelling gaat en het

ethisch niet verantwoord is een groep mensen gecontroleerd bloot te stellen aan hard geluid” (Blokzijl et al., 2021, p. 10).

Ook kennis over een mogelijke oplossing is niet aanwezig, aangezien het niet mogelijk is om wetenschappelijk te onderbouwen in welke mate het terugdringen van hard geluid invloed heeft op gehoorschade.

Over de oorzaak van gehoorschade is dan ook geen consensus: “causaliteit tussen gehoorschade en versterkte muziek is een terugkerend debat” (Blokzijl et al., 2021, p. 10). Echter, de betrokken partijen zijn het er wel over eens dat de Nederlandse maximum geluidsnorm stapsgewijs gelijk moet worden getrokken met de aanstaande WHO-richtlijn.

De betrokken actoren zijn, onder andere, ministeries, gemeenten, sportscholen, en festivals. Het is niet duidelijk wie van de betrokken actoren verantwoordelijk is voor controle op het naleven van een maximum geluidsnorm.

**Tabel 6**

*Categorisering casus 5*

Kennis probleem	Niet beschikbaar	Ongestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Niet aanwezig	
Consensus probleem	Niet aanwezig	
Consensus oplossing	Niet aanwezig	
Betrokken actoren	Onbekend	

#### 4.6 Evaluatie Wet taken meteorologie en seismologie

De Wet taken meteorologie en seismologie, in werking getreden in 2016, deelt taken toe aan de minister van Infrastructuur en Waterstaat over het informeren van de samenleving op het gebied van weer, klimaat, en bodem beweging (Oostdijk et al., 2021).

Over het belang van kwalitatief goede kennisvoorziening op het gebied van meteorologie en seismologie is eenduidige kennis beschikbaar.

Er is consensus tussen de direct betrokken actoren over de doelen die worden nagestreefd. Echter, de betrokken actoren zijn het oneens over de vraag of het probleem van publieke of private aard is. Hierop volgt de vraag of het aan de markt of de overheid is om in een oplossing te voorzien.

De directe betrokkenen zijn weerbedrijven en het KNMI, waartussen op dit moment nauwelijks samenwerking is, ondanks de wederzijdse afhankelijkheid: weerbedrijven zijn afhankelijk van het meetnetwerk en de modellen van het KNMI, het KNMI is afhankelijk van het brede bereik dat weerbedrijven hebben bij het publiek.

**Tabel 7**

*Categorisering casus 6*

Kennis probleem	Beschikbaar	Slecht gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Niet aanwezig	
Consensus oplossing	Niet aanwezig	
Betrokken actoren	Bekend	

#### 4.7 Evaluatie Convenant onbedwelmd slachten volgens religieuze riten

In 2012 is dit Convenant overeengekomen met als doel verbeteringen door te voeren op het gebied van dierenwelzijn bij het slachtproces wanneer gebruik wordt gemaakt van israëlitische of islamitische slachtmethoden (Deloitte, 2021).

Over het beleidsprobleem is kwantitatieve data beschikbaar betreffende de hoeveelheid slachtingen die jaarlijks plaatsvindt in Nederland. Over de invloed van onbedwelmd slachten op dieren is wetenschappelijke kennis tot op zeker hoogte beschikbaar, maar continu in ontwikkeling. Wet- en regelgeving wordt aangepast aan de hand van deze ontwikkelingen.

Er is bij de direct betrokken partijen consensus over de doelstelling van het evaluandum:

“Verbeteringen door te voeren op het gebied van dierenwelzijn bij het slachtproces, waarbij moet worden uitgegaan van de optimale omstandigheden in relatie tot het welzijn van het dier, op zodanige wijze die verenigbaar is met de israëlitische of islamitische ritus” (Deloitte, 2021, p. 11).

Een eventueel (gebrek aan) consensus vanuit de samenleving en andere partijen is niet meegenomen in de evaluatie.

Er zijn vier partijen vertegenwoordigd in het convenant. Daarbij hebben de partijen rekening te houden met Nederlandse en Europese wet- en regelgeving.

**Tabel 8***Categorisering casus 7*

Kennis probleem	Beschikbaar	Gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

#### 4.8 Periodieke evaluatie Gezondheidsraad 2017-2021

In 1902 werd de Gezondheidsraad ingesteld bij Gezondheidszorgwet om de ministers en de Staten-Generaal in te lichten over de stand der wetenschap betreffende de volksgezondheid (Van Mil et al., 2022).

De kennis over het probleem en de oplossing staan niet ter discussie.

Betreffende de consensus over het beleidsprobleem en de oplossing hiervoor geven “zowel betrokkenen van binnen de Gezondheidsraad als belanghebbenden buiten de Gezondheidsraad aan dat de raad op gedegen en betrouwbare wijze adviseert op basis van de stand van de wetenschap” (Van Mil et al., 2022, p. 32).

Enkel de Raad zelf is direct betrokken bij het beleid. Daarnaast maken verschillende partijen gebruik van de adviezen van de Raad, namelijk de Eerste en Tweede Kamer, en verschillende gezondheidsactoren.

**Tabel 9***Categorisering casus 8*

Kennis probleem	Beschikbaar	Gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

## 4.9 Evaluatie Convenant e-lending

In 2019 hebben zes organisaties afspraken gemaakt over het (uit)lenen van e-books om een breed aanbod te creëren via de landelijke digitale openbare bibliotheek (Mulder et al., 2021).

Over het probleem en de omvang hiervan is kwantitatieve data beschikbaar. Deze data betreft de hoeveelheid e-books die beschikbaar is, hoeveel e-books worden uitgeleend, en hoeveel mensen gebruik maken van e-books.

Over de doelstellingen zijn de betrokken partijen het eens: “alle partijen geven in gesprekken aan tevreden te zijn met de doelstellingen van het Convenant. Deze doelstellingen zijn en blijven relevant voor partijen” (Mulder et al., 2021, p. 19).

Bij het opstellen van het Convenant zijn acht partijen betrokken geweest: de Koninklijke Bibliotheek, de Vereniging Openbare Bibliotheken, Stichting Lira, Stichting Pictoright, Stichting Leenrecht, de Auteursbond, de Groep Algemene Uitgevers, en het Ministerie van Onderwijs, Cultuur en Wetenschap.

**Tabel 10**

*Categorisering casus 9*

Kennis probleem	Beschikbaar	Gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

## 4.10 Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant 2016-2021 (onderdeel Milieu)

In 2016 is de beleidsnota: Provinciaal Milieu- en waterplan (PMWP) vastgesteld met als doel het Brabantse milieu- en waterbeleid te versterken (Bernoster et al., 2021).

“Betrokkenen geven in overwegende mate aan beperkt kennis te hebben van beschikbare middelen, de inzet daarvan en de daarmee verband houdende doelmatigheid van de activiteiten die voortvloeien uit het PMWP” (Bernoster et al., 2021, p. 5). Uit de evaluatie wordt niet duidelijk welke kennis er is omtrent het probleem zelf.

In het beleid worden verschillende doelen gesteld. Echter, deze doelen zijn niet meetbaar gemaakt. Hierdoor is het voor betrokkenen dan ook niet duidelijk in hoeverre doelstellingen zijn

behaald. Het is onduidelijk of dit komt doordat er onvoldoende kennis beschikbaar is, of dat partijen het niet eens kunnen worden over de specifieke doelen.

Naast de provincies, zijn ook gemeenten betrokken. Verder is het onduidelijk welke actoren direct of indirect betrokken zijn bij dit beleid.

**Tabel 11**

*Categorisering casus 10*

Kennis probleem	Niet beschikbaar	Ongestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Niet aanwezig	
Consensus probleem	Niet aanwezig	
Consensus oplossing	Niet aanwezig	
Betrokken actoren	Onbekend	

#### 4.11 Beleidsevaluatieonderzoek Streekomroepvorming

Eén van de manieren om te voorzien in een lokaal toereikend media-aanbod is het vormen van een streekomroep (Van Schelven et al., 2021). Het idee achter streekomroepvorming is dat lokale omroepen professioneler kunnen werken, waardoor de kwaliteit en continuïteit van de lokale journalistiek gerealiseerd en versterkt kan worden.

De kennis betreffende het beleidsprobleem is beschikbaar door middel van een onderzoek van de Raad voor Cultuur en de Raad voor het Openbaar Bestuur. Zij concluderen dat “lokale media onmisbaar zijn. Zij vervullen een cruciale rol in lokale gemeenschappen [...]” (Van Schelven et al., 2021, p. 8). Betreft de oplossing van het probleem stellen de Raden dat er in lokale media moet worden geïnvesteerd en dat samenwerking tussen omroepen moet worden gestimuleerd: “de Raden zien kansen om te investeren in een netwerk van mediafondsen als impuls voor regionale en lokale kwaliteitsjournalistiek” (Van Schelven et al., 2021, p. 8).

Uit de evaluatie blijkt dat de wetgevende en uitvoerende partijen gelijkgestemd zijn als het aankomt op de probleemstelling. Echter, de omroepen kunnen zich minder goed vinden in de voorgestelde oplossing: “verschillende omroepen hebben het beeld dat de NLPO-streekomroepvorming (en specifiek het fuseren van lokale omroepen) probeert te forceren. [...] Bij deze omroepen is behoefte aan meer aandacht voor de pluriformiteit van streken en meer ruimte voor maatwerk” (Van Schelven et al., 2021, p. 25).

De actoren die direct zijn betrokken bij dit beleid zijn het Ministerie van Onderwijs, Cultuur, en Wetenschap, de Raad voor Cultuur, de Raad voor het Openbaar Bestuur, de VNG, en de individuele gemeenten. Indirect betrokken actoren zijn de lokale omroepen, en regionale en landelijke publieke media-instellingen.

**Tabel 12**

*Categorisering casus 11*

Kennis probleem	Beschikbaar	Slecht gestructureerd probleem
Kennis oplossing	Beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Niet aanwezig	
Betrokken actoren	Bekend	

#### 4.12 Ministerie van BZK Evaluatie Nationaal Energie Bespaarfonds 2014-2019

Het Nationaal Energiebespaarfonds (NEF, nu het Nationaal Warmtefonds) verstrekt sinds 2014 de Energiebespaarlening; een lening gericht op particuliere woningbezitters, VvE's, en scholen om de verduurzaming van hun woning of gebouwen te kunnen financieren (PWC, 2021). Het opzetten van een financieel instrument voor verduurzaming is onderdeel van een duurzaamheidspakket met als doel burgers in staat te stellen te investeren in woningverduurzaming en als stimulant voor de bouwsector.

Er is kennis beschikbaar over in welke mate er gebruik is gemaakt van het fonds. Echter, de data die inzicht geeft op de daadwerkelijk gerealiseerde energiebesparing is zeer beperkt. Verschillende onderzoeken hanteren verschillende methoden en maatstaven. Hierdoor is er beperkt kennis beschikbaar over de werkzaamheid van de oplossing.

Er is consensus tussen de betrokken partijen over het probleem dat het NEF beoogt op te lossen. De actoren die direct en indirect zijn betrokken hebben consensus bereikt over het nut en noodzaak van dit beleid. Dit wijst zich uit de financiële steun die de betrokken actoren beschikbaar stellen.

De actoren die betrokken zijn bij dit beleid zijn onder andere het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Rabobank, ASN Bank, Council of Europe Development Bank, en de European Investment Bank.

**Tabel 13***Categorisering casus 12*

Kennis probleem	Beschikbaar	Matig gestructureerd probleem
Kennis oplossing	Niet beschikbaar	
Consensus kennis	Aanwezig	
Consensus probleem	Aanwezig	
Consensus oplossing	Aanwezig	
Betrokken actoren	Bekend	

**4.13 Conclusie**

Bovenstaande beleidsevaluaties zijn geanalyseerd aan de hand van de kennis die beschikbaar is over het probleem en de oplossing, de consensus die er tussen de betrokken actoren is over de kennis, het probleem, en de oplossing, en tot slot, of het bekend is welke actoren betrokken zijn.

De evaluaties waarbij kennis beschikbaar en consensus aanwezig is, vallen in de gestructureerde categorie. Van de geselecteerde evaluaties zijn dit de Periodieke evaluatie gezondheidsraad, de Evaluatie van de Aanwijzingen inzake de rijksinspecties, de Evaluatie van het Convenant onbedweld slachten volgens religieuze riten, en de Evaluatie van het Convenant e-lending.

De evaluaties waarbij er wel kennis beschikbaar is, maar geen consensus vallen in de slecht gestructureerde categorie. Het Beleidsevaluatieonderzoek Streekomroepvorming, en de Evaluatie Wet taken meteorologie en seismologie vallen in deze categorie.

De Evaluatie Nationaal Energie Bespaarfonds, de Evaluatie Transitie Autoriteit en Jeugd, en de Evaluatie Radioactief Afval vallen in de derde categorie, de matig gestructureerde categorie. Dit wil zeggen dat er consensus is, maar dat niet alle kennis beschikbaar is.

Tot slot kan het voorkomen dat er zowel geen kennis als consensus beschikbaar is. Deze problemen zijn ongestructureerd, en hebben de hoogste mate van wickedness. Van de geselecteerde evaluaties vallen de Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant, Sturing op duurzaamheid, en Preventie gehoorschade in deze categorie.

De categorisering van de geselecteerde beleidsevaluaties vormt de basis voor de rest van dit onderzoek. Een visuele weergave van de geselecteerde beleidsevaluaties op een schaal van wickedness is te vinden in figuur 2.



## Figuur 2

### *Beleidsvaluaties op een schaal van wickedness*

Gestructureerd	Slecht gestructureerd	Matig gestructureerd	Ongestructureerd
<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Beleidsvaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>
<i>Evaluatie van de Aanwijzingen inzake de rijksinspecties</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Sturing op duurzaamheid</i>
<i>Evaluatie Convenant onbedweld slachten volgens religieuze riten</i>		<i>Evaluatie Radioactief afval</i>	<i>Preventie gehoorschade.</i>
<i>Evaluatie Convenant e-lending</i>			

## 5. De Inrichting

Op basis van de categorisering in figuur 2 wordt in dit hoofdstuk de tweede deelvraag beantwoord: *In welke mate speelt de gradatie van wickedness en rol bij de inrichting van de rapportage van de geselecteerde beleidsevaluaties?* Een meer uitgebreide versie van de gepresenteerde tabellen is te vinden in bijlage 2.

### 5.1 Gestructureerde categorie

In tabel 14 is de inrichting van de rapportages van de beleidsevaluaties in de gestructureerde categorie weergegeven.

**Tabel 14**

*Inrichting gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedweld slachten volgens religieuze ritten</i>	<i>Evaluatie Convenant E-lending</i>
Inhoud	Inhoud	Woord vooraf	Inhoud
Managementsamen- vatting	Managementsamen- vatting	Managementsamen- vatting	Inleiding
Inleiding	Inleiding	Inhoudsopgave	Resultaten
Terugblik	Onderzoeksaanpak	Introductie	Ervaringen
Realisatie publieke waarde	De vormgeving van de Aanwijzingen	Evaluatie van afspraken over normen ter verbetering van het dierenwelzijn rond het doden	Conclusies en aanbevelingen
Conclusies en aanbevelingen	De uitwerking van de Aanwijzingen	Evaluatie van afspraken niet direct gerelateerd aan het doden	
	De praktijkwerking van de Aanwijzingen	Conclusies en aanbevelingen	
	Conclusies en aanbevelingen		

Van de vier evaluaties starten drie met een inhoudsopgave en één met een woord vooraf. Drie van de vier evaluaties hebben een managementsamenvatting, waarvan twee evaluaties deze laten aansluiten op de inhoudsopgave en één op het woord vooraf. Inhoudelijk zijn twee van de drie managementsamenvattingen meer uitgebreid dan slechts het noemen van de aanleiding, hoofdvragen, conclusies, en aanbevelingen. Deze evaluaties benoemen ook de gehanteerde methoden, bevindingen, conclusies, en aanbevelingen in de managementsamenvatting.

Alle vier de evaluaties maken gebruik van een inleidend hoofdstuk. De inhoud van de inleidingen varieert. Drie van de vier evaluaties beschrijven de achtergrond van het beleid, waarvan twee dit combineren met de aanleiding van de evaluatie. De vierde evaluatie combineert de aanleiding van het onderzoek met de context van het beleid. Eén evaluatie gaat in het inleidende hoofdstuk in op de betrokken partijen en een andere bespreekt de details van het geëvalueerde beleid. Drie evaluaties bespreken de gehanteerde onderzoeks aanpak waarvan één evaluatie dit combineert met de gehanteerde methoden.

Het hoofdstuk dat volgt na de inleiding is in één van de evaluaties een terugblik naar een vorige evaluatie van het beleid en in een andere evaluatie de onderzoeks aanpak. De twee overige evaluaties gaan na de inleiding direct in op de resultaten van de evaluatie.

Betreffende de hoofdstukken waarin de resultaten worden besproken, bespreekt één van de evaluaties de resultaten aan de hand van het publieke waarde model van Mark Moore. Een andere evaluatie beschrijft de resultaten aan de hand van de vormgeving, uitwerking, en praktijkwerking van het beleid. De twee andere evaluaties voeren in het resultaten onderdeel een inhoudelijke evaluatie uit zonder specifiek analysekader.

Na de hoofdstukken waarin de resultaten worden besproken, en in een enkel geval na het hoofdstuk waarin de ervaringen van de betrokken actoren worden besproken, volgen de conclusies en aanbevelingen. De eerste evaluatie bespreekt in dit hoofdstuk de antwoorden op de onderzoeksvragen, waarvan één van de drie hoofdvragen ook wordt voorzien van een aanbeveling. Een andere evaluatie heeft de indeling van het slothoofdstuk gespiegeld aan de indeling van de resultaten, waarbij per thema de conclusies en aanbevelingen worden besproken. De derde evaluatie bespreekt in de conclusie de doeltreffend- en doelmatigheid van het beleid, gevolgd door aanbevelingen. De laatste evaluatie bespreekt de conclusies gevolgd door de aanbevelingen.

## 5.2 Slecht gestructureerde categorie

In tabel 15 is de tweede categorie beleidsevaluaties weergegeven die wordt geanalyseerd aan de hand van de inrichting: de slecht gestructureerde categorie.

**Tabel 15**

*Inrichting slecht gestructureerde categorie*

<i>Beleidsvaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Inhoudsopgave	Inhoudsopgave
Inleiding	Inleiding
Achtergrond	Een maatschappelijk doel, een stelsel, en een wet
Stand van zaken	Werking van het stelsel van meteorologie en seismologie
Betrokken partijen	Werking van de Wet taken meteorologie en seismologie
Conclusie en aanbevelingen	Conclusies en aanbevelingen
	Beantwoording onderzoeksvragen

De twee evaluaties in de slecht gestructureerde categorie starten allebei met een inhoudsopgave, gevolgd door een inleiding. Terwijl de ene evaluatie de inleiding opent met een aanleiding van het onderzoek, gevolgd door de onderzoeksvragen en de scope van het onderzoek, start de andere evaluatie met de doel- en vraagstelling van het onderzoek. Beide evaluaties beschrijven vervolgens de gehanteerde onderzoeks aanpak, gevolgd door een leeswijzer.

Het eerste rapport volgt de inleiding op met een hoofdstuk over de achtergrond van het beleid. De andere evaluatie gaat in het tweede hoofdstuk in op het maatschappelijke doel, het stelsel, en de wetgeving omtrent het beleid.

De eerste evaluatie vervolgt het rapport met een weergave van de huidige stand van zaken, waaronder succes- en faalfactoren voor het al dan niet behalen van het doel van het beleid. De tweede evaluatie bespreekt in het derde hoofdstuk de werking van het stelsel van het beleid aan de hand van het publieke waarde model van Mark Moore.

De eerste evaluatie bespreekt vervolgens de rol van de betrokken actoren, alsmede de manier waarop andere betrokkenen die rol ervaren. De tweede evaluatie bespreekt de werking

van het beleid in de praktijk. Aanvullend worden de rollen, taken, en onderlinge verwachtingen van de betrokken actoren besproken.

In beide rapporten worden de conclusies en aanbevelingen besproken. Voor het eerste rapport is dit onderdeel het slothoofdstuk. Hierin worden de conclusies als antwoorden op de onderzoeksvragen gegeven, waarna de aanbevelingen worden gepresenteerd. De andere evaluatie bespreekt de conclusies, gevolgd door de aanbevelingen. In deze evaluatie volgen de antwoorden op de onderzoeksvragen in een apart hoofdstuk.

### 5.3 Matig gestructureerde categorie

Tabel 16 weergeeft een overzicht van de inrichting van de evaluaties in de matig gestructureerde categorie.

**Tabel 16**

*Inrichting matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Management samenvatting	Inhoud	Inhoudsopgave
Inhoudsopgave	Conclusies en aanbevelingen	Inleiding
Introductie	Inleiding	Het beleid in perspectief
Onderzoeksaanpak	De context van de TAJ	Uitvoeringspraktijk in 2022
Bevindingen: Effect van het fonds	Inrichting en positionering van de TAJ	Ervaringen met radioactief afvalbeleid
Bevindingen: Functioneren van het fonds	Taakuitvoering	Aandachtspunten voor het nieuwe Nationaal Programma
Conclusies en aanbevelingen	Effecten van het werk van de TAJ	

Van de drie evaluaties in de matig gestructureerde categorie starten twee met een inhoudsopgave. De andere evaluatie presenteert een managementsamenvatting voorafgaand aan de inhoudsopgave.

Twee van de drie evaluaties vervolgen het rapport met een introductie of inleiding. In dit hoofdstuk bespreekt één van de twee slechts de aanleiding van het onderzoek en de andere

de aanleiding, vraagstelling, onderzoeksverantwoording, en de leeswijzer. De derde evaluatie vervolgt de inhoudsopgave met een bespreking van de opgestelde conclusies en aanbevelingen, waarbij eerst de algemene conclusies worden gepresenteerd, gevolgd door deelconclusies. Vervolgens wordt een aantal lessen voor de toekomst besproken. Deze evaluatie bespreekt na de conclusies en aanbevelingen het inleidende hoofdstuk. In dit hoofdstuk worden de aanleiding, het doel, de vraag, de aanpak, en de leeswijzer besproken.

De eerste evaluatie vervolgt het rapport met een beschrijving van de bevindingen omtrent de effecten en het functioneren van het beleid. Per bevinding worden conclusies en aanbevelingen geformuleerd. De evaluatie sluit af met een hoofdstuk waarin algemene conclusies en aanbevelingen worden gegeven. Tot slot worden de onderzoeksvragen beantwoord, gevolgd door een aantal aanbevelingen per deelvraag.

De tweede evaluatie vervolgt het rapport met een beschrijving van de inrichting en positionering van het beleid, gevolgd door de taakuitvoering en de effecten. Per hoofdstuk worden deelconclusies geformuleerd.

De derde evaluatie gaat na de inleiding in op de geschiedenis van het beleid. Vervolgens wordt de uitvoeringspraktijk in 2022 geanalyseerd. In dit hoofdstuk worden ook de betrokken actoren en hun ervaringen met het beleid besproken. De evaluatie sluit af met aandachtspunten voor nieuw beleid.

#### 5.4 Ongestructureerde categorie

Tot slot, de inrichting van evaluaties in de ongestructureerde categorie (tabel 17).

**Tabel 17**  
*Inrichting ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord- Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Inhoud	Managementsamenvatting	Inhoudsopgave
Conclusies en aanbevelingen	Inhoudsopgave	Inleiding
Inleiding	Inleiding	Aanpakken en doorlopen onderzoeksproces
Opbouw PMWP	Feiten over het beleid en de sturing	Onderzoeksactiviteiten

Doeltreffendheid	Bevindingen over het beleid en de sturing	Stand van de wetenschap
Doelmatigheid	Conclusies	Scenario's met maatregelen
Opvolging aanbevelingen Tussenevaluatie en Statenmededeling	Aanbevelingen	Weging van de scenario's
		Uitvoeringsagenda
		Handreiking voor afgewogen besluitvorming

Van de drie evaluaties in deze categorie starten twee met een inhoudsopgave en één met een managementsamenvatting, gevolgd door de inhoudsopgave.

Twee van de drie evaluaties vervolgen het rapport met een inleiding. In dit hoofdstuk geeft één evaluatie de achtergrond van het beleid, gevolgd door de onderzoeksvragen en de leeswijzer. De andere evaluatie bespreekt in de inleiding de onderzoeksvragen en een leeswijzer. De derde evaluatie bespreekt aansluitend op de inhoudsopgave de conclusies en de aanbevelingen betreffende de doeltreffend- en doelmatigheid. Dit hoofdstuk wordt opgevolgd door een inleiding waarin de aanleiding, het doel en de scope, en de onderzoeks aanpak worden besproken, gevolgd door een leeswijzer.

Het eerste rapport gaat verder met een bespreking van de opbouw van het geëvalueerde beleid, gevolgd door een hoofdstuk toegewijd aan de doeltreffend- en doelmatigheid. Het rapport sluit af met een bespreking van de mate waarin eerdere aanbevelingen zijn opgevolgd.

In het tweede rapport wordt de inleiding opgevolgd met een hoofdstuk waarin feiten over het beleid en de sturing worden beschreven aan de hand van de doelen en middelen, organisatie en sturing, en resultaten en effecten van het beleid. Het hoofdstuk sluit af met feiten over de manier waarop de beleidsvoerende organisatie leert. Het volgende hoofdstuk gaat in op de bevindingen omtrent de ervaringen van actoren met het beleid en de sturing. De resultaten worden opgevolgd door conclusies betreffende het beleid in het algemeen en deelconclusies over de verschillende elementen binnen het beleid. Het rapport sluit af met een hoofdstuk waarin enkele aanbevelingen worden geformuleerd.

Het derde rapport volgt de inleiding op met een hoofdstuk waarin de aanpak en het onderzoeksproces nader worden beschreven. Dit hoofdstuk wordt opgevolgd door een beschrijving van de wetenschappelijke kennis die beschikbaar is over het beleidsprobleem. Het vierde hoofdstuk beschrijft een aantal scenario's voor het ontwikkelen van nieuw beleid,

waarna deze tegen elkaar worden afgewogen. Deze afweging vindt plaats aan de hand van een inschatting van de effectiviteit en impact, de administratieve en uitvoeringslasten, en de kosten-batenverhouding. Dit hoofdstuk wordt opgevolgd door een beschrijving van de wijzen waarop de verschillende scenario's kunnen worden gerealiseerd. Het rapport wordt afgesloten met een handreiking voor de beleidsmaker voor afgewogen besluitvorming.

## 5.5 Analyse

In de wetenschappelijke literatuur wordt onderscheid gemaakt tussen zes onderdelen van een rapportage van beleidsevaluaties: managementsamenvatting, inleiding, evaluatiedesign, resultaten, conclusie, en aanbevelingen.

De managementsamenvatting, volgens de literatuur een essentieel onderdeel om de kern van het rapport weer te geven, wordt in slechts vier rapporten gebruikt, verdeeld over de gestructureerde, matig, en ongestructureerde categorie.

Het tweede onderdeel, de inleiding, dient volgens de literatuur de aanleiding, de betrokken actoren, de afkadering van het onderzoek, en een leeswijzer te bevatten. Alle geselecteerde rapporten maken gebruik van een inleidend hoofdstuk, maar de inhoud verschilt. De aanleiding wordt in evaluaties in alle categorieën besproken. De betrokken actoren krijgen in één evaluatie in de gestructureerde categorie een plaats in de inleiding. In de overige categorieën krijgen de betrokken actoren met name een plek in het resultaten onderdeel.

Volgens de literatuur dient het evaluatiedesign in een apart hoofdstuk te worden besproken. Dit onderdeel behelst de onderzoeksvragen, methoden, en het evaluatiekader. Echter, in de praktijk wordt dit onderdeel in negen evaluaties in de gestructureerde, slecht, matig, en ongestructureerde categorie geïntegreerd in de inleiding. Terwijl drie evaluaties verspreid over de gestructureerde, matig, en ongestructureerde categorie het evaluatiedesign in een apart hoofdstuk behandelen.

Het derde onderdeel, een presentatie van de resultaten, wordt toegepast in alle geselecteerde evaluaties. Er zijn vanuit de literatuur geen verdere eisen voor dit onderdeel.

Aansluitend op de resultaten volgen de conclusies. Ten eerste, alle evaluaties in de gestructureerde en slecht gestructureerde categorie sluiten het rapport af met conclusies in combinatie met aanbevelingen. Eén rapportage in de slecht gestructureerde categorie beantwoordt de onderzoeksvragen in een apart hoofdstuk. Ten tweede, in één rapportage in de matig gestructureerde categorie worden de conclusies op het eind gepresenteerd, in combinatie met de aanbevelingen. In de andere rapportage worden de conclusies en aanbevelingen aan het



begin van het rapport gepresenteerd, op de plaats van de managementsamenvatting. De derde rapportage in de matig gestructureerde categorie maakt geen gebruik van conclusies, maar presenteert aandachtspunten voor vervolgleid. Tot slot, in de categorie met de hoogste mate van wickedness worden de conclusies en aanbevelingen in één rapport aan het begin gepresenteerd in de plaats van een managementsamenvatting en in één rapport aan het eind. In het laatste rapport wordt een handleiding voor besluitvorming gepresenteerd, in plaats van conclusies en aanbevelingen.

Er is in de literatuur geen consensus over het gebruik van aanbevelingen. Echter, in alle categorieën worden aanbevelingen gepresenteerd, in combinatie met de conclusies.

## 5.6 Conclusie

Op basis van bovenstaande analyse concludeer ik ten eerste dat de mate waarin gebruik wordt gemaakt van een managementsamenvatting niet in lijn is met de waarde die hieraan wordt toegekend vanuit de literatuur. Slechts vier rapporten maken gebruik van een onderdeel dat essentieel wordt geacht. Echter, deze evaluaties zijn wel verdeeld over drie categorieën: de gestructureerde, matig, en ongestructureerde categorie. Op basis hiervan concludeer ik dat wickedness geen rol speelt bij het hanteren van een managementsamenvatting.

Ten tweede, in alle categorieën worden de elementen van de inleiding in de inleiding besproken, behalve de betrokken actoren. Dit element kwam alleen in de gestructureerde categorie terug in de inleiding. In de andere drie categorieën worden de betrokken actoren besproken in het resultaten onderdeel. Op basis hiervan concludeer ik dat wickedness, betreffende de inleiding, enkel een rol speelt bij het bespreken van de betrokken actoren. Daarbij wordt in de literatuur onderscheid gemaakt tussen de inleiding en het evaluatiedesign. Echter, in de praktijk is het evaluatiedesign veelal onderdeel van de inleiding. De mate van wickedness speelt geen rol bij het gebruik, danwel de inrichting van het resultaten onderdeel van de rapportage.

Tot slot maakt de literatuur onderscheid tussen de conclusies en aanbevelingen, waarbij er geen consensus is over het nut en noodzaak van aanbevelingen. In de praktijk worden conclusies altijd gecombineerd met aanbevelingen, ongeacht de mate van wickedness. De mate van wickedness speelt wel een rol bij de vormgeving van dit onderdeel: in de twee categorieën met een lage mate van wickedness wordt dit onderdeel altijd op het eind van het rapport geplaatst, terwijl er in de categorieën met een hogere mate van wickedness meer variatie is: het hoofdstuk wordt zowel aan het eind als aan het begin van het rapport geplaatst, of volledig weggelaten en vervangen door een handreiking of aandachtspunten voor de opdrachtgever.

## 6. Het Evaluatieproces

Op basis van de categorisering opgesteld in hoofdstuk 4 wordt in dit hoofdstuk de derde deelvraag beantwoord: *In welke mate speelt de gradatie van wickedness een rol bij het evaluatieproces gehanteerd in de geselecteerde beleidsevaluaties?* De volledige uitwerking van onderstaande tabellen is te vinden in bijlage 3.

### 6.1 Gestructureerde categorie

De eerste categorie beleidsevaluaties die wordt geanalyseerd aan de hand van het gehanteerde evaluatieproces is de gestructureerde categorie (tabel 18).

**Tabel 18**

*Evaluatieproces gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedweld slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E-lending</i>
Doel	Doel- en vraagstelling	Doel- en vraagstelling	Afkadering
Vraagstelling	Structuur van de evaluatie	Ophalen en verwerken kwalitatieve en kwantitatieve resultaten	Ophalen en verwerken kwalitatieve en kwantitatieve resultaten
Analysekader		Beoordeling in de vorm van conclusies	Beoordeling in de vorm van conclusies
Ophalen, verwerken en analyseren kwalitatieve en kwantitatieve resultaten	Analysekader		
Toetsing bevindingen met de opdrachtgever	Ophalen kwalitatieve en resultaten		

Beoordeling in de vorm van conclusies	Verwerking en analyse van bevindingen		
---------------------------------------	---------------------------------------	--	--

De gehanteerde aanpak wordt door drie van de vier evaluaties besproken in de inleiding, de vierde evaluatie heeft er een apart hoofdstuk voor ingericht. Drie van de vier evaluaties beschrijven een doel- en vraagstelling, de vierde evaluatie hanteert een afkadering waarbinnen het onderzoek plaatsvindt.

Twee van de vier evaluaties werken met een analysekader. Eén hiervan hanteert het publieke waarde model van Mark Moore en de andere hanteert een eigen analysekader op basis van diverse aspecten van onafhankelijkheid: oordeelsvorming en inzet van instrumenten, positionering, procedures en processen, en, tot slot, oordeelsvorming in beeldvorming. Eén van de evaluaties volgt de structuur van het beleid ter evaluatie.

Vervolgens gaan alle evaluaties van start met het verzamelen van kwalitatieve data. Drie van de vier evaluaties verzamelen ook kwantitatieve data.

Eén van de vier evaluaties verwerft en analyseert de data, waarna de bevindingen worden getoetst bij de opdrachtgever alvorens een beoordeling plaatsvindt in de vorm van conclusies. De overige evaluaties verwerken en analyseren de bevindingen, waarna de evaluatie direct wordt afgesloten met een beoordeling van de bevindingen.

## 6.2 Slecht gestructureerde categorie

In tabel 19 is het evaluatieproces van de evaluaties in de slecht gestructureerde categorie weergegeven.

**Tabel 19**

*Evaluatieproces slecht gestructureerde categorie*

<i>Beleidsvaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Onderzoeksvraag en scope	Doel- en vraagstelling
Verkenning	Analyse beleidstheorie, dan wel analysekader
Verdieping	Toetsing beleidstheorie met opdrachtgever en klankbordgroep
Rapportage	Ophalen kwalitatieve data

	Analyse bevindingen aan de hand van beleidstheorie, analysekader
	Beoordeling bevindingen in de vorm van conclusies
	Toetsing bevindingen met de betrokken actoren
	Aanvullende kwalitatieve data ophalen
	Nieuwe versie van het rapport opstellen
	Nieuwe versie valideren door klankbordgroep
	Oplevering definitief rapport

Beide evaluaties starten, als onderdeel van de inleiding, met een vaststelling van de onderzoeksvraag. Naast de onderzoeksvraag stelt de ene evaluatie een scope vast en de andere een doelstelling.

De eerste evaluatie hanteert een aanpak bestaande uit drie stappen: verkenning, verdieping, en rapportage. In de verkenning- en verdiepingsfase wordt kwalitatieve en kwantitatieve data verzameld. De rapportage fase bestaat uit het verwerken, analyseren, en beoordelen van de data. Bij het beoordelen van de data is onderscheid gemaakt tussen feiten, ervaringen, en eigen beoordelingen. Het opgestelde rapport is getoetst bij een begeleidingscommissie.

De tweede evaluatie start na de vaststelling van de doel- en vraagstelling met een analyse van de beleidstheorie. Deze theorie is getoetst bij de opdrachtgever en de klankbordgroep, en fungeert als een analysekader. Vervolgens is kwalitatieve data opgehaald en geanalyseerd aan de hand van het analysekader. Een beoordeling van de data vindt plaats in de vorm van conclusies en wordt getoetst bij betrokken actoren. Op basis van de toetsing, is aanvullende data verzameld, waarna een nieuw rapport is opgesteld. De aangescherpte versie van het rapport is vervolgens gevalideerd door de klankbordgroep, waarna het rapport definitief werd opgeleverd.

### 6.3 Matig gestructureerde categorie

Een samenvatting van het gehanteerde evaluatieproces van de derde categorie beleidsevaluaties is te vinden in tabel 20.

**Tabel 20**

*Evaluatieproces matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Evaluatievragen	Doelstelling	Vraagstelling
Evaluatiekader	Vraagstelling	Oriëntatie (plan van aanpak, evaluatiekader)
Ophalen kwalitatieve en kwantitatieve data	Ophalen kwalitatieve en kwantitatieve resultaten	Gegevensverzameling
Aparte beoordeling resultaten per deelvraag in de vorm van deelconclusies (en aanbevelingen)	Toetsen resultaten in reflectiesessie	Analyse aan de hand van evaluatiekader
Beantwoorden onderzoeksvragen	Aparte beoordeling resultaten per deelvraag in de vorm van deelconclusies	Bevindingen en conclusies
Aanbevelingen	Beantwoorden onderzoeksvragen	Tussenrapportage
	Aanbevelingen	Tussentijdse bevindingen besproken en gevalideerd met begeleidingscommissie
		Expertsessie
		Aanscherping onderzoekresultaten
		Definitieve rapportage
		Validatie rapportage met begeleidingscommissie

Van de drie beleidsevaluatie beschrijven twee de gehanteerde aanpak in de inleiding. In de andere evaluatie is hier een apart hoofdstuk voor ingericht. Alle evaluaties hanteren een onderzoeksvraag, waarvan één evaluatie ook het doel van de evaluatie beschrijft.

Eén van de evaluaties is opgedeeld in afgekaderde fasen: oriëntatie, gegevensverzameling, analyse en tussenrapportage, en expertsessie en eindrapportage. De eerste fase focust op het opstellen van een plan van aanpak bestaande uit de aanpak, planning, werkafspraken, en het evaluatiekader. Eén van de overige twee evaluaties maakt ook gebruik van een evaluatiekader: het input-throughput-output-outcome-impact model.

De verzameling van kwalitatieve data komt voor in alle evaluaties. Twee van de drie evaluaties verzamelen ook kwantitatieve data.

Eén van de evaluaties volgt de dataverzamelingsfase op met het beoordelen van de resultaten per deelvraag in de vorm van deelconclusies en aanbevelingen. De tweede evaluatie toetst de opgehaalde resultaten in een reflectiesessie met betrokken actoren. Beide evaluaties beantwoorden daarna de onderzoeksvragen en stellen aanbevelingen op.

De laatste evaluatie volgt de dataverzamelingsfase op met een analyse aan de hand van het evaluatiekader, waarna bevindingen en conclusies worden opgesteld. Deze bevindingen en conclusies worden vastgelegd in een tussenrapportage en besproken en gevalideerd met een begeleidingscommissie. Aanvullend wordt een expertsessie georganiseerd op een specifiek thema van het beleid. Aan de hand van deze sessie worden de resultaten aangescherpt en vastgelegd in een definitieve rapportage. De definitieve rapportage wordt, tot slot, gevalideerd met de begeleidingscommissie.

#### 6.4 Ongestructureerde categorie

De laatste categorie die wordt geanalyseerd op basis van het gehanteerde evaluatieproces is de ongestructureerde categorie (tabel 21).

**Tabel 21**

*Evaluatieproces ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord- Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Doel, scope en onderzoeksvragen	Onderzoeksvragen	Onderzoeksvragen

Feitenscan	Analysekader	Oprichting begeleidingscommissie
Ophalen kwalitatieve data	Ophalen kwalitatieve data	Plan van aanpak
Inhoudelijke analyse	Rapporteren bevindingen en conclusies	Reflectie Plan van Aanpak door begeleidingscommissie
Voorlopige conclusies ter beantwoording onderzoeksvragen	Valideren resultaten en uitwerking aanbevelingen met betrokken ambtenaren	Ophalen kwalitatieve data
Hoor- en wederhoor met betrokken medewerkers in een reflectiesessie		Uitwerking mogelijke scenario's
Aanscherping bevindingen		Toetsing scenario's bij relevante partijen
Een externe reflectiesessie kon niet worden georganiseerd i.v.m. onderzoekstechnische beperkingen		Toetsing resultaten bij begeleidingscommissie
		Gezamenlijke analyse van scenario's met begeleidingscommissie
		Opstellen uitvoeringsagenda's
		Rapportage resultaten

Van de drie evaluaties in deze categorie maakt één gebruik van een apart hoofdstuk om de gehanteerde aanpak te beschrijven. De andere twee evaluaties beschrijven de aanpak als onderdeel van de inleiding, waarvan één hier een aparte paragraaf voor inricht.

Alle evaluaties stellen de onderzoeksvragen vast. Eén evaluatie beschrijft ook het doel en de scope.

Na het vaststellen van de onderzoeksvragen, het doel, en de scope gaat de eerste evaluatie verder met een feitenscan van documenten aangeleverd door de opdrachtgever. De

tweede evaluatie stelt een analysekader vast: het input-throughput-output-outcome-impact model. Beide evaluaties gaan vervolgens over tot het ophalen van kwalitatieve data.

De derde evaluatie gaat na het vaststellen van de onderzoeksvragen verder met het oprichten van een begeleidingscommissie en een plan van aanpak. De begeleidingscommissie krijgt de kans om te reflecteren op het plan van aanpak. Na deze stap gaat de derde evaluatie ook over op het ophalen van kwalitatieve data.

Na het verzamelen van data gaat de eerste evaluatie over op een inhoudelijke analyse, waarna een voorlopige conclusie ter beantwoording van de deelvragen wordt opgesteld. Betrokken medewerkers worden gevraagd hierop te reflecteren in een gezamenlijke sessie, waarna de bevindingen worden aangescherpt.

Na het verzamelen van kwalitatieve data gaat de tweede evaluatie over op het rapporteren en beoordelen van de bevindingen, waarna deze worden gevalideerd door betrokken ambtenaren. Tot slot wordt in een samenwerking tussen de evaluatoren en de betrokken ambtenaren een aantal aanbevelingen uitgewerkt.

De derde evaluatie werkt, na het verzamelen van data, een aantal scenario's uit die vervolgens worden getoetst bij de relevante partijen en de begeleidingscommissie. In samenwerking met de begeleidingscommissie vindt een analyse van de scenario's plaats. Aan de hand hiervan worden uitvoeringsagenda's opgesteld door de evaluatoren, waarna de resultaten worden vastgelegd in een rapportage.

## 6.5 Analyse

In de wetenschappelijke literatuur worden vier fasen van het evaluatieproces onderscheiden: structurering, dataverzameling, data-analyse, en beoordeling.

In de eerste fase, structurering, wordt het onderzoek afgebakend en worden de vervolgstappen besproken. In de gestructureerde categorie wordt het onderzoek met name afgebakend aan de hand van doel- en vraagstellingen en evaluatie onderwerpen. In de slecht gestructureerde categorie worden de evaluaties afgebakend aan de hand van een onderzoeksvraag, scope, en een doelstelling. In de matig gestructureerde categorie worden zowel evaluatievragen, en een doel- en vraagstelling gebruikt. In de laatste categorie vinden evaluaties plaats aan de hand van een doel, scope, en onderzoeksvragen. Het tweede element dat besproken dient te worden in de structurering zijn de vervolgstappen. Dit wordt in elf van de twaalf geselecteerde evaluaties besproken na de afkadering van het onderzoek. In een enkel geval, bij een evaluatie in de ongestructureerde categorie, wordt per hoofdstuk besproken welke stappen zijn ondernomen om tot de gepresenteerde resultaten te komen.



Volgens de literatuur is de tweede fase in het evaluatieproces de dataverzamelingsfase. In deze fase wordt alle informatie verzameld die nodig is om het evaluandum te analyseren en beoordelen. In de gestructureerde categorie wordt de dataverzameling, in overeenstemming met de literatuur, gehanteerd als tweede fase in het evaluatieproces. Hierbij wordt zowel bestaande als nieuwe data verzameld. Alle evaluaties in deze categorie maken gebruik van kwalitatieve data. Drie van de vier maken ook gebruik van kwantitatieve data. In de slecht gestructureerde categorie wordt in één geval de dataverzameling gehanteerd als tweede fase en in de andere evaluatie zowel als vierde en als achtste fase. Bij één van deze evaluaties wordt alleen kwalitatieve data verzameld, bij de andere zowel kwalitatieve als kwantitatieve data. In beide evaluaties wordt gebruik gemaakt van bestaande en nieuwe data. In de matig gestructureerde categorie is de dataverzameling de derde fase in het evaluatieproces. Alle evaluaties maken gebruik van bestaande en nieuwe data. Twee van deze evaluaties verzamelen zowel kwalitatieve als kwantitatieve data. De derde evaluatie verzamelt enkel kwalitatieve data. In de ongestructureerde categorie, wordt de dataverzameling in twee evaluaties gehanteerd als derde fase en een enkele keer als vijfde fase. Alle evaluaties in deze categorie verzamelen kwalitatieve data en maken gebruik van zowel bestaande als nieuwe data.

De derde fase in het evaluatieproces is de analysefase. In deze fase wordt data geanalyseerd aan de hand van evaluatievragen of -criteria. In de gestructureerde categorie wordt opgehaalde data bij twee van de vier evaluaties geanalyseerd. In beide gevallen is deze fase samengevoegd met een andere fase van het proces: één keer met de dataverzamelingsfase en één keer met de beoordelingsfase. Deze analyses vinden plaats aan de hand van een evaluatiekader: Mark Moore's publieke waarde model en een zelf opgesteld kader. De andere twee evaluaties in deze categorie verwerken alleen de opgehaalde data in de rapportage. In de slecht gestructureerde categorie vindt de analysefase, samengevoegd met de beoordelingsfase, plaats na de dataverzamelingsfase. Er is in het rapport geen analysekader opgenomen. De andere evaluatie in deze categorie hanteert twee analysefasen. De eerste analysefase focust op de beleidstheorie, aan de hand waarvan een analysekader wordt opgesteld. De tweede analysefase vindt plaats na een toetsing van de opgehaalde data met de opdrachtgever en klankbordgroep. In de matig gestructureerde categorie vindt in alle gevallen een aparte analysefase plaats. In twee evaluaties vindt deze fase plaats direct na de dataverzamelingsfase, bij de derde evaluatie wordt de opgehaalde data eerst getoetst in een reflectiesessie. Twee evaluaties werken met een evaluatiekader, bij de derde evaluatie fungeren de onderzoeksvragen als evaluatievragen. In de ongestructureerde categorie wordt de analysefase in twee evaluaties gecombineerd met een andere fase en in één evaluatie staat de analyse op zichzelf. Eén evaluatie

maakt gebruik van een evaluatiekader, een andere evaluatie hanteert evaluatievragen. In twee evaluaties volgt de analysefase op de dataverzamelingsfase. In één evaluatie vindt de analyse plaats in samenspraak met de begeleidingscommissie, nadat de opgehaalde resultaten zijn uitgewerkt en getoetst.

De laatste fase die in de literatuur wordt onderscheiden is de beoordelingsfase. In deze fase wordt een waardeoordeel geveld over het evaluandum. In de gestructureerde categorie is dit de laatste fase in het evaluatieproces. In drie evaluaties vindt deze fase plaats na de analysefase en bij de vierde evaluatie na een toetsing van de analyse. In de slecht gestructureerde categorie wordt de fase in één evaluatie samengevoegd met de analysefase. In de andere evaluatie vindt deze fase op zichzelf plaats. In beide gevallen wordt de beoordeling getoetst. Bij de ene evaluatie wordt de rapportage op basis hiervan aangepast. Bij de andere evaluatie is het onduidelijk welk gevolg de toetsing heeft voor de inhoud van de rapportage. In de matig gestructureerde categorie vindt de beoordelingsfase bij twee evaluaties plaats nadat de opgehaalde data is getoetst. Bij de derde evaluatie worden de bevindingen na de beoordeling getoetst. In de ongestructureerde categorie wordt de beoordelingsfase in twee evaluaties uitgevoerd na de analysefase, waarna een reflectie- of validatiesessie plaatsvindt. Bij de derde evaluatie wordt de data zowel na de dataverzamelingsfase, als na een toetsing en analyse met een begeleidingsgroep beoordeeld.

## 6.6 Conclusie

Op basis van bovenstaande analyse concludeer ik dat de mate van wickedness geen rol speelt bij de structurering van het evaluatieonderzoek. In alle categorieën wordt het onderzoek afgebakend en de vervolgstappen besproken.

Met betrekking tot de tweede fase, de dataverzameling, concludeer ik, ten eerste, dat de mate van wickedness geen rol speelt bij het hanteren van deze fase, dit gebeurt in alle categorieën. Ten tweede concludeer ik dat de mate van wickedness wel een rol speelt bij het moment waarop deze fase plaatsvindt in het evaluatieproces. In de gestructureerde categorie is de dataverzamelingsfase de tweede stap, in lijn met de literatuur. Naarmate de gradatie van wickedness toeneemt, vindt deze fase verder in het proces plaats. Ten derde concludeer ik dat de mate van wickedness geen rol speelt bij het gebruiken van bestaande of nieuwe data: alle evaluaties maken gebruik van beide. Tot slot concludeer ik met betrekking tot de dataverzamelingsfase dat de mate van wickedness wel een rol speelt bij het type data dat wordt gebruikt: enkel in de ongestructureerde categorie worden evaluaties alleen gebaseerd op

kwalitatieve data, in alle andere categorieën wordt gebruik gemaakt van kwalitatieve en kwantitatieve data.

Met betrekking tot de analysefase concludeer ik dat de mate van wickedness geen rol speelt betreft de aanwezigheid van de analysefase: deze fase komt voor in alle categorieën. Ten tweede speelt de mate van wickedness geen rol bij het samenvoegen van de analysefase met een andere fase van het evaluatieproces: dit gebeurt in alle categorieën. Echter, de mate van wickedness speelt wel een rol bij het hanteren van de analysefase als een aparte fase. Dit gebeurt namelijk alleen in de drie categorieën met een hogere mate van wickedness. Met betrekking tot het moment in het proces waarop de analysefase plaatsvindt, concludeer ik dat de mate van wickedness een rol speelt. In de drie categorieën met een hogere mate van wickedness vindt een toetsfase plaats voorafgaand aan de analysefase. In de gestructureerde categorie vindt de analysefase altijd plaats na de dataverzamelingsfase. Met betrekking tot het hanteren van een analysekader concludeer ik dat de mate van wickedness geen rol speelt: dit komt voor in alle categorieën. Echter, de mate van wickedness speelt wel een rol bij het gebruik van evaluatievragen: dit komt alleen voor in de twee categorieën met de hoogste mate van wickedness. Tot slot speelt de mate van wickedness een rol bij het betrekken van externen in de analysefase: dit gebeurt enkel in de ongestructureerde categorie.

De laatste fase die in de literatuur naar voren komt, is de beoordelingsfase. Op basis van bovenstaande analyse concludeer ik dat de mate van wickedness geen rol speelt bij het hanteren van de beoordelingsfase: dit gebeurt in alle categorieën. De mate van wickedness speelt wel een rol bij het moment waarop deze fase plaatsvindt in het evaluatieproces. In de literatuur wordt gesteld dat de beoordelingsfase de laatste fase is, echter, in de drie categorieën met een hogere mate van wickedness wordt de beoordeling getoetst, waarna er een wijziging van de rapportage plaats kan vinden.

## 7. De Instrumenten

In dit hoofdstuk wordt de vierde deelvraag beantwoord: *In welke mate speelt de gradatie van wickedness een rol bij de instrumenten gehanteerd in de geselecteerde beleidsevaluaties?* Een meer uitgebreide versie van onderstaande tabellen is te vinden in bijlage 4.

### 7.1 Gestructureerde categorie

De eerste categorie die wordt geanalyseerd aan de hand van de gehanteerde instrumenten is de gestructureerde categorie (tabel 22).

**Tabel 22**

*Instrumenten gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedwelmd slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E-lending</i>
Deskresearch	Deskresearch	Deskresearch	Documentenanalyse
Casussen	Individuele interviews	Gesprekken	Interviews
(groeps)Interviews	Groepsinterviews	Praktijkbezoeken	
Enquête			
Begeleidingscommissie			
Schriftelijke input			

Alle evaluaties in deze categorie maken gebruik van een vorm van deskresearch. Twee van de vier evaluaties gebruiken dit instrument om een eerste beeld te creëren van het beleid. Drie van die vier evaluaties benoemen ook welke documenten zijn gebruikt ter analyse. Dit gaat met name om jaarverslagen, openbare- en interne documenten van de opdrachtgever en/of publieke instellingen zoals toezichthouders, en eerder uitgevoerde evaluaties.

Naast de documentenanalyse maken alle evaluaties gebruik van (groeps)interviews. In alle evaluaties worden individuele gesprekken gevoerd. Twee van de vier evaluaties maken ook gebruik van groepsgesprekken.

Twee van de vier evaluaties maken gebruik van aanvullende instrumenten. De ene evaluatie maakt gebruik van zes casussen om inzicht te krijgen in de werkwijze van het beleid. Daarnaast wordt gebruik gemaakt van een enquête en krijgen organisaties de mogelijkheid om schriftelijke input te leveren voor het onderzoek. Tot slot maakt deze evaluatie gebruik van een begeleidingscommissie. De andere evaluatie maakt naast deskresearch en interviews gebruik van praktijkbezoeken aan uitvoerende organisaties om de werking van het beleid in de praktijk te zien.

## 7.2 Slecht gestructureerde categorie

In tabel 23 is de tweede categorie beleidsevaluaties weergegeven die wordt geanalyseerd aan de hand van de gehanteerde instrumenten.

**Tabel 23**

*Instrumenten slecht gestructureerde categorie*

<i>Beleidsevaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Documentenanalyse	Documentenstudie
Verkennde gesprekken	Interviews
Online informatie uitvraag bij lokale omroepen	Diepte-interviews
Telefonische interviews met gemeentebambtenaren	Klankbordgroep
Casusanalyse	Begeleidingscommissie
	Aanvullende interviews

Beide evaluaties maken gebruik van een documentenanalyse. Waarbij de één dit gebruikt om een beeld te creëren van de huidige stand van zaken, gebruikt de andere het om de beleidstheorie te analyseren.

Daarnaast maken beide evaluaties gebruik van interviews. De eerste evaluatie maakt een onderscheid tussen verkennende gesprekken en telefonische interviews. De andere evaluatie maakt een onderscheid tussen interviews aan het begin van het proces ter analyse van de beleidstheorie, diepte-interviews om de werking van het beleid in de praktijk in kaart te brengen, en aanvullende interviews aan het einde van het evaluatieproces.

Aanvullend op de documentenanalyse en interviews, maakt de eerste evaluatie ook gebruik van een online informatie uitvraag en een analyse van zes casussen om een verdiepend beeld te geven van het beleid. De andere evaluatie maakt naast een documentenanalyse en interviews gebruik van een klankbordgroep en een begeleidingscommissie.

### 7.3 Matig gestructureerde categorie

De derde categorie die wordt geanalyseerd op basis van de gehanteerde instrumenten is de matig gestructureerde categorie (tabel 24).

**Tabel 24**

*Instrumenten matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Documentstudie	Documentstudie	Documentstudie
Verdiepende interviews	Interviews	Diepte-interviews
Validatie	Reflectiesessie	Begeleidingscommissie
	Begeleidingscommissie	Expertsessie

Alle evaluaties in deze categorie starten met een documentenanalyse van bestaande onderzoeksrapporten, jaarrapportages, en beleidsdocumenten.

Aanvullend op de documentenanalyse maken de evaluaties gebruik van verdiepende/diepte interviews met ministeries, gemeenten, en andere betrokken organisaties.

Ten derde maken alle evaluaties gebruik van een vorm van validatie. De eerste evaluatie doet dit door middel van aanvullende gesprekken met de geïnterviewde personen. De tweede evaluatie middels een reflectiesessie met organisaties die indirect betrokken zijn bij het geëvalueerde beleid. De derde evaluatie valideert de bevindingen via een begeleidingscommissie.

De derde evaluatie maakt, aanvullend op bovenstaande methoden, gebruik van een expertsessie ter verdieping op een specifiek thema van het geëvalueerde beleid.

## 7.4 Ongestructureerde categorie

De laatste categorie van beleidsevaluaties die wordt geanalyseerd op basis van de gehanteerde instrumenten is de ongestructureerde categorie (tabel 25).

**Tabel 25**

*Instrumenten ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade</i>
Feitenscan	Bureaustudie	Vastleggen feitelijke stand van zaken drie experts
Vragenlijsten	(groeps)Interviews	Groepsinterviews
Interviews	Sessies met betrokken ambtenaren	Begeleidingscommissie
Interne reflectiesessie (Hoor- en wederhoor van bevindingen)		

Twee van de bovenstaande drie evaluaties starten met een vorm van documentenanalyse. De eerste evaluatie voert een feitenscan uit op aangeleverde documenten. De tweede evaluatie verzamelt relevante sturingsinformatie op basis van beleidsdocumenten. De derde evaluatie start, in samenwerking met drie experts, met het vastleggen van de feitelijke stand van zaken.

Als tweede methode wordt bij twee van de drie evaluaties gebruik gemaakt van (groeps)interviews. In het geval van de ene evaluatie worden zowel interne als externe betrokkenen geïnterviewd. Voor de andere evaluatie worden gesprekken gevoerd met de betrokkenen van het beleid en organisaties die verantwoordelijk zijn voor de handhaving van het beleid. De derde evaluatie maakt gebruik van vragenlijsten voorafgaand aan de interviews.

Tot slot maken alle evaluaties gebruik van een vorm van validatie. De eerste evaluatie maakt gebruik van een interne reflectiesessie om bevindingen te toetsen. In de tweede evaluatie worden de resultaten gevalideerd met betrokken ambtenaren, waarna de evaluatoren gezamenlijk met de betrokken ambtenaren aanbevelingen uitwerken. De derde evaluatie toetst bevindingen en weegt verschillende scenario's af in samenspraak met een begeleidingscommissie.

## 7.5 Analyse

In de literatuur worden de documentenanalyse, het interview, de focusgroep, het expertpanel, en de workshop onderscheiden als evaluatie-instrumenten voor externe evaluatoren.

De documentenanalyse en interviews worden toegepast in alle categorieën, terwijl de focusgroep alleen wordt toegepast in de ongestructureerde categorie. Het expertpanel wordt toegepast in de matig en ongestructureerde categorie. In geen van de geselecteerde evaluaties wordt gebruik gemaakt van een workshop.

Aanvullend op de instrumenten die in de literatuur worden onderscheiden, kunnen elf instrumenten uit de geselecteerde evaluaties worden gedestilleerd: schriftelijke input, praktijkbezoek, verkennend gesprek, online informatie uitvraag, klankbordgroep, validatiesessie, feiten scan, casusanalyse, enquête, reflectiesessie, en begeleidingscommissie. In de gestructureerde categorie wordt, naast de data-analyse en interviews, ook gebruik gemaakt van een casusanalyse, enquête, begeleidingscommissie, schriftelijke input, en praktijkbezoek. In de slecht gestructureerde categorie wordt gebruik gemaakt van een casusanalyse, begeleidingscommissie, verkennend gesprek, online uitvraag, en een klankbordgroep. In de matig gestructureerde categorie wordt gebruik gemaakt van een begeleidingscommissie, validatiesessie, en reflectiesessie, naast de instrumenten die uit de literatuur naar voren komen. In de ongestructureerde categorie worden een enquête, begeleidingscommissie, reflectiesessie, en feiten scan als aanvullende instrumenten gebruikt.

## 7.6 Conclusie

Met betrekking tot de instrumenten die naar voren komen in de literatuur concludeer ik ten eerste dat de mate van wickedness geen rol speelt bij uitvoeren van een documentenanalyse en interviews: beide worden toegepast in alle categorieën. Ten tweede concludeer ik dat de rol van wickedness wel een rol speelt bij de focusgroep en het expert panel. De focusgroep wordt enkel toegepast in de slecht gestructureerde categorie. Het expert panel wordt gebruikt in de twee categorieën met de hoogste mate van wickedness.

Met betrekking tot de aanvullend geïdentificeerde instrumenten concludeer ik ten eerste dat de mate van wickedness een rol speelt bij de schriftelijke input en het praktijkbezoek: deze instrumenten komen alleen voor in de gestructureerde categorie. Ten tweede speelt de mate van wickedness een rol bij het verkennende gesprek, de online informatie uitvraag, en de klankbordgroep. Deze instrumenten worden enkel toegepast in de slecht gestructureerde categorie. Ten derde, speelt de mate van wickedness een rol bij de validatiesessie: dit wordt enkel gebruikt in de matig gestructureerde categorie. Ten vierde, de feiten scan wordt



alleen gebruik in de ongestructureerde categorie. Op basis hiervan concludeer ik dat de mate van wickedness een rol speelt bij deze instrumenten. Ten vijfde, de mate van wickedness speelt een rol bij de casusanalyse. Dit instrument wordt enkel toegepast in de twee categorieën met de laagste mate van wickedness. De enquête wordt toegepast in zowel de gestructureerde als de ongestructureerde categorie. Op basis hiervan concludeer ik dat de mate van wickedness een rol speelt bij de toepassing van de enquête: dit gebeurt enkel bij de categorieën op de uitersten van de schaal van wickedness. De reflectiesessies worden toegepast in de matig en ongestructureerde categorie, de twee categorieën met de hoogste mate van wickedness. Op basis hiervan concludeer ik dat wickedness een rol speelt bij het hanteren van reflectiesessies. Tot slot, de mate van wickedness speelt geen rol bij het gebruik van de begeleidingscommissie, dit wordt toegepast in alle categorieën.

## 8. Doelmatig- & Doeltreffendheid

In dit hoofdstuk beantwoord ik de laatste deelvraag: *In welke mate speelt de gradatie van wickedness en rol bij de mate waarin de geselecteerde beleidsevaluaties een uitspraak doen over de doelmatig- en doeltreffendheid van het beleid?* Wederom vormt de categorisering uit hoofdstuk 4 de basis voor het beantwoorden van de deelvraag. Onderstaande tabellen geven een samengevatte versie weer van bijlage 5.

### 8.1 Gestructureerde categorie

Allereerst wordt de gestructureerde categorie geanalyseerd aan de hand van de mate waarin de evaluaties een uitspraak doen over de doelmatig- en doeltreffendheid van het beleid (tabel 26).

**Tabel 26**

*Doelmatig- en doeltreffendheid gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedweld slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E-lending</i>
Doelmatigheid			
	De Aanwijzingen zijn als doelmatig te karakteriseren.	De grootste kosten [...] zijn de kosten voor het permanent toezicht [...] en worden als doelmatigheid beschouwd.	
Doeltreffendheid			
Onze conclusie is dat de publieke waarde van de Gezondheidsraad is gelegen in advisering [...].	[...]de Aanwijzingen spelen een belangrijke rol in het waarborgen van onafhankelijk opererende rijksinspecties.	Het Convenant heeft geleid tot een aantoonbare verbetering van het dierenwelzijn in de praktijk [...]	

	<p>Toch is het niet vanzelfsprekend dat rijksinspecties voldoende onafhankelijk te werk kunnen gaan [...]</p>		
--	---	--	--

Van de vier evaluaties besteden er twee aandacht aan de doelmatigheid van het geëvalueerde beleid. In beide gevallen wordt beargumenteerd dat het gevoerde beleid als doelmatig kan worden beschouwd.

Wat betreft de doeltreffendheid doen drie van de vier evaluaties een uitspraak. Eén van de evaluaties stelt dat het beleid doeltreffend is geweest. De tweede evaluatie stelt dat het beleid bijdraagt aan het gestelde doel, maar plaatst hier ook kanttekeningen bij. De derde evaluatie beschrijft de wijze waarop het beleid bijdraagt aan het doel, maar niet in welke mate dit gebeurt.

## 8.2 Slecht gestructureerde categorie

In tabel 27 is de slecht gestructureerde categorie geanalyseerd met betrekking tot de mate waarin een uitspraak wordt gedaan over de doelmatig- en doeltreffendheid.

**Tabel 27**

*Doelmatig- en doeltreffendheid slecht gestructureerde categorie*

<i>Beleidsvaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
<b>Doelmatigheid</b>	
<p>Streekomroepvorming is een middel dat bijdraagt aan het haalbaar en betaalbaar maken van een LTMA in Nederlandse gemeenten [...].</p> <p>[...] streekomroepvorming alleen is niet voldoende voor het behalen van de kwaliteitsdoelstellingen. Het grootste</p>	<p>Het proces bij maatschappij-ontwrichtend weer of calamiteiten kan verbeterd worden om de publieke veiligheid bij extreem weer beter te borgen.</p>

knelpunt voor het kunnen verzorgen van een LTMA is de beperkte financiering van lokale publieke omroepen.	
Doeltreffendheid	
	Wij concluderen dat de Wtms in de praktijk een doeltreffend instrument is en bijdraagt aan de realisatie van de beoogde publieke waarde [...]

Beide evaluaties besteden aandacht aan de doelmatigheid van het beleid. De eerste evaluatie stelt dat het beleid doelmatig is, maar plaatst kanttekeningen bij de vraag of het gevoerde beleid de enige manier is waarop het doel behaald kan worden. De tweede evaluatie doet geen uitspraak over de doelmatigheid en stelt alleen dat het verbeterd kan worden.

Eén van de twee evaluaties besteedt aandacht aan de doeltreffendheid. Deze evaluatie stelt dat het beleid doeltreffend is. De evaluatie onderbouwt de uitspraak, maar plaats ook kanttekeningen: het beleid zou nog doeltreffender kunnen zijn.

### 8.3 Matig gestructureerde categorie

Een samenvatting van de derde categorie met betrekking tot de mate waarin de evaluaties een uitspraak doen over de doelmatig- en doeltreffendheid is te vinden in tabel 28.

**Tabel 28**

*Doelmatig- en doeltreffendheid matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Doelmatigheid		
	De TAJ heeft [...] doelmatig gefunctioneerd.  In bredere zin zijn er echter wel kanttekeningen te plaatsen bij de [...]	

	doelmatigheid van de activiteiten van de TAJ.	
Doeltreffendheid		
[...] concluderen wij dat het NEF [...] over het algemeen doeltreffend is geweest, maar sommige fases (m.n. de opstart) en activiteiten (o.a. scholen en laadpalen) minder doeltreffend zijn gebleken.	De TAJ is effectief geweest in het realiseren van zorgcontinuïteit.  In bredere zin zijn er echter wel kanttekeningen te plaatsen bij de effectiviteit [...] van de activiteiten van de TAJ.	In conclusie kan gesteld worden dat het beleidskader effectief en overzichtelijk is ingericht.

Eén van de drie evaluaties doet een uitspraak over de doelmatigheid. De evaluatie stelt dat het beleid, met kanttekeningen, doelmatig is. Van de overige twee evaluaties stelt één evaluatie het onderzoeken van de doelmatigheid wel als doel, maar doet er verder geen uitspraak over.

Alle evaluaties in deze categorie doen een uitspraak over de doeltreffendheid van het beleid, namelijk dat het beleid doeltreffend is. In twee gevallen worden er kanttekeningen bij deze conclusie geplaatst.

#### 8.4 Ongestructureerde categorie

Tot slot worden evaluaties in de ongestructureerde categorie geanalyseerd op basis van de mate waarin zij een uitspraak doen over de doelmatig- en doeltreffendheid (tabel 29).

**Tabel 29**

*Doelmatig- en doeltreffendheid ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Doelmatigheid		
Wij kunnen geen conclusie trekken over de vraag of het PMWP doelmatig is geweest.	[...] De mate waarin deze inzet doelmatig is geweest, kan niet afdoende worden	

	beoordeeld zonder verdere operationalisering van ambities, doelen, resultaten en activiteiten en de samenhang daartussen.	
Doeltreffendheid		
Wij concluderen dan ook dat het aannemelijk is dat het PMWP in beperkte mate doeltreffend is geweest [...]	Op deze vraag is geen overall conclusie te trekken.  [...] het is ook duidelijk dat de onzekerheden en afhankelijkheden bij deze transitie zo groot zijn, en de gestelde termijn zo lang, dat een harde uitspraak over doeltreffendheid altijd lastig zal zijn.	

Van de drie evaluaties besteden twee aandacht aan de doelmatig- en doeltreffendheid van het geëvalueerde beleid.

Betreffende de doelmatigheid stellen beide evaluaties dat het niet mogelijk is een uitspraak te doen. Eén van de evaluaties wijdt dit aan het gebrek aan operationalisering van de doelen die het beleid beoogt na te streven.

Betreffende de doeltreffendheid wordt in één evaluatie geconcludeerd dat het aannemelijk is dat het beleid in beperkte mate doeltreffend is geweest. De andere evaluatie trekt geen conclusie en stelt dat het probleem zodanig onzeker, afhankelijk, en langdurig is dat een uitspraak doen over doeltreffendheid altijd lastig zal zijn.

## 8.5 Analyse

In de gestructureerde categorie wordt in twee van de vier evaluaties aandacht besteed aan de doelmatigheid. In beide gevallen wordt geconcludeerd dat het beleid doelmatig is. In beide evaluaties in de slecht gestructureerde categorie wordt het evalueren van de doelmatigheid als doel gesteld. In één evaluatie wordt geconcludeerd dat het beleid doelmatig is, terwijl de andere

evaluatie geen conclusie trekt. In de matig gestructureerde categorie wordt in twee van de drie evaluaties aandacht besteed aan de doelmatigheid. In één van deze evaluaties wordt geconcludeerd dat het beleid doelmatig is, maar dat er ook kanttekeningen kunnen worden geplaatst bij deze conclusie. In de andere evaluatie wordt geen conclusie getrokken over de doelmatigheid, ondanks dat dit wel een doel is van de evaluatie. In de laatste categorie, de ongestructureerde categorie, wordt in twee van de drie evaluaties aandacht besteed aan de doelmatigheid. In beide evaluaties wordt geconcludeerd dat het niet mogelijk is om vast te stellen of het beleid doelmatig is. In één evaluatie wordt aangegeven dat dit niet mogelijk is omdat het beleid geen duidelijke operationalisering hanteert.

De doeltreffendheid wordt besproken in drie van de vier evaluaties in de gestructureerde categorie. Twee van de drie concluderen dat het beleid doeltreffend is, waarvan er één kanttekeningen plaats bij deze conclusie. De derde evaluatie doet geen uitspraak over de mate waarin het beleid doeltreffend is. In de slecht gestructureerde categorie besteedt één evaluatie aandacht aan de doeltreffendheid. Hierbij wordt geconcludeerd dat het beleid doeltreffend is, maar er worden ook kanttekeningen bij deze constatering geplaatst. Alle evaluaties in de matig gestructureerde categorie besteden aandacht aan de doeltreffendheid en concluderen dat het beleid doeltreffend is, waarvan twee evaluaties ook kanttekeningen plaatsen bij deze conclusie. In de ongestructureerde categorie besteden twee van de drie evaluaties aandacht aan de doeltreffendheid. In beide gevallen wordt er geen conclusie getrokken over de doeltreffendheid. Bij één evaluatie wordt beargumenteerd dat het niet mogelijk is om een conclusie te trekken over de doeltreffendheid omdat het probleem te onzeker, afhankelijk, en langdurig is.

## 8.6 Conclusie

Op basis van bovenstaande analyse concludeer ik dat de mate van wickedness geen rol speelt bij het ten doel stellen van het evalueren van de doelmatigheid: in alle categorieën wordt de doelmatigheid van het beleid geëvalueerd. Ten tweede concludeer ik dat de mate van wickedness wel een rol speelt bij de mate waarin een uitspraak wordt gedaan over de doelmatigheid: enkel in de gestructureerde categorie wordt altijd een conclusie getrokken over de doelmatigheid van het beleid wanneer dit een doel van de evaluatie is. Daarbij speelt wickedness een rol bij de erkenning van de onmogelijkheid conclusies te trekken over de doelmatigheid. Het evalueren van de doelmatigheid wordt in de drie gradaties met een hogere mate van wickedness als doel gesteld, ondanks dat hier niet in alle gevallen een conclusie over

wordt getrokken. Alleen in de ongestructureerde categorie wordt specifiek aangegeven dat het inderdaad niet mogelijk is een uitspraak te doen over de doelmatigheid en waarom.

Met betrekking tot de doeltreffendheid concludeer ik dat de mate van wickedness geen rol speelt bij het ten doel stellen de doeltreffendheid te evalueren: dit gebeurt in elke categorie. Daarbij speelt de mate van wickedness geen rol bij het trekken van conclusies over de doeltreffendheid: dit gebeurt in alle categorieën. Echter, de mate van wickedness speelt wel een rol bij de inhoud van deze conclusie. In de gestructureerde, slecht, en matig gestructureerde categorie wordt geconcludeerd dat het beleid, eventueel met kanttekeningen, doeltreffend is. Enkel in de ongestructureerde categorie wordt geconcludeerd dat het niet mogelijk is om vast te stellen of het beleid doeltreffend is.



## 9. Conclusie

De hoofdvraag van dit onderzoek is: *In welke mate speelt de gradatie van wickedness van problemen een rol bij beleidsevaluaties uitgevoerd door externe, commerciële adviesbureaus?*

Om deze vraag te beantwoorden heb ik twaalf beleidsevaluaties geselecteerd en ingedeeld in de gestructureerde, slecht, matig, of ongestructureerde categorie. Vervolgens heb ik vier onderdelen van deze evaluaties onderzocht: de inrichting, het evaluatieproces, de gehanteerde instrumenten, en de mate waarin een uitspraak wordt gedaan over de doelmatig- en doeltreffendheid van het beleid. Op basis hiervan concludeer ik dat de mate van wickedness een rol speelt bij al deze elementen. Dit antwoord licht ik verder toe in onderstaande alinea's.

Betreffende de inrichting speelt de mate van wickedness een rol bij de plek die betrokken actoren krijgen in de rapportage van de evaluatie. In de gestructureerde categorie worden de betrokken actoren alleen in de inleiding besproken. Terwijl de categorieën met een hogere mate van wickedness dit een onderdeel maken van het resultatenonderdeel. Dit komt overeen met de literatuur: bij problemen met een hoge mate van wickedness spelen de betrokken actoren een belangrijke rol. De mate van wickedness speelt ook een rol bij het bespreken van het leerproces van de beleidsvoerende organisatie: dit gebeurt enkel in de ongestructureerde categorie. Daarbij speelt de mate van wickedness een rol bij de plaats en vormgeving van de conclusies en aanbevelingen. In de twee categorieën met de hoogste mate van wickedness worden de conclusies vervangen door een alternatief afsluitend hoofdstuk, zoals een overzicht van aandachtspunten of een handleiding. Daarbij kunnen de conclusies en aanbevelingen in deze categorieën zowel aan het eind als aan het begin van een rapportage worden geplaatst, terwijl de conclusies en aanbevelingen in de twee categorieën met een lagere mate van wickedness altijd op het eind worden geplaatst.

Met betrekking tot het evaluatieproces speelt de mate van wickedness een rol bij het moment waarop de dataverzameling-, data-analyse, en de beoordeling plaatsvinden. Hoe hoger de mate van wickedness, hoe meer variatie in het moment en de wederkerigheid waarin deze fasen aan bod komen. Ten tweede wordt er bij de hoogste mate van wickedness enkel gebruik gemaakt van kwalitatieve data, terwijl er bij de drie categorieën met een lagere mate van wickedness een combinatie van kwalitatieve en kwantitatieve data wordt gebruikt. Ten derde speelt de mate van wickedness een rol bij het hanteren van evaluatievragen in plaats van een evaluatiekader: dit gebeurt enkel in de matig en ongestructureerde categorie. Ten vierde worden betrokken actoren alleen meegenomen in de analyse in de categorie met de hoogste mate van

wickedness. Tot slot kan er een aanvullende fase worden geïdentificeerd: de toetsfase. De mate van wickedness speelt een rol bij de frequentie waarin deze fase wordt toegepast: hoe meer wicked het probleem, hoe vaker een toetsfase plaatsvindt. Het toetsen van de beoordeling van de data gebeurt alleen in de drie categorieën met een hogere mate van wickedness. In deze categorieën komt met wisselende mate voor dat wordt erkend dat de toetsing invloed heeft op de inhoud van de uiteindelijke rapportage. Gelinkt aan deze bevinding is het kritiekpunt dat beleidsevaluaties uitgevoerd door externen onafhankelijkheid missen. Echter, uit dit onderzoek kan niet worden opgemaakt in welke mate de toetsing effect heeft gehad op de inhoud van de rapportage.

Betreffende de gehanteerde instrumenten speelt de mate van wickedness een rol bij het gebruik van praktijkbezoeken, deze worden enkel ingezet in de gestructureerde categorie. Ten tweede worden klankbordgroepen alleen gebruikt in de slecht gestructureerde categorie en validatiesessies alleen in de matig gestructureerde categorie. De mate van wickedness speelt ook een rol bij het gebruik van de feitenscan, dit wordt enkel gebruikt in de ongestructureerde categorie. Tot slot speelt de mate van wickedness een rol bij reflectiesessies en expertpanels, deze instrumenten worden enkel toegepast in de twee categorieën met de hoogste mate van wickedness.

Tot slot speelt de mate van wickedness een rol bij het vaststellen van de doelmatig- en doeltreffendheid. In de gestructureerde categorie wordt altijd een oordeel geveld over de doelmatig- en doeltreffendheid wanneer dit als doel wordt gesteld. Terwijl evaluaties in de slecht en matig gestructureerde categorie het evalueren van de doelmatig- en doeltreffendheid wel als doel stellen, maar er uitspraak over doen. In de ongestructureerde categorie wordt geconcludeerd dat het niet mogelijk is om de doelmatig- en doeltreffendheid vast te stellen, omdat het probleem dat het beleid beoogd aan te pakken dusdanig onzeker en afhankelijk is. Deze conclusie sluit aan bij de literatuur: het beoordelen van de mate waarin doelen efficiënt en effectief zijn behaald, is enkel geschikt voor problemen die gestructureerd zijn. Desondanks wordt het evalueren hiervan alsnog als doel gesteld bij problemen met een hogere mate van wickedness.

## 10. Discussie

Er zijn een aantal elementen die ik, naar aanleiding van dit onderzoek, wil aanstippen.

Ten eerste is er op het gebied van externe, commerciële adviesbureaus nauwelijks wetenschappelijke literatuur beschikbaar. Terwijl dit concept een cruciaal onderdeel is van dit onderzoek, kan het maar beperkt wetenschappelijk worden onderbouwd.

Ten tweede, op het gebied van wicked problems is een ruime hoeveelheid literatuur beschikbaar. Echter, bij het toetsen van de literatuur komt naar voren dat de realiteit dusdanig complex is dat deze literatuur niet volledig aansluit. Het gevolg hiervan is dat de basis van dit onderzoek – de categorisering van beleidsevaluaties op een schaal van wickedness – gedeeltelijk is gestoeld op mijn eigen interpretatie van de beschikbare data en literatuur. Dit heeft, zoals vermeld in de methodologie, gevolgen voor de herhaalbaarheid van dit onderzoek. Daarbij wordt het uit de literatuur niet duidelijk wat de precieze rol is van actoren in de slecht en matig gestructureerde categorie. Daarom is dit onderdeel van ondergeschikte waarde geweest ten opzichte van de twee andere dimensies: kennis en consensus.

Ten derde heb ik, als gevolg van de beperkte beschikbare tijd en middelen, in dit onderzoek slechts een aantal belangrijke onderdelen van beleidsevaluaties onderzocht. Zoals te zien in het theoretisch kader bestaan beleidsevaluaties uit meer onderdelen zoals het evaluandum, het evaluatiemotief, en het moment in de beleidscyclus, bijvoorbeeld.

Ten vierde heb ik ondervonden dat de literatuur over beleidsevaluaties slechts in beperkte mate aansluit op hogere categorieën van wickedness. De basiselementen van de inrichting van rapportages sluiten goed aan bij de categorieën met een relatief lage mate van wickedness. Echter, bij een hoge mate van wickedness wordt een meer uiteenlopende indeling toegepast. Dit is ook het geval bij het evaluatieproces. De stappen die in de literatuur worden gepresenteerd, komen terug in alle categorieën van wickedness. Echter, de literatuur houdt zowel geen rekening met een wederkerig verloop van dit proces, als met de toepassing van een toetsfase. Met betrekking tot de evaluatie-instrumenten voor externe evaluatoren was er relatief weinig literatuur beschikbaar, waardoor ik elf aanvullende instrumenten heb geïdentificeerd. Tot slot, dit onderzoek bevestigt dat het evalueren van doelmatig- en doeltreffendheid bij een hogere mate van wickedness nagenoeg onmogelijk is. Ondanks dat dit een bekend gegeven is in de literatuur over wicked problems, is dit minder gangbaar in de literatuur over beleidsevaluaties. Daarin wordt het evalueren van de doelmatig- en doeltreffendheid als standaard uitgangspunt genomen, ongeacht de mate van wickedness.

## 11. Aanbevelingen

### 11.1 Aanbevelingen voor externe, commerciële adviesbureaus

Op basis van de conclusie van dit onderzoek - de mate van wickedness speelt een rol bij alle onderzochte elementen van beleidsevaluaties - doe ik zes aanbevelingen aan externe, commerciële adviesbureaus. Ten eerste, stel de mate van wickedness vast zodat de onderzochte elementen van beleidsevaluaties hierop kunnen worden aangepast. Ten tweede, in alle categorieën, maar met name in de ongestructureerde categorie, worden beleidsevaluaties gebaseerd op kwalitatief onderzoek. Adviesbureaus zouden daarom moeten investeren in de kwaliteit en de toepassing hiervan. Ten derde, bespreek bij beleidsevaluaties met een hogere mate van wickedness de verwachtingen en perspectieven van de betrokken actoren. Ten vierde, houdt rekening met een wederkerig evaluatieproces, inclusief toetsfase, naarmate wickedness toeneemt. Ten vijfde, bij problemen met een hogere mate van wickedness is het evalueren van de doelmatig- en doeltreffendheid nagenoeg onmogelijk. In plaats daarvan zouden het leerproces en het proces waarmee het beleid tot stand is gekomen, onderwerp moeten zijn van de evaluatie. Tot slot, elk rapport zou een managementsamenvatting moeten bevatten.

### 11.2 Aanbevelingen voor opdrachtgevers van beleidsevaluaties

Naast de aanbevelingen aan externe, commerciële adviesbureaus doe ik ook een aanbeveling aan opdrachtgevers van beleidsevaluaties: het evalueren van de doelmatig- en doeltreffendheid is, naar mate de wickedness toeneemt, steeds minder goed mogelijk omdat problemen met een hoge mate van wickedness zeer afhankelijk, omvangrijk, en langdurig zijn. In plaats daarvan zou de evaluatievraag moeten focussen op het proces waarmee het beleid tot stand is gekomen.

### 11.3 Aanbevelingen voor toekomstig wetenschappelijk onderzoek

Tot slot, toekomstig wetenschappelijk onderzoek zou de bevindingen uit dit onderzoek met andere methoden, op een grotere schaal kunnen toetsen. De selectie van externe evaluaties kan worden uitgebreid, en eventueel worden aangevuld met een vergelijking tussen interne en publieke, externe evaluatoren. Daarbij kan worden gekeken naar aanvullende elementen van beleidsevaluaties, zoals het evaluandum of het evaluatiemotief. Maar toekomstig onderzoek zou ook kunnen kijken naar de toegevoegde waarde van praktijkbezoeken voor problemen met een hogere mate van wickedness. Aangezien deze problemen zeer contextafhankelijk en uniek zijn, zou dit instrument bij kunnen dragen aan het begrip van het probleem.

## 12. Literatuurlijst

Alford, J., & Head, B. (2017). Wicked and less wicked problems: A typology and a contingency framework. *Policy and Society*, 36(3), 397-413.

Balthasar, A. (2006). The effects of the institutional design on the utilization of evaluation: Evidenced using qualitative comparative analysis (QCA). *The International Journal of Theory, Research and Practice*, 12(3).

Bekkers, V. (2017) *Beleid in beweging. Achtergronden, benaderingen, fasen en aspecten van beleid in de publieke sector* (3e druk). Boom Bestuurskunde.

Bongers, F.J., & Den Hertog, P. (2007). Vernieuwing van methoden voor beleidsevaluatie. *Bestuurswetenschappen*, 3, 29-44.

Bouterse, M., & Pattyn, V. (2018). Welke factoren bevorderen of belemmeren het gebruik van beleidsevaluaties? Resultaten van een studie bij de Inspectie Ontwikkelingssamenwerking en Beleidsevaluatie. *Beleidsonderzoek Online*.

<https://doi.org/10.5553/BO/221335502018000003001>

Bovens, M. A. P., 'T Hart, P., Van Twist, M. J. W., Van den Berg, C. F. Van der Steen, M. A., & Tummers, L. G. (2017). *Openbaar Bestuur. Beleid, organisatie en politiek* (9e druk). Wolters Kluwer.

Bressers, J., & Hoogerwerf, A. (1995). *Beleidsevaluatie* (3<sup>e</sup> ed). Samsom Tjeenk Willink.

De Peuter, B., De Smedt, J., & Bouckaert, G. (2007). *Handleiding beleidsevaluatie Deel 1: Evaluatiedesign en -management* (Nr. D/2006/10106/013). Bestuurlijke organisatie Vlaanderen. Geraadpleegd op 6 mei 2022, van [https://steunpuntbov.be/rapport/s0205020\\_deel1.pdf](https://steunpuntbov.be/rapport/s0205020_deel1.pdf)

De Roo, K. (2002). Beleidsevaluatie: een gebruiksaanwijzing. *Natuurfocus*, 1(3), 97-102.

Ducket, D., Feliciano, D., Martin-Ortega, J., & Munoz-Rojas, J. (2016). Tackling wicked environmental problems: The discourses and its influence on praxis in Scotland. *Landscape and Urban Planning*, 154, 44-56.

Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, 26(2), 171–186. <https://doi.org/10.2307/2392467>

Haarhuis, K., & Parapuf, A. (2016). *Evaluatievermogen bij beleidsdepartementen. Praktijken rond uitvoering en gebruik van ex post beleids- en wetsevaluaties*. Geraadpleegd op 15 juni 2022, van [https://www.kcbr.nl/sites/default/files/juni\\_2016\\_evaluatievermogen\\_eindrapport.pdf](https://www.kcbr.nl/sites/default/files/juni_2016_evaluatievermogen_eindrapport.pdf)

Hajer, M. A. (1989). Discourscoalities in politiek en beleid. *Beleidswetenschap*, 1989(3), 242-263.

Hall, P. A. (1993). Policy Paradigms, Social Learning, and the State. *Comparative Politics*, 275-296.

Head, B. W. (2008). Wicked Problems in Public Policy. *Public Policy*, 3(2), 101-118.

Head, B. W., & Alford, J. (2015). Wicked Problems: Implications for Public Policy and Management. *Administration and Society*, 47(6), 711–739.

Hoogerwerf, A. (1987). Beleid berust op veronderstellingen: de beleidstheorie. In P. B. Lehning & J. B. D. Simonis (Reds.), *Handboek beleidswetenschap* (pp. 23-40). Boom.

Hoogerwerf, A. (2003). Beleidsvoorbereiding: het ontwerpen van beleid. In Hoogerwerf, A. & M. Herweijer (red.), *Overheidsbeleid* (7<sup>e</sup> ed) (pp. 89-107). Kluwer.

Houppermans, M. (2018). Wat is de toegevoegde waarde van de onafhankelijk deskundige bij beleidsdoorlichtingen? *Beleidsonderzoek Online*. Geraadpleegd op 15 juni 2022, van [https://www.beleidsonderzoekonline.nl/tijdschrift/bs0/2018/6/BO\\_2213-3550\\_2018\\_000\\_005\\_001/fullscreen](https://www.beleidsonderzoekonline.nl/tijdschrift/bs0/2018/6/BO_2213-3550_2018_000_005_001/fullscreen)

In 't Veld, R.J. (2010). *Kennisdemocratie, opkomend stormtij*. Academic Service.

Klijn, E.H., & Koppenjan, J. (2021). Besluitvormingen management in netwerken: een multi-actor perspectief op sturing. In T. Abma & R.J. in 't Veld (red.), *Handboek beleidswetenschap* (pp. 179-193). Boom.

Knottnerus, A., De Goede, P., & Van der Knaap, P. (2016). Systematisch leren van evalueren. *Bestuurskunde*, (25)2, 6-18. <https://doi.org/10.5553/Bk/092733872016025002002>

Korsten, A. F. A. (2013). *Wat ervan terecht komt: Zicht op beleidsevaluatie*. Geraadpleegd op 25 juni 2022, van:

<https://www.arnokorsten.nl/PDF/Onderzoek/ABC%20van%20beleidsevaluatie.pdf>

Korsten, A.F.A. (2016a). *Het intelligent aanpakken van beleidsproblemen*. Geraadpleegd op 15 juni, van

<https://arnokorsten.nl/PDF/Beleid/Beleidsproblemen%20intelligent%20aanpakken.pdf>

Korsten, A.F.A. (2016b). *Wicked problems in het kort*. Geraadpleegd op 22 juni 2022, van

<https://www.arnokorsten.nl/PDF/Beleid/Wicked%20problems%20in%20het%20kort.pdf>

Korsten, A. (2019). Omgaan met 'wicked problems'. *Beleidsonderzoek Online*. Geraadpleegd op 15 juni 2022, van [https://www.beleidsonderzoekonline.nl/tijdschrift/bs0/2019/3/BO\\_2213-3550\\_2019\\_000\\_002\\_001.pdf](https://www.beleidsonderzoekonline.nl/tijdschrift/bs0/2019/3/BO_2213-3550_2019_000_002_001.pdf)

Mathison, S. (2018). Does evaluation contribute to the public good? *Sage Journals*, 24(1), 113-119. <https://doi.org/10.1177/1356389017749278>

Ministerie van Financiën. (z.d.). *Evaluatiestelsel*. Geraadpleegd op 15 juni 2022, van <https://www.rijksfinancien.nl/beleidsevaluatie/evaluatiestelsel#:~:text=Ministeries%20zijn%20verplicht%20om%20beleid,regelgeving%20om%20subsidies%20te%20evalueren>

Mullins, L. J. (2002). *Management and Organisational Behaviour* (6e ed). Prentice-Hall.

Orde van organisatiekundigen en -adviseurs. (z.d.) *De kerncode voor kennisintensieve dienstverlening*. Geraadpleegd op 21 juni 2022, van <https://www.ooa.nl/wp-content/uploads/2022/04/Gedragcode.pdf>

Peters, B.G. (2017). What is so wicked about wicked problems: A conceptual analysis and a research program. *Policy and Society*, 36(3), 385-396.

Pierson, P. (2000). Increasing Returns, Path Dependence, and the Study of Politics. *American Political Science Review*, 94(2), 251-267. [https://www-jstor-org.eur.idm.oclc.org/stable/pdf/2586011.pdf?refreqid=excelsior%3A605db9087f66ab9d37a2bf1d2b1e363a&ab\\_segments=&origin=](https://www-jstor-org.eur.idm.oclc.org/stable/pdf/2586011.pdf?refreqid=excelsior%3A605db9087f66ab9d37a2bf1d2b1e363a&ab_segments=&origin=)

Pressman, J. L., & Wildavsky, A. (1973). *Implementation: How Great Expectations in Washington are Dashed in Oakland*. University of California.

Pröpper, I.M. (1998). Terugkoppeling: de vertaling van evaluatie naar beleid. In A. Hoogerwerf, & M. Herweijer (Reds.) *Overheidsbeleid: een inleiding in de beleidswetenschap* (pp. 141-162). Samsom H.D. Tjeenk Willink.

Ritchey, T. (2013). Wicked Problems: Modelling Social Messes with Morphological Analysis. *Swedish Morphological Society*, 2(1), 1-8.

Rittel, H. & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), 155-169.

Scott, W. R. (2003). *Organizations: Rational, Natural, and Open Systems* (5e ed.). Prentice Hall.

Simon, H. A. (1976). *Administrative behavior: A study of decision-making processes in administrative organization* (3e ed). Free Press.

Stolk, R., Wesseling, H., & Van der Beek, P. (2021). *De jacht op publieke waarde. Hoe overheden opgaven formuleren en realiseren*. Boom Bestuurskunde.



Stone, D. (2003). *The policy paradox*. Norton.

Strikwerda, J. (1998). Het anti-model van het adviesvak. *Management Consultant*, 42- 45.

Strikwerda, J. (2004). Is er nog behoefte aan professionele organisatieadviseurs? *Opleiding & Ontwikkeling*, 17(6), 9-12.

Terry, E. (2016). *De situationele benadering. Een onderzoek naar de betekenis en mogelijkheden van beleidsevaluatie in de ruimtelijke planning* [Proefschrift, Universiteit Gent]. Ghent University Academic Bibliography.

Termeer, K. (1993). *Dynamiek en inertie rondom mestbeleid*. VUGA.

Van Berkel, K., & Manickam, A. (2019). *Wicked world. Systeeminnovatie voor complexe vraagstukken*. Noordhoff Uitgevers.

Van den Berg, C. F., Van der Steen, M. A., & Tummers, L. G. (2021). *Openbaar bestuur, beleid, organisatie en politiek* (10e ed.). Wolters Kluwer

Van der Knaap, P., Pattyn, V. & Hanemaayer, D. (2020). *Beleidsevaluatie in theorie en praktijk* (1<sup>e</sup> ed.). Boom.

Van der Zwan, A., & Noordegraaf, M. (1998). Onderzoek van een complexe casus – Een gemeentevervoerbedrijf als voorbeeld. *Bestuurskunde*, 7(2), 62-74.

Van Noije, L. (2017). Mode in onderzoek. *Mens & Maatschappij*, 92(1), 1-6.  
<https://doi.org/10.5117/MEM2017.1.NOIJ>

Van Thiel, S. (2021). *Bestuurskundig onderzoek: een methodologische inleiding* (4<sup>e</sup> ed.). Uitgeverij Coutinho.

Verhoeven, N. (2011). *Wat is onderzoek? Praktijkboek voor methoden en technieken* (6<sup>e</sup> ed.). Boom.

Verlet, D., & Devos, C. (2008). Hoofdstuk 2: Prestatiemeting: doel of middel? Over de niet-evidente meting van prestaties binnen de publieke sector. In D. Verlet, & C. Devos (Reds.), *Over beleidsevaluatie: van theorie naar praktijk en terug* (pp. 44-67). Studiedienst van de Vlaamse Regering.

Wat is consulting? (z.d.). <https://www.consultancy.nl/carriere/wat-is-consulting>

### 13. Geraadpleegde Casussen

Askari, L., Grootelaar, H., Hoekstra, I., Schievels, J., & Van Veldhuisen, A. (2019). *Evaluatie Transitie Autoriteit Jeugd* (GV534). Andersson Elffers Felix.

Bernoster, C., Smeets, B., & Cohn, R. (2021). *Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant 2016-2021 (onderdeel Milieu)*. KWINK Groep.

Blokzijl, L., Boot, J., Kok, M., & Wierenga, W. (2021). *Preventie gehoorschade. Aanpak problematische blootstelling versterkte muziek*. Berenschot B.V.

Deloitte. (2021). *Evaluatie Convenant onbedwelmd slachten volgens religieuze riten*.

Kohlmann, R., Overgaauw, N., & Oost, M. (2022). *Sturing op duurzaamheid. Collegeonderzoek naar de sturing van gemeente Apeldoorn op klimaatadaptatie, circulaire economie en de energietransitie* (66353). Berenschot B.V.

Mulder, J., Janssen, L., & Van Mil, B. (2021). *Evaluatie Convenant e-lending*. KWINK Groep.

Oostdijk, A., Rikhof, M., Thijssen, N., & Wierenga, W. (2021). *Evaluatie Wet taken meteorologie en seismologie* (65385). Berenschot B.V.

Oostdijk, A., Van Polen, S., & Thijssen, N. (2022). *Evaluatie Radioactief afval. Een evaluatie van de Nota Radioactief afval en het Nationaal Programma Radioactief afval*. Berenschot B.V.

PWC. (2021). *Ministerie van BZK Evaluatie Nationaal Energie Bespaarfonds 2014-2019* (2021-0366/SM/el/lk).

Van Schelven, R., Mulder, J., Hofman, K., Blijleven, W., & Van Dijck, I. (2021). *Beleidsvaluatieonderzoek Streekomroepvorming*. KWINK Groep.

Van Mil, B., In 't Veld, R., Mulder, J., Hofman, K., & Ploum, M. (2022). *Periodieke evaluatie Gezondheidsraad 2017-2021*. KWINK Groep.

Van Veldhuisen, A., Schelfhout, D., Schievels, J., & Vloedbeld, C. (2020). *Evaluatie van de Aanwijzingen inzake de rijksinspecties (GB287)*. Andersson Elffers Felix.

## Bijlage 1: Methodologie

**Tabel 30**

*Deductieve coderingslijst*

Thema	Concept	Indicator
Gestructureerde categorie	Kennis over het probleem en de oplossing	Beschikbaar
	Consensus over het probleem en het doel	Aanwezig
	Betrokken actoren	Bekend
Slecht gestructureerde categorie	Kennis over het probleem en de oplossing	Beschikbaar
	Consensus over het probleem en het doel	Niet aanwezig
	Betrokken actoren	(On)bekend
Matig gestructureerde categorie	Kennis over het probleem en de oplossing	Niet beschikbaar
	Consensus over het probleem en het doel	Aanwezig
	Betrokken actoren	(On)bekend
Ongestructureerde categorie	Kennis over het probleem en de oplossing	Niet beschikbaar
	Consensus over het probleem en het doel	Niet aanwezig
	Betrokken actoren	Onbekend
Inrichting	Managementsamenvatting	
	Inleiding	Aanleiding
		Betrokken actoren
		Afkadering
		Leeswijzer
	Evaluatiedesign	Vraagstelling
		Evaluatiecriteria
		Evaluatieonderwerp
	Resultaten	
	Conclusie	Beoordeling
		Samenvatting
Aanbevelingen		

Instrumenten	Documentenanalyse	
	Interviews	
	Focusgroep	
	Expertpanel	
	Workshops	

**Tabel 31**

*Inductieve coderingslijst*

Thema	Indicator
Inrichting	Inhoudsopgave
Instrumenten	Klankbordgroep
	Enquete
	Begeleidingscommissie
	Casusanalyse
	Praktijkbezoek
	Validatiesessie
	Reflectiesessie
Feitenscan	
Doelmatigheid	Efficiency
Doeltreffendheid	Effectiviteit

## Bijlage 2: De Inrichting

**Tabel 32**

*Inrichting gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedwelmd slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E- lending</i>
Inhoud	Inhoud	Woord vooraf	Inhoud
Managementsamenvatting	Managementsamenvatting	Managementsamenvatting	Inleiding
Inleiding	Inleiding	Inhoudsopgave	Resultaten
Terugblik	Onderzoeksaanpak	Introductie	Ervaringen
Realisatie publieke waarde	De vormgeving van de Aanwijzingen	Evaluatie van afspraken over normen ter verbetering van het dierenwelzijn rond het doden	Conclusies en aanbevelingen
Conclusies en aanbevelingen	De uitwerking van de Aanwijzingen	Evaluatie van afspraken niet direct gerelateerd aan het doden	
	De praktijkwerking van de Aanwijzingen	Conclusies en aanbevelingen	
	Conclusies en aanbevelingen		

**Tabel 33***Inrichting slecht gestructureerde categorie*

<i>Beleidsevaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Inhoudsopgave	Inhoudsopgave
Inleiding	Inleiding
Achtergrond	Een maatschappelijk doel, een stelsel en een wet
Stand van zaken	Werking van het stelsel van meteorologie en seismologie
Betrokken partijen	Werking van de Wet taken meteorologie en seismologie
Conclusie en aanbevelingen	Conclusies en aanbevelingen
	Beantwoording onderzoeksvragen

**Tabel 34***Inrichting matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Management samenvatting	Inhoud	Inhoudsopgave
Inhoudsopgave	Conclusies en aanbevelingen	Inleiding
Introductie	Inleiding	Het beleid in perspectief
Onderzoeksaanpak	De context van de TAJ	Uitvoeringspraktijk anno 2022
Bevindingen: Effect van het fonds	Inrichting en positionering van de TAJ	Ervaringen met radioactief afvalbeleid
Bevindingen: Functioneren van het fonds	Taakuitvoering	Aandachtspunten voor het nieuwe Nationaal Programma
Conclusies en aanbevelingen	Effecten van het werk van de TAJ	



**Tabel 35***Inrichting ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Inhoud	Managementsamenvatting	Inhoudsopgave
Conclusies en aanbevelingen	Inhoudsopgave	Inleiding
Inleiding	Inleiding	Aanpakken en doorlopen onderzoeksproces
Opbouw PMWP	Feiten over het beleid en de sturing	Onderzoeksactiviteiten
Doeltreffendheid	Bevindingen over het beleid en de sturing	Stand van de wetenschap
Doelmatigheid	Conclusies	Scenario's met maatregelen
Opvolging aanbevelingen Tussenevaluatie en Statenmededeling	Aanbevelingen	Weging van de scenario's
		Uitvoeringsagenda
		Handreiking voor afgewogen besluitvorming

## Bijlage 3: Het Evaluatieproces

**Tabel 36**

*Proces gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedwelmd slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E- lending</i>
Doel	Doel- en vraagstelling	Doel- en vraagstelling	Afkadering
Vraagstelling	Structuur van de evaluatie	Ophalen en verwerken kwalitatieve en kwantitatieve resultaten	Ophalen en verwerken kwalitatieve en kwantitatieve resultaten
Analysekader		Beoordeling in de vorm van conclusies	Beoordeling in de vorm van conclusies
Ophalen, verwerken en analyseren kwalitatieve en kwantitatieve resultaten	Analysekader		
Toetsing bevindingen met de opdrachtgever	Ophalen kwalitatieve en resultaten		
Beoordeling in de vorm van conclusies	Verwerking en analyse van bevindingen		
	Beoordeling in de vorm van conclusies		

**Tabel 37***Proces slecht gestructureerde categorie*

<i>Beleidsevaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Onderzoeksvraag en scope	Doel- en vraagstelling
Verkenning	Analyse beleidstheorie, dan wel analysekader
Verdieping	Toetsing beleidstheorie met opdrachtgever en klankbordgroep
Rapportage	Ophalen kwalitatieve data
	Analyse bevindingen aan de hand van beleidstheorie, analysekader
	Beoordeling bevindingen in de vorm van conclusies
	Toetsing bevindingen met de betrokken actoren
	Aanvullende kwalitatieve data ophalen
	Nieuwe versie van het rapport opstellen
	Nieuwe versie valideren door klankbordgroep
	Oplevering definitief rapport

**Tabel 38***Proces matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Evaluatievragen	Doelstelling	Vraagstelling
Evaluatiekader	Vraagstelling	Orientatie (plan van aanpak, evaluatiekader)
Ophalen kwalitatieve en kwantitatieve data	Ophalen kwalitatieve en kwantitatieve resultaten	Gegevensverzamelng
Aparte beoordeling resultaten per deelvraag in de vorm van deelconclusies (en aanbevelingen)	Toetsen resultaten in reflectiesessie	Analyse aan de hand van evaluatiekader
Beantwoorden onderzoeksvragen	Aparte beoordeling resultaten per deelvraag in de vorm van deelconclusies	Bevindingen en conclusies
Aanbevelingen	Beantwoorden onderzoeksvragen	Tussenrapportage
	Aanbevelingen	Tussentijdse bevindingen besproken en gevalideerd met begeleidingscommissie
		Expertsessie
		Aanscherping onderzoeksresultaten
		Definitieve rapportage
		Validatie rapportage met begeleidingscommissie

**Tabel 39***Proces ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Doel, scope en onderzoeksvragen	Onderzoeksvragen	Onderzoeksvragen
Feitenscan	Analysekader	Oprichting begeleidingscommissie
Ophalen kwalitatieve data	input, throughput, output, outcome	Plan van aanpak
Inhoudelijke analyse	Ophalen kwalitatieve data	Reflectie Plan van Aanpak door begeleidingscommissie
Voorlopige conclusies ter beantwoording onderzoeksvragen	Rapporteren bevindingen en conclusies	Ophalen kwalitatieve data
Hoor- en wederhoor met betrokken medewerkers in een reflectiesessie	Valideren resultaten en uitwerking aanbevelingen met betrokken ambtenaren	Uitwerking mogelijke scenario's
Aanscherping bevindingen		Toetsing scenario's bij relevante partijen
Een externe reflectiesessie kon niet worden georganiseerd ivm onderzoekstechnische beperkingen		Toetsing resultaten bij begeleidingscommissie
		Gezamenlijke analyse van scenario's met begeleidingscommissie

		Opstellen uitvoeringsagenda's
		Rapportage resultaten

## Bijlage 4: De Instrumenten

**Tabel 40**

*Instrumenten gestructureerde categorie*

<i>Periodieke evaluatie Gezondheidsraad</i>	<i>Evaluatie van de Aanwijzingen inzake rijksinspecties</i>	<i>Evaluatie Convenant onbedwelmd slachten volgens religieuze riten</i>	<i>Evaluatie Convenant E- lending</i>
Deskresearch (doel: een eerste beeld van de feiten en cijfers over de Gezondheidsraad en zijn ontwikkeling tijdens de evaluatieperiode)	Desk research (doel: opzet en de uitwerking van de Aanwijzingen in beeld gebracht)	Desk research	documentenanalyse
jaarverslagen, jaarrekeningen, eerdere evaluaties, kabinetsreacties, relevante mediaberichten, Kamerstukken en andere bronnen	de Aanwijzingen, aanvullende regelingen en bepalingen voor het waarborgen van de onafhankelijkheid van de inspectie, opgesteld door individuele departementen	Openbaar beschikbare documenten (kamerstukken, wetteksten, verdragen, documenten van de toezichthouder), interne documenten van verschillende convenants partijen, de toezichthouder en de wetenschappelijke adviescommissie.	8 interviews
6 casussen (doel: aanvullend inzicht te vergaren over de advisering van de Gezondheidsraad en de effecten daarvan bij	Individuele interviews (doel: feitenmateriaal dat in de documentstudie is	Gesprekken	

ontvangers, wetenschappelijke peers, belanghebbenden en de maatschappij)	verzameld te duiden, en te verrijken met ervaringen over de betekenis en werking van de Aanwijzingen in de praktijk)		
(groeps)interviews (doel: aanvulling zijn op de feiten, cijfers en beelden uit de deskresearch)	27 stuks met Inspecteuren-generaals, vertegenwoordigers departementen, vertegenwoordigers van Bureau Inspectieraad, onafhankelijke experts. vooraf een interviewleidraad opgestuurd, achteraf een gespreksverslag ter verificatie gestuurd.	Convenantspartijen,(2 gesprekken met iedere convenantspartij, een aan het begin en een aan het eind ter validatie en verfijning van de bevindingen) wetenschappelijke adviescommissie, toezichthouder (dossierhouder, twee coördinatoren, vier toezichthouders) en slachthuizen/slachters.	
Gezondheidsraad (13 afdelingen, 33 gesprekspartners), VWS (3 afdelingen/6 respondenten), SZW (1), LNV (1), WRR (1), ROB (1), RVS (1), College ter beoordeling van geneesmiddelen (2), RIVM (1)	Groepsinterviews (doel: inzicht te krijgen in de (gedrags)dynamiek aan zowel de kant van inspecties als de kant van ministeries, en de rol die de Aanwijzingen daar al dan niet bij spelen)	Praktijkbezoeken aan twee slachthuizen waar onbedwelmd geslacht wordt	



Enquete	Eén groepsinterview met vertegenwoordigers van de ‘eigenaarsrol’, Drie groepsinterviews met inspecteurs die een belangrijke rol spelen in de interactie met het departement, Twee groepsinterviews met inspecteurs die werkzaam zijn in het primaire toezichtproces		
Begeleidingscommissie			
Gezondheidsdraad (6 deelnemers), VWS (1 deelnemer), WRR (1 deelnemer)			
Schriftelijke input (door 6 organisaties)			

**Tabel 41***Instrumenten slecht gestructureerde categorie*

<i>Beleidsevaluatieonderzoek Streekomroepvorming</i>	<i>Evaluatie Wet taken meteorologie en seismologie</i>
Documentenanalyse (doel: een feitelijk beeld van alle (nog te vormen) streekomroepen te krijgen)	Documentenstudie (analyse beleidstheorie)
Verkennde gesprekken (5: Regionale Publieke Omroepen (RPO), vertegenwoordigers van twee streekomroepen, het Commissariaat voor de Media (hierna: Commissariaat) en met het Stimuleringsfonds voor de Journalistiek)	Interviews (analyse beleidstheorie) t ministerie van IenW en EZK en het KNMI
Online informatievraag bij lokale omroepen (93 respondenten uit 51 streken)	Diepteinterviews (doel: de werking van de Wtms in de praktijk te onderzoeken aan de hand van ervaringen van betrokkenen in het stelsel van meteorologie en seismologie.) 27 semigestructureerde diepte-interviews afgenomen met betrokken stakeholders
Telefonische interviews met ambtenaren uit 69 gemeenten	Klankbordgroep (IenW - DG Water, MET Office, Weather Impact, Rijksuniversiteit Groningen, Unie van Waterschappen, Instituut Fysieke Veiligheid
Casusanalyse, 6 casussen (doel: verdieping te brengen aan en kleuring te geven aan de algemene beelden die ontstaan zijn uit de vorige onderdelen van het onderzoek.)	Begeleidingscommissie (doel: De bevindingen bespreken) KNMI en het Ministerie van IenW en in een sessie met weerbedrijven, de klankbordgroep en het opdrachtgeversberaad, waarin de opdrachtgevers van het KNMI zitting hebben
	Aanvullende interviews

**Tabel 42***Instrumenten matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Documentstudie (bestaande onderzoeksrapporten)	Documentstudie (jaarrapportages en interne beleidsdocumenten)	Documentstudie (documenten die onderdeel uitmaken van het radioactief afvalbeleid, verwante notities, documenten die nodig zijn voor contextuele duiding van het thema radioactief afval)
Verdiepende interviews (17 interviews gehouden met (mede)uitvoerders, BZK, partners, overige financiers, toezichhouders en onderzoekers)	Interviews (de commissieleden, medewerkers van het ondersteunend bureau, brancheorganisaties en gemeenten. De interviews hadden een semigestructureerd karakter en zijn gevoerd aan de hand van een vooraf opgestelde topiclijst die was afgeleid van het evaluatiekader.)	Diepte-interviews (doel: meningen en ervaringen verwerven)
Validatie (in de analyse fase hebben wij aanvullende gesprekken gevoerd met SVn, de Algemene Rekenkamer en BZK om de	Reflectiesessie (Op basis van de voorlopige resultaten uit de interviews is een reflectiesessie georganiseerd met vijf vertegenwoordigers van jeugdhulpaanbieders en gemeenten die in de praktijk met de TAJ te maken hebben gehad.	Begeleidingscommissie (doel: valideren (tussentijdse) bevindingen)

berekeningen en bevindingen te toetsen)	Deze reflectiesessie had tot doel de voorlopige bevindingen te toetsen, aan te scherpen en aan te vullen.)	
	Begeleidingscommissie (samengesteld door vws, bestaande uit VWS (2), J&V (1), Financien (1), VNG, BGZJ, ABD)	Expertsessie (doel: verdieping kennis op specifiek thema)

**Tabel 43**

*Instrumenten ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
Feitenscan (op door provincie aangeleverde relevante beleidsdocumenten)	Bureaustudie (belangrijke beleidsdocumenten en sturingsinformatie)	vastleggen feitelijke stand van zaken met behulp van experts in Notitie(3 in totaal)
Vragenlijsten (voor betrokken provinciale medewerkers)	(groeps)interviews (20 in totaal, met 18 interne betrokkenen en 11 externe partners (zoals energiecoöperaties, bestuurlijke partners en uitvoeringsorganisaties)	Groepsinterviews (7 Convenantpartners. • Aanbieders van versterkte muziek, die geen lid zijn van het convenant. • Organisaties die verantwoordelijk zijn voor handhaving: IGJ, omgevingsdiensten, gemeenten. • De zorgsector: JGZ/GGD, KNO-artsen, audiciens. • Ministerie van VWS.)

Interviews (met programmabeheersers Milieu, Energie en Circulaire Economie)	Sessies met betrokken ambtenaren (validatie resultaten en aanbevelingen verder uitgewerkt)	Begeleidingscommissie (doel: toetsen bevindingen en scenario's afwegen)
Interne reflectiesessie (Hoor- en wederhoor van bevindingen)		

## Bijlage 5: Doelmatig- & Doeltreffendheid

**Tabel 44**

*Doelmatig- en doeltreffendheid gestructureerde evaluaties*

Periodieke evaluatie Gezondheidsraad	Evaluatie van de Aanwijzingen inzake rijksinspecties	Evaluatie Convenant onbedweld slachten volgens religieuze ritens	Evaluatie Convenant e- lending
Doelmatigheid is geen vraagstuk n de evaluatie	Zijn de aanwijzingen doelmatig in relatie tot het functioneren van de rijksinspecties?	De grootste kosten van dit Convenant zijn de kosten voor het permanent toezicht. Deze kosten worden als doelmatigheid beschouwd. [...]	Doelmatigheid is geen vraagstuk n de evaluatie
	[...[ we concluderen we deze evaluatie dat de Aanwijzingen een belangrijk middel zijn om de onafhankelijkheid van rijksinspecties te markeren en borgen.  [...] we concluderen dat de administratieve lasten en andere (ongewenste) neveneffecten van de Aanwijzingen (zeer) beperkt zijn.		

	In die zin zijn de Aanwijzingen dan ook als doelmatig te karakteriseren		
In hoeverre is de Gezondheidsraad effectief in het realiseren van zijn publieke waarde?	Worden de doelstellingen van de aanwijzingen met de daarvoor bestaande instrumenten in de praktijk bereikt?	Het Convenant heeft geleid tot een aantoonbare verbetering van het dierenwelzijn in de praktijk. De normen geformuleerd in het convenant (artikel 2 en de technische verfijningen) worden in de praktijk nageleefd en zijn een verbetering van het dierenwelzijn ten opzichte van de nul-meting in 2014.	Doeltreffendheid is geen vraagstuk in de evaluatie
Onze conclusie is dat de publieke waarde van de Gezondheidsraad is gelegen in advisering vanuit de combinatie van enerzijds de toegevoegde waarde van de wetenschappelijke	[...] In algemene zin kan dan ook worden geconcludeerd dat de Aanwijzingen een belangrijke rol spelen in het waarborgen van onafhankelijk opererende rijksinspecties.  Toch is het niet vanzelfsprekend dat		

<p>invalshoek en anderzijds de kunst van het binden en benutten van een grote hoeveelheid expertise bij een groot aantal wetenschappers en experts.</p>	<p>rijksinspecties voldoende onafhankelijk te werk kunnen gaan. De Aanwijzingen dekken dus niet alle risico's die zich in de praktijk voor kunnen doen met het oog op onafhankelijkheid.</p> <p>[...]</p>		
---	---	--	--

**Tabel 45**

*Doelmatig- en doeltreffendheid slecht gestructureerde categorie*

<p><i>Beleidsvaluatieonderzoek Streekomroepvorming</i></p>	<p><i>Evaluatie Wet taken meteorologie en seismologie</i></p>
<p>De effectiviteit van streekomroepen als instrument om te komen tot een relevante en levensvatbare lokale omroep.</p>	<p>Is het proces bij maatschappij-ontwrichtend weer/calamiteiten zoals bedoeld in artikel 5 van de Wtms efficiënt en doeltreffend in het licht van publieke doelen (veiligheid, bereikbaarheid, leefomgeving en welvaart)?</p>
<p>Streekomroepvorming is een middel dat bijdraagt aan het haalbaar en betaalbaar maken van een LTMA in Nederlandse gemeenten. Uit onderzoek blijkt dat streekomroepen beter in staat zijn om de kwaliteitsdoelstellingen te behalen om een LTMA te verzorgen.</p>	<p>Het proces bij maatschappij-ontwrichtend weer of calamiteiten kan verbeterd worden om de publieke veiligheid bij extreem weer beter te borgen.</p>



<p>Samenwerking in een streekomroep zorgt voor schaalvergroting, een professionelere organisatie, meer aanbod op verschillende kanalen en veelal een samenvoeging van budgetten.</p> <p>[...] streekomroepvorming alleen is niet voldoende voor het behalen van de kwaliteitsdoelstellingen. Het grootste knelpunt voor het kunnen verzorgen van een LTMA is de beperkte financiering van lokale publieke omroepen.</p> <p>De tijdelijke stimulansen die zijn opgezet voor de versterking van de lokale media zijn kleine impulsen die versterkend kunnen werken maar ook tijdelijk zijn.</p>	
<p>Doeltreffendheid is geen onderdeel van de evaluatie</p>	<p>Wat is de doeltreffendheid van de Wtms in de praktijk in het licht van de zorgtaken van de minister in het licht van het publieke belang (veiligheid, bereikbaarheid, leefomgeving en welvaart)?</p>
	<p>Wij concluderen dat de Wtms in de praktijk een doeltreffend instrument is en bijdraagt aan de realisatie van de beoogde publieke waarde (zie paragraaf 5.1). [...]</p>

**Tabel 46**

*Doelmatig- en doeltreffendheid matig gestructureerde categorie*

<i>Evaluatie Nationaal Energie Bespaarfonds</i>	<i>Evaluatie Transitie Autoriteit Jeugd</i>	<i>Evaluatie Radioactief afval</i>
Doelmatigheid is geen onderdeel van de evaluatie	In hoeverre is de TAJ conform het instellingsbesluit voldoende doeltreffend en doelmatig geweest [...]	Daarnaast wordt de [...] doelmatigheid van het beleid geanalyseerd.
	De TAJ heeft als organisatie binnen de kaders van haar opdracht doelmatig gefunctioneerd.  In bredere zin zijn er echter wel kanttekeningen te plaatsen bij de effectiviteit en doelmatigheid van de activiteiten van de TAJ.	Geen eenduidig antwoord.
Het effect van het fonds (doeltreffendheid)	In hoeverre is de TAJ conform het instellingsbesluit voldoende doeltreffend en doelmatig geweest [...]:	<i>In hoeverre worden met het huidige beleid en de uitwerking van het beleid de beoogde doelstellingen bereikt [...]?</i>
[...] concluderen wij dat het NEF in de periode 2014-2019 over het algemeen doeltreffend is geweest, maar sommige fases (m.n. de opstart) en activiteiten (o.a. scholen en	De TAJ is effectief geweest in het realiseren van zorgcontinuïteit.  In bredere zin zijn er echter wel kanttekeningen te plaatsen bij de effectiviteit en	In conclusie kan gesteld worden dat het beleidskader effectief en overzichtelijk is ingericht.

laadpalen) minder doeltreffend zijn gebleken.	doelmatigheid van de activiteiten van de TAJ.	
---	---	--

**Tabel 47**

*Doelmatig- en doeltreffendheid ongestructureerde categorie*

<i>Evaluatie Provinciaal Milieu- en Waterplan Noord-Brabant</i>	<i>Sturing op duurzaamheid</i>	<i>Preventie gehoorschade.</i>
In hoeverre is het PMWP [...] doelmatig geweest?	<i>In hoeverre is (de uitvoering van) het duurzaamheidsbeleid in Apeldoorn [...] doelmatig, [...]?</i>  <i>In hoeverre is de uitvoering van de uitvoeringsagenda energietransitie in Apeldoorn [...] doelmatig [...]?</i>	Doelmatigheid is geen onderdeel van de evaluatie
Wij kunnen geen conclusie trekken over de vraag of het PMWP doelmatig is geweest.	Voor energietransitie geldt dat de uitvoeringsagenda de afgelopen jaren sterk is doorontwikkeld. De mate waarin deze inzet doelmatig is geweest, kan niet afdoende worden beoordeeld zonder verdere operationalisering van ambities, doelen, resultaten en activiteiten en de samenhang daartussen.	
In hoeverre is het PMWP doeltreffend [...] geweest?	In hoeverre is (de uitvoering van) het duurzaamheidsbeleid en van de uitvoeringsagenda Energietransitie doeltreffend?	Doeltreffendheid is geen onderdeel van de evaluatie

<p>Wij concluderen dan ook dat het aannemelijk is dat het PMWP in beperkte mate doeltreffend is geweest met een duidelijk verschil in doeltreffendheid tussen de wettelijke taken en de provinciale ambities.</p>	<p>Op deze vraag is geen overall conclusie te trekken.</p> <p>[...] het is ook duidelijk dat de onzekerheden en afhankelijkheden bij deze transitie zo groot zijn, en de gestelde termijn zo lang, dat een harde uitspraak over doeltreffendheid altijd lastig zal zijn.</p>	
---	--	--