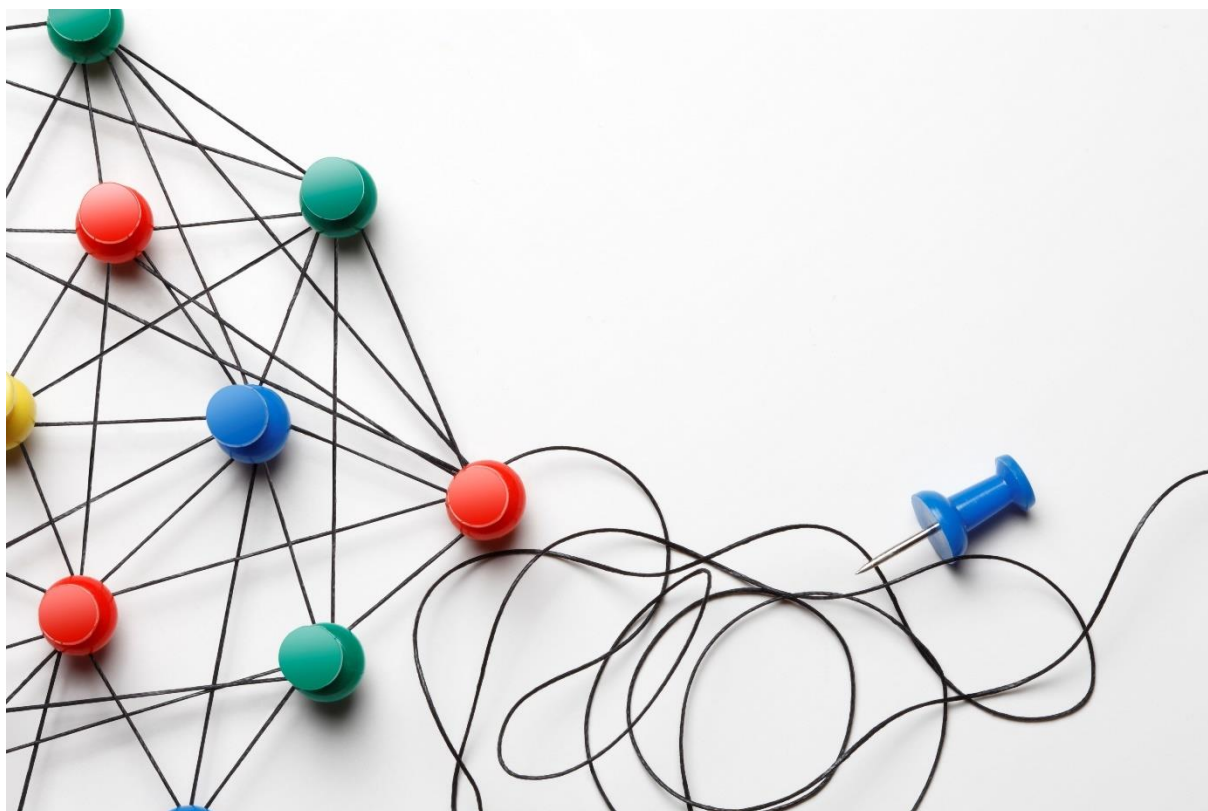


Een zoektocht naar de (on)mogelijkheid tot Responsief Auditen...

... en de reflecterende weg daarnaartoe.



dhr. F. Schuiten MA
studentnummer: 572036
Den Haag, 2 december 2022

Referaat t.b.v. behalen van de
Post-master opleiding *Internal Audit & Advisory*
Erasmus Universiteit Rotterdam
Erasmus School of Accounting & Assurance

Onder begeleiding van: dhr. R. De Korte RA RE RO CIA

Openbaar, met bronvermelding
Aantal woorden: 17449

Voorwoord

Ha lezer!

Na anderhalf jaar afstudeeronderzoek kan ik met enige trots dit referaat voorleggen. Wat allemaal is begonnen met een nieuwsgierigheid is een verkenning – een heuse zoektocht – geworden. Het resultaat ervan heb ik in de volgende veertig pagina's weergegeven.

Die weergave was niet alleen een inhoudelijke zoektocht, maar ook een zoektocht naar de juiste presentatie: de structuur en volgordelijkheid, de doelgroep, de mate van reflectie in het document zelf, enzovoorts. Ik neem je daarom in twee delen mee. Het eerste deel heb ik geschreven voor internal auditprofessionals die iets anders willen dan onderzoeken met een normenkader en voor hen die geïnteresseerd zijn in een reflectie op het primaat van de grondbeginselen van objectiviteit en onafhankelijkheid binnen een internal auditfunctie. Gebaseerd op Tineke Abma's Responsieve Evaluatie heb ik een plan geschreven voor een internal auditor om toe te kunnen passen. Maar let wel: vooralsnog niet als audit. Op die stelling reflecteer ik ook in het eerste deel van dit referaat. Het tweede deel heb ik geschreven voor hen die geïnteresseerd zijn in de achtergrond van dat verhaal. In de beknopte achtergronden van wetenschapsfilosofie en evaluatietheorie in relatie tot internal auditing; in een kritische beschouwing van de grondpijlers van de Responsieve Evaluatie; in de aanpak en resultaten van mijn praktijkvalidatie; in de beperkingen van het referaat en de wegen voor toekomstig onderzoek.

De illustratie op de voorpagina geeft vervolgens m.i. passend zowel procesmatig als inhoudelijk de kern van deze zoektocht weer. In verschillende reflectieve cycli heb ik vorm kunnen geven aan een plan voor een responsieve aanpak binnen een internal auditfunctie én daarop kunnen reflecteren met anderen. Het is een wirwar geweest met telkens nieuwe informatie, maar – achteraf bezien – met duidelijke kernpunten (de punaises), die dan vervolgens een nieuwe cyclus startten. Het laat ook zien dat er nog een weg te gaan is (de liggende punaise en overgebleven draad). Inhoudelijk interpreteer ik de illustratie als verschillende werkelijkheden die in continue interactie en verbinding met elkaar staan en op die manier vorm krijgen; via de draden interacteren deze punaises met elkaar.

Ik heb veel te danken aan Ron als begeleider. Hij gaf precies wat ik nodig had: een stopsignaal wanneer het te vaag, te abstract werd, of ik weer een nieuwe doos van Pandora wilde opentrekken; iemand die het visionaire realistisch kon houden. Maar ook iemand die je doorstuurt naar anderen als hij het niet weet, of even niet de geschikte begeleider was. Dat heeft mij echt geholpen om dit referaat tot een – in ieder geval voor mij – mooi einde te kunnen brengen. Een einde van onze samenwerking is het zeker niet.

En dan wil ik ook een dankwoord aan Karen geven, die mijn waterval over wetenschapsfilosofie, evaluatietheorie en het auditvak meermaals heeft aangehoord, verschillende versies van dit referaat heeft doorgenomen, maar bovenal mijn eigen aannames zeer kritisch heeft bevraagd. Een flinke bijdrage aan mijn reflexiviteit als onderzoeker is het zeker geweest.

Nu volgt eerst een samenvatting van het referaat, dan de inhoudsopgave. En vervolgens neem ik je mee in het plan van aanpak (eerste deel) en de methodologische verantwoording (tweede deel).

Ik wens je veel leesplezier.

Hartelijke groet,

Frank

De onmogelijkheid van Responsief Auditen en de mogelijkheid van responsieve internal auditactiviteiten

Een samenvattende weergave van het referaat.

Dit afstudeeronderzoek startte met de vraag hoe een andere wijze van auditen die meervoudige werkelijkheden veronderstelt, de gereedheidskist van de internal auditor kan verrijken. De deductieve methodologie, gestoeld op aannames van een enkelvoudige werkelijkheid, hebben mij nooit gepast. Daarom wilde ik verkennen of en hoe het anders kon. Daarvoor heb ik uitgebreid literatuuronderzoek uitgevoerd, met een lichte empirische validatie.

Uit mijn literatuuronderzoek bleek dat een andere manier van onderzoeken dan met een deductieve methodologie, zoals we die doorgaans binnen het auditwerkveld kennen, mogelijk is en dat deze binnen de evaluatiegemeenschap, als aanpalend werkgebied van auditing, reeds wordt toegepast met de Responsieve Evaluatie. Een aanpak die de verhalen van alle belanghebbenden bij een onderzoek centraal stelt in verkennende gesprekken, verhalenworkshops en een dialoggroep, waaruit die meervoudige werkelijkheden blijken. Een aanpak waarbij niet consensus centraal staat, maar wederzijds begrip. Een aanpak waarin niet een inhoudelijk-technische expertrol voor de onderzoeker is weggelegd, maar die van verhalenanalist, Socratische gids, educator en facilitator. Voor een specifieke casus heb ik op basis van inzichten vanuit wetenschappelijke literatuur een plan van aanpak geschreven en dat vervolgens licht gevalideerd bij verschillende senior collega's en managementteamleden binnen de Auditdienst Rijk.

Reflecterend blijkt echter dat het niet mogelijk is om een wijze van *auditen* toe te passen die meervoudige werkelijkheden veronderstelt. De objectivistische auditlogica waarin de beginselen van onafhankelijkheid en objectiviteit regievoeren, botsen met de grondbeginselen van een responsieve aanpak – en in het verlengde daarvan met iedere aanpak die meervoudige werkelijkheden veronderstelt. Daarentegen is er wel ruimte om deze responsieve aanpak toe te passen *binnen de variëteit van activiteiten van de internal audit functie (IAF)*. Deze omvatten immers meer dan alleen auditen. Er is ruimte voor een andere onderzoeksaanpak die gebaseerd kan zijn op andere dan – en dus los kan staan van – de fundamentele principes die regeren binnen de hiervoor genoemde objectivistische auditlogica. Er is ruimte om weg te stappen van het ideaal van een eenduidige werkelijkheid, die objectiveerbaar kan worden gemeten of vastgesteld. Er is ruimte om in nabijheid en interactie met de organisatie tot inzichten te komen en daarin meer een faciliterende, educatieve, Socratische rol te vervullen in plaats van een inhoudelijke, objectieve expertrol. Pragmatisch gezien vereist deze aanpak wel specifieke onderzoeksvaardigheden die een *behavioural auditor* toepast en een fundamenteel andere grondhouding die de *trusted advisor* als muziek in de oren klinkt.

Dat roept bredere vragen op voor de IAF. In hoeverre is het nog mogelijk (dan wel raadzaam) om verschillende activiteiten toe te voegen aan die van de IAF? Floreren we dan in verscheidenheid of verliezen we focus? In hoeverre is het de taak van de internal auditor nog om die aanvullende activiteiten uit te voeren? Of sterker: zijn we internal auditors – en daarmee gehouden aan beroepsregels – die ook andersoortige onderzoeken uitvoeren en consultancytaken kunnen vervullen? In hoeverre moeten internal auditors daarin rolvast blijven vasthouden aan de principes van onafhankelijkheid en objectiviteit? Of zijn we allicht meer breed georiënteerde onderzoekers en adviseurs die verschillende activiteiten uitvoeren, waaronder internal audits, en ook alleen dan als internal auditor fungeren? Ik heb hier (nog) geen antwoord op; verder onderzoek is nodig.

Het referaat toont dat de Responsieve Audit, momenteel binnen de specifieke context van de Auditdienst Rijk, (nog) niet mogelijk is. De bijdrage geeft inzichten voor het brede internal auditvak door de objectivistische grondbeginselen ter discussie te stellen voor activiteiten van de IAF, anders dan auditen. Bovendien blijkt dat deze aanpak uitermate geschikt is voor ambigue vraagstukken, waar internal auditfuncties voor kunnen worden gevraagd.

Inhoudsopgave

Het begin van een zoektocht.....	5
----------------------------------	---

Eerste deel: Plan van Aanpak voor een Mogelijke Responsieve Audit

1. Audit (beknopt) kritisch beschouwd vanuit wetenschapsfilosofie en evaluatietheorie.....	7
2. Stappenplan voor een mogelijke Responsieve Audit bij het ministerie van Onderwijs, Cultuur en Wetenschap.....	9
3. Reflectie op de (on)mogelijkheid van Responsief Auditen	14
<i>Pragmatische zijde van onderzoeksvaardigheden en -houding.....</i>	14
<i>Responsieve onderzoeksaanpak als auditaanpak?.....</i>	14
<i>Responsieve aanpak binnen de ADR als IAF?.....</i>	15
Einde in zicht?.....	17

Tweede deel: Methodologische Verantwoording

Methodologische verantwoording met vijf reflectieve cycli.....	19
Eerste cyclus: het plan (op hoofdlijnen) voor een afstudeeronderzoek.....	20
Tweede cyclus: Audit bezien vanuit wetenschappelijke literatuur.....	22
<i>Audit omvat een te onderzoeken realiteit met een normenkader leidend tot een oordeel</i>	22
<i>Auditing als een objectieve activiteit– waarheid, ideaal, of utopie?</i>	22
Derde cyclus: Hoe audit en evaluatie dichter bij elkaar liggen dan je denkt	26
<i>Audit en evaluatie als verwante vakgebieden</i>	26
<i>Drie Generaties van Evaluatie</i>	26
<i>Vierde evaluatiegeneratie gegrond in postmodernisme.....</i>	27
<i>Kritiek richting aannames over macht, actie en consensus</i>	28
Vierde cyclus: Abma's Responsieve Evaluatie.....	30
<i>Responsieve Evaluatie: wat doe je kortgezegd?.....</i>	30
<i>Responsieve Evaluatie: waarvoor staat het?.....</i>	31
Vijfde cyclus: Van Responsieve Evaluatie naar een aanpak voor een Responsieve Audit?	38
<i>Empirische validatie met tweede versie van plan van aanpak als resultaat.....</i>	38
<i>Beperkte methodologische fundering leidt tot noodzaak voor aanvullende empirische validatie</i>	39
<i>Uitkomsten van de aanvullende empirische validatie, qua actiegerichte doelstelling</i>	40
<i>Uitkomsten van de aanvullende empirische validatie, qua reflectieve doelstelling.....</i>	41
<i>De academische korrels zout bij deze zoektocht.....</i>	41
<i>De weg naar toekomstige zoektochten</i>	42
Literatuurlijst	44

Het begin van een zoektocht...

Iedere audit is een te starten zoektocht. Naar een mogelijk probleem in de interne beheersing; naar causale verbanden voor dat probleem; of naar impactvolle aanbevelingen. Het zal veel lezers – bij wie ik enig aanverwantschap tot het auditveld veronderstel – daarom niet verbazen dat dit sluitstuk van de post-masterstudie ook een zoektocht was. Die startte voor mij bij een nieuwsgierigheid over de boeiende, postmoderne notie van meerdere, tegelijkertijd bestaande sociale werkelijkheden, versus de notie van een enkelvoudige werkelijkheid die objectief en onafhankelijk kan worden onderzocht.

Objectiviteit en onafhankelijkheid worden als fundamenteën van het auditwerkveld beschouwd. Een ‘klassiek’ uitgevoerde audit veronderstelt dat auditors – met een vooraf opgesteld normenkader – een objectief oordeel kunnen verstrekken. De hiermee gepaarde vanzelfsprekendheid kan in twijfel worden getrokken (bijv. Francis, 1994; Nordin, 2022). Vanuit een nieuwsgierigheid om aan deze fundamenteën te schudden, ben ik deze zoektocht begonnen met de vraag of een auditbenadering, die meervoudige werkelijkheden veronderstelt, mogelijk is. Het gaat daarbij niet alleen om het gebruiken van een *andere methode*, maar om een methode met *andere fundamenteën* (vgl. Lincoln & Guba, 1985).

In de afgelopen jaren is de toegevoegde waarde van internal auditafdelingen (IAF) veelvuldig onderzocht en wordt geen rooskleurig toekomstbeeld gegeven (bijv. Lenz & Hahn, 2015; Türetken e.a., 2020). Daarom pleiten Eulrich en Lenz (2020) voor een IAF die niet langer alleen assurance-werkzaamheden verricht, maar om de randen van het vakgebied op te zoeken om zó van meerwaarde te blijven. Met deze bijdrage wil ik proberen de gevestigde grenzen – de hiervoor genoemde fundamenteën – op te zoeken. Dat is geen eenvoudige zoektocht geweest, maar heeft geleid tot tastbaar resultaat: een onderzoeksplan, gestoeld op responsieve principes (Abma, 2001, 2005, 2006).

Met het eerste onderdeel van deze bijdrage (aangegeven met blauwe titels) neem ik je mee in deze verkenning. Daarmee heeft dit eerste onderdeel twee doelen: een beschrijving van een responsieve aanpak die je praktisch kan toepassen; en een kritische beschouwing over het gebruik en de onderliggende fundamenteën van deze methode binnen de context van het auditveld. Deze tweeledige doelstelling geeft ook de relevantie aan: een praktisch toepasbare methode enerzijds, een mogelijke bijdrage aan het debat rondom de meerwaarde van de IAF anderzijds.

In de volgende drie hoofdstukken geef ik antwoord op de vraag *“Hoe kan een sociaal constructionistische methode van onderzoek de gereedschapskist van de internal auditor verrijken?”* Daartoe heb ik uitgebreid literatuuronderzoek en een lichte empirische validatie uitgevoerd. In dit onderdeel schets ik de resultaten daarvan, aangevuld met een analytisch-beschrijvende reflectie. Dat doe ik allereerst met een bondige beschouwing van wetenschappelijke literatuur over audit, (wetenschaps)filosofie en evaluatietheorie. Daarna volgt een uiteenzetting van een plan van aanpak gestoeld op responsieve principes. Deze bijdrage sluit af met een reflectie over de plek van deze aanpak in de discussie over de meerwaarde van de IAF. Voor de uitgebreide toelichting op het uitgevoerde literatuuronderzoek en de empirische validatie, verwijs ik naar het tweede deel van deze bijdrage: de methodologische verantwoording (aangegeven met groene titels). Waar ik in deze bijdrage kort en bondig refereer naar literatuur en inzichten van professionals, biedt de verantwoording een rijker beeld.

EERSTE DEEL

Plan van

Aanpak

1

Audit (beknopt) kritisch beschouwd vanuit wetenschapsfilosofie en evaluatietheorie

Onderstaand beschouw ik het auditveld in vogelvlucht. Daardoor doe ik geenszins recht aan de onderliggende complexiteit. De methodologische verantwoording biedt meer diepgang. Eerst beschrijf ik drie wereldbeelden vanuit de wetenschapsfilosofie, gelinkt aan auditing en evaluatietheorie. Vervolgens beschrijf ik de kernprincipes en fasen van de Responsieve Evaluatie.

De dominante auditvorm stoelt op aannames die voortkomen uit wetenschapsfilosofische stromingen (Van Twist & De Korte, 2022). Iedere audit biedt inzicht in een fenomeen (vgl. Schuiten & De Korte, 2022). Daaraan kan zekerheid worden ontleend, hetzij expliciet door het afgeven van assurance, hetzij impliciet wanneer een auditprofessional een onderzoek uitvoert (Driessen & Molenkamp, 2012). Hieronder schuilt een objectief wereldbeeld: er is één werkelijkheid die we (zo) objectief (mogelijk) waarnemen. De auditaanpak is deductief met een vooraf opgesteld normenkader. Een auditor streeft het ideaal van objectiviteit na. Velen herkennen zich hierin, ongetwijfeld.

De grootste uitzondering is de *behavioural audit* (Van der Meulen & Otten, 2013a, 2013b), die een subjectief wereldbeeld veronderstelt: er is sprake van meervoudige werkelijkheden (Boonstra, 2000). Er wordt – cognitieve psychologie volgend – verondersteld dat deze realiteiten worden weerspiegeld in mentale modellen die menselijk gedrag aansturen. De aanpak is inductief: verzamelde onderzoeksgegevens zijn leidend, niet een vooraf opgesteld normenkader. Een meer intersubjectief wereldbeeld is ook mogelijk (Brown, 2017). De aanname van meervoudige werkelijkheden houdt stand. Verschillend is echter dat wordt verondersteld dat realiteiten in en dankzij menselijke interactie (lees: in dialoog) ontstaan (Brown, 2017; Cunliffe, 2008; Hosking, 2011). Ze bestaan niet los van deze interactie (Burr & Dick, 2017). Binnen het auditveld zien we hier nog geen verschijningsvorm van.

De evaluatiegemeenschap biedt als aanverwant vakgebied inzichten. Hoewel evaluatie en audit verschillen, delen zij overeenkomsten (vgl. Bemelmans-Videc, e.a., 2017; Mayne, 2006; Naidoo, 2020). Daarnaast heeft de evaluatiegemeenschap voor eenzelfde verwondering gestaan. In hun magnum opus beschrijven Guba en Lincoln (1989) drie evaluatiegeneraties: eerst werden evaluaties als objectieve metingen gebruikt; vervolgens werden deze aangevuld met kwalitatieve beschrijvingen; ten slotte werd ook een evaluatie-oordeel toegevoegd. De gelijkenissen met het auditvak zijn overduidelijk.

Met hun vierde evaluatiegeneratie pleitten zij voor een paradigmawisseling, waarbij niet langer een objectief, onafhankelijk oordeel, maar onderlinge interactie en onderhandelde betekenisgeving centraal stond. Anderen noemen dit ook wel responsief evalueren (Stake, 1975; Visse, e.a., 2011). Dit leidde kritiek, ook vanuit gelijkgestemden, over het gebrek aan aandacht voor machtsrelaties (Vanderplaat, 1995), het gebrek aan actie en hun semi-queeste naar consensus (bijv. Gregory, 2000; Laughlin & Broadbent, 1996). Als antwoord ontstond een specifieke stroom binnen de responsieve evaluatie die verhalen en dialoog centraal stelt, waarvan de grondslagen overeenkomen met het intersubjectieve wereldbeeld (Abma, 2001, 2005, 2006; Abma & Widdershoven, 2005; Baur, e.a.,

2010; Edley, 2001; Niessen, e.a., 2008; Visse e.a., 2011; Van der Haar & Hosking, 2004; Widdershoven, 2001).¹

De Responsieve Evaluatie bestaat uit vier fases (Abma, 2005; Abma & Widdershoven, 2005; Visse, e.a., 2011). Allereerst moeten evaluatoren een goed beeld krijgen van de context, problematiek en relevante belanghebbenden. Vooral dat laatste is relevant, omdat stakeholders de meest centrale rol spelen. In de tweede fase gaat de evaluator hen het gesprek aan, om hun verhalen omtrent het evaluatievraagstuk in kaart te brengen, tot saturatie is bereikt. Deze verhalen worden verdiept in verschillende focusgroepen met groepen belanghebbenden. Uiteindelijk worden deze verschillende verhalen nog samengebracht in een dialooggroep met vertegenwoordigers van de verschillende belanghebbenden.

Hét kernconcept is dialoog (Greene, 2001; Schwandt, 2001). Immers: in dialoog bestaan en ontstaan realiteiten. Het uiteindelijke doel is het stimuleren van wederkerig begrip, zodat onderling leren en het formuleren van vervolgacties kunnen worden gestimuleerd (Baur, e.a., 2010). Dat betekent ook dat je rol als onderzoeker anders is: je bent allesbehalve een inhoudelijk expert. Je analyseert verhalen; bewaakt de kwaliteit van de oprechte dialoog; faciliteert onderling begrip; fungeert als socratische gids; en hebt de rol van een educator (Abma & Widdershoven, 2011). Je bent je ervan bewust dat een evaluatie een interventie op zichzelf is en bovenal een activiteit waarin sociale verhoudingen continu aandacht behoeven.

De verschillen met de klassiek uitgevoerde audit zijn groot. Een vergelijking van de drie kerncomponenten van een toetsend onderzoek illustreert dit. Bij een klassiek uitgevoerde audit analyseert de auditor met een vooraf opgesteld normenkader zo objectief mogelijk de waar te nemen realiteit om tot bevindingen en/of conclusies te komen. Hoewel de Responsieve Evaluatie werkt met een stappenplan, verschijnt het daadwerkelijke inhoudelijke ontwerp pas tijdens het onderzoek. Er is geen vooraf opgesteld normenkader; normen en waarden (evenals bevindingen en/of conclusies) komen tot stand in de dialoog met en tussen belanghebbenden. Daarin staan de meervoudige realiteiten centraal. Om deze ten tonele te laten treden, vertellen belanghebbenden hun verhalen daarbij. Onderstaande tabel vat dit samen.

Hierbij veronderstel ik de kerncomponenten van toetsend onderzoek als regels die bepalen wat als audit kan worden beschouwd. Hoewel de invulling van de drie componenten op geheel andere fundamenteën is gebaseerd, lijkt het mogelijk om een toetsend onderzoek uit te voeren met de grondbeginselen van de Responsieve Evaluatie. Het volgende hoofdstuk zet hiervan het stappenplan uiteen voor een concrete casus.

Kerncomponent (Verschuren & Doorewaard, 2021)	Klassiek uitgevoerde audit (Bos, De Korte & Otten, 2017)	Responsieve Evaluatie, als toetsend onderzoek
Norm	Vooraf opgesteld normenkader.	Normen en waarden komen tijdens het onderzoek intersubjectief in dialoog tussen belanghebbenden tot stand.
Onderzoeksobject	Een eenduidige realiteit: het (te meten) auditobject.	Meervoudige realiteiten die via verhalen ten tonele treden.
Bevindingen en conclusie/oordeel	Bevindingen vloeien voort vanuit de vergelijking van normenkader en auditobject; een conclusie of oordeel wordt zo objectiveerbaar mogelijk gemaakt door triangulatie en afspraken met de opdrachtgever.	Bevindingen en conclusies of oordelen komen intersubjectief in dialoog tussen belanghebbenden tot stand.

Tabel 1. Vergelijking van de auteur tussen de klassiek uitgevoerde audit en de responsieve evaluatie, gezien vanuit kerncomponenten van een toetsend onderzoek.

¹ In de rest van deze bijdrage duid ik deze specifieke stroom met hoofdletters, om het verschil met de bovenliggende benadering te duiden.

2

Stappenplan voor een mogelijke Responsieve Audit bij het ministerie van Onderwijs, Cultuur en Wetenschap

In dit hoofdstuk beschrijf ik een concreet stappenplan voor een mogelijke Responsieve Audit. Hierin beschrijf ik de vier stappen van de Responsieve Evaluatie, toegepast op deze specifieke casus. Bij iedere stap licht ik het doel, de werkzaamheden en rol als onderzoeker toe. Dit is het resultaat van literatuuronderzoek en verschillende gesprekken en focusgroepen. Dit heb ik toegelicht in de methodologische verantwoording.

Context: verantwoordingsmechanismen voor de culturele basisinfrastructuur

Met de culturele basisinfrastructuur wordt jaarlijks ruim 375 miljoen euro met Rijksgelden gesubsidieerd in de culturele sector. Binnen het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) is een tweetal directies hiervoor verantwoordelijk: de directies Erfgoed en Kunsten (E&K) en Media en Cultuur (M&C). De Raad voor Cultuur adviseert het ministerie over het cultuurbeleid en de subsidietoekenning. Zij heeft meermaals geadviseerd de huidige verantwoordingssystematiek te vereenvoudigen, vanwege de meer bureaucratiserende context (bijv., Raad voor Cultuur, 2022, p. 35). Momenteel leggen culturele instellingen naast hun bestuursverslag en jaarrekening ook verantwoording af met prestatie-indicatoren (Regeling op het specifiek cultuurbeleid, artikel 2.15). Vanuit de culturele sector is een roep deze systematiek aan te passen, maar onduidelijk is wat de grondredenen hiervoor zijn. Het ministerie heeft de Auditdienst Rijk (ADR) gevraagd hen te begeleiden in deze zoektocht om eveneens te komen tot mogelijke alternatieve wijze van verantwoorden. Samen met een ADR-collega ben ik in de eerste fase betrokken. Ik mocht een voorstel doen voor een plan van aanpak. Dat voorstel deel ik hieronder.

Uitwerking per stap van het plan voor de Responsieve Audit

Onderstaand licht ik het plan per fase nader toe. Voor een schematische weergave verwijs ik naar figuur 1 aan het einde van dit hoofdstuk.



Het doel van deze stap is inzicht te verkrijgen in de problematiek, geldende kaders en belanghebbenden door de ogen van het ministerie. Het eigenaarschap van dit traject ligt bij hen. Daarom is het belangrijk ambtenaren al in deze stap te betrekken. Naast hun betrokkenheid, is “*prolonged engagement*” bij de onderzoeker belangrijk (Abma, 2005, pp. 92-93). Door intensief kennis te maken met de context, creëert de onderzoeker begrip voor de leefwereld van het ministerie en de culturele instellingen. Eén van de onderzoekers heeft ervaring binnen het beleidsveld, de context van het ministerie en specifiek de twee directies. Daarnaast biedt deze verkennende fase een invulling van de *prolonged engagement* voor de andere onderzoekers.

Probleemdefinitie. De initiële probleemdefinitie met onderliggende vragen is het startpunt van het gezamenlijke onderzoek. Deze wordt in twee stappen geformuleerd. Allereerst worden verkennende

gesprekken gevoerd met ambtenaren bij het ministerie en andere (Rijks)partners, alsook een analyse van relevante documentatie verricht. Dit leidt tot een eerste versie van de probleemdefinitie die wordt gevalideerd en verrijkt tijdens de workshop om belanghebbenden te inventariseren (zie tabel 2).

Kaders om binnen te blijven. Vervolgens is het belangrijk om het kader waarbinnen mogelijke wijzigingen, verbeteringen en/of oplossingen moeten blijven, in kaart te brengen. Hiervoor wordt een analyse van relevante wet- en regelgeving uitgevoerd en verkennende gesprekken gevoerd met drie juridische en financiële experts binnen het ministerie. De rol van de onderzoekers is louter een kritisch-bevragende, op basis van hun analyse van wet- en regelgeving; de experts zijn leidend voor het vormgeven van de kaderstelling.

Workshops stakeholderanalyse. De inventarisatie van belanghebbenden is cruciaal, omdat hun (verschillende) verhalen, perspectieven en wereldbeelden de ruggengraat vormen van de aanpak. Het is een teken van het interactieve karakter, waarin werkelijk *samen* met de belanghebbenden de audit wordt gedaan. Dit staat in schril contrast met een audit waar *met gebruik van* belanghebbenden wordt gewerkt. Deze lijst van (groepen) belanghebbenden is initieel. Hun wordt namelijk in de volgende fase gevraagd welke (mogelijke andere) belanghebbenden belangrijk zijn om mee te nemen in dit traject (Abma, 2005). Het kan zijn dat de lijst dus nog wordt uitgebreid. Er worden twee sessies begeleid waarvoor ambtenaren worden uitgenodigd die relaties met de culturele instellingen onderhouden, omdat zij een goede inschatting kunnen maken van belanghebbenden. Onderstaande tabel geeft op hoofdlijnen de sessie weer. Indien de workshops niet leiden tot eenduidige groepen, worden de resultaten met de contactpersonen en gedelegeerd opdrachtgever besproken.

Workshoponderdelen

Probleemdefinitie. Deelnemers noteren welke woorden zij herkennen, willen wijzigen, toevoegen of verwijderen. Deelnemers lichten hun bijdragen toe en voeren hierover het gesprek. Wijzigingen worden doorgevoerd. De ADR-begeleiders bewaren de verschillende versies van de probleemdefinitie.

Individuele stakeholders beschrijven en beoordelen. Deelnemers benoemen zoveel stakeholders als ze maar kunnen en kwalificeren deze met een vierpuntsschaal langs de assen *belang*, *bevoegd*, *bekwaam* en *bereid* (vgl. Balane, e.a., 2020).

Stakeholders groeperen en selecteren. Deelnemers lezen alle stakeholders met kwalificatie; bespreken verschillen; en komen tot een groepering. Daarbij beschrijven zij op alle assen wat de stakeholders bindt *ten aanzien van* de subsidieverantwoording.

Tabel 2. Onderdelen workshop stakeholderanalyse.



Deze fase schetst de eerste contouren van de verhalen rondom de huidige en gewenste verantwoordingsystematiek. Het interactieve karakter krijgt meer invulling. De individuele gesprekken inventariseren de persoonlijke, concrete ervaringen, perspectieven en zorgen (Guba & Lincoln, 1989; Abma, 2005). Deze verhalen zijn weergaven van de leefwereld van belanghebbenden: “[P]ersonal narratives are useful for bringing the hidden into view, and making the contradictions, complexities, and ambiguities associated with multiple truths more accessible [...]” (Burgess-Limerick & Burgess-Limerick, 1998, p. 64).

Qua methodiek werk je niet met – voor velen gebruikelijke – topic interviews, waarin verschillende, van tevoren geselecteerde onderwerpen leidend zijn. De punten die betrokkenen naar voren brengen zijn belangrijker; hun belevingen en ervaringen van concrete gebeurtenissen staan centraal; dit is wat Abma en Widdershoven (2005) bedoelen met “*conversational interviews*” (p. 93). Daarom werken we met één startvraag: “Kun je me meenemen naar de laatste keer dat je de verantwoording voor de subsidie hebt moeten opstellen?” Deze manier van gesprekvoering vereist technieken die ook in de *behavioural audit* worden toegepast.

Hiervoor gelden allereerst bekende standaardtechnieken (bijv., hummen, bevestigen en bruggetjes). Peters (2023) benoemt eveneens de nuttige techniek van circulair bevragen, gebruikt onder systeempychologen. Ten slotte beschrijft hij hoe Dilts’ logische niveaus (i.e., omgeving, gedrag, competentie, overtuigingen, identiteit en connectie) kunnen worden gebruikt om ervoor te zorgen dat de interviewer “alle niveaus evenredig [bestrijkt] en probeert ze in hun samenhang te begrijpen” (Peters, 2023). Deze methoden zijn dezelfde als in een behavioural audit. Het verschil zit in het feit dat de interviewer bij een responsieve aanpak geen neutrale luisteraar is, maar een actieve deelnemer die ook diens visie of begrip kan inbrengen (Abma, 2003).

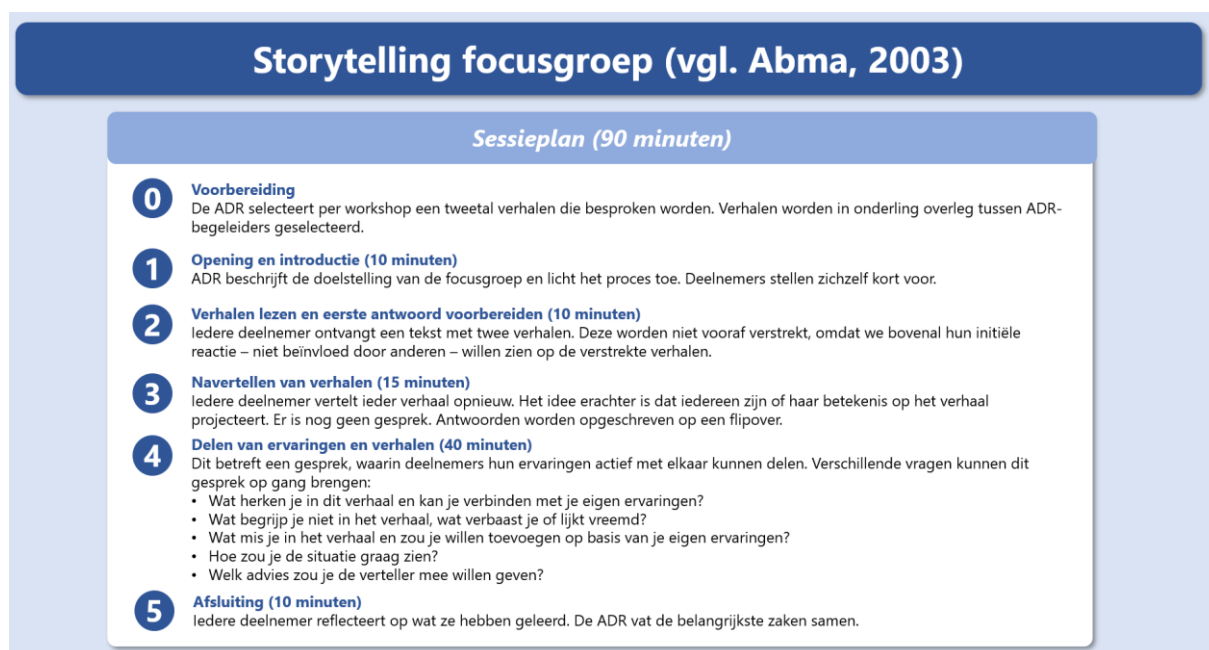
De gesprekken worden met meerdere belanghebbenden per groep gevoerd. Het selectie criterium is maximale variëteit, versus het gangbare criterium van representativiteit (Abma, 2005). Het aantal gesprekken is daarmee van tevoren onbekend, omdat je gesprekken blijft voeren tot saturatie is bereikt: data wordt meermaals herhaald en nieuwe inzichten wegblijven (Saunders, e.a., 2018).

De gesprekken worden opgenomen en uitgewerkt tot verhalen in verslagvorm die zo dicht mogelijk bij het taalgebruik van de betrokkene blijven. Maar: “*The meaning of stories is not given, and this implies that we [have] to interpret them.*” (Abma, 2003, p. 227). Daarom wordt dit verhaal voor hoor/wederhoor teruggelegd. Het is immers het verhaal van de betrokkene, niet van de onderzoeker. Pas na goedkeuring analyseren de onderzoekers deze verhalen inductief met thematische analyse en destilleren zo onderliggende thematieken, een interpretatieve analyse van thema’s in verhalen (Roman & Waterschoot, 2018). Daarnaast wordt ook een narratieve analyse uitgevoerd naar de structuur van het plot, naar de geconstrueerde relationele verhoudingen (Abma, 2001; Czarniawska, 2004; Soderberg, 2006; Sools, 2012; Van der Arend, 2007). Daarom is het niet alleen een analyse van *wat* maar ook *hoe* het verhaal wordt verteld. Dat is ook het voornaamste verschil met de *behavioural audit*.



In deze fase worden homogene focusgroepen georganiseerd, zodat belanghebbenden binnen hun homogene groep kennis kunnen maken met elkaars verhalen en onderliggende thematieken. Alle betrokkenen vanuit de gespreksfase worden hierbij betrokken. Verhalen kunnen herkenning oproepen, tot verbazing spreken of volledig verkeerd vallen. Alles is mogelijk. De storytelling focusgroepen richten zich primair op verdieping en verrijking van de verschillende verhalen (Abma, 2005; Visse e.a., 2011). Deze focusgroep draagt in meerdere vormen bij aan leerprocessen binnen organisaties: meerdere perspectieven verschijnen; taboes kunnen worden besproken; reflectie wordt gestimuleerd; een cultuur van nieuwsgierigheid door vragen te stellen wordt bevorderd; en er wordt expliciet stilgestaan bij onderliggende normen, waarden en standpunten (Abma, 2003). De dialoog onder belanghebbenden is een aanvullende vorm van validatie van de verhalen door te bezien in hoeverre verhalen worden herkend. In meer interveniërende zin is dit de eerste confrontatie van

verschillende perspectieven onder deelnemers. Het uiteindelijke doel is om onder belanghebbenden de dialoog te faciliteren en onderling begrip te creëren. Dit kan leiden tot veranderingen in ervaringen, perspectieven of zelfs onderliggende wereldbeelden.



Figuur 1. Sessieplan storytelling focusgroep (vgl. Abma, 2003).

Qua voorbereiding kiezen de onderzoekers maximaal drie verhalen, waarbij wordt gelet op: spreiding van de thema's, aansluiting met de leefwereld van alle deelnemers, en mogelijke verschillen tussen verhalen die tot confrontatie kunnen leiden. Zie bovenstaande figuur voor een sessieplan (vgl. Abma, 2003). De focusgroep wordt opgenomen; er wordt een verslag geschreven, waarbij we zo dicht mogelijk bij het taalgebruik van de betrokkenen blijven. Dit verslag wordt voor hoor/wederhoor teruggelegd voor de eigen uitspraken van belanghebbenden en de weergave van de dialoog. Bij deze verslaglegging is belangrijk dat niet iedereen het inhoudelijk eens hoeft te zijn over elkaars uitspraken. Immers: dat zou leiden tot een afgewaterde vorm van consensus; hetgeen deze benadering juist wil vermijden. De diversiteit van perspectieven staat voorop. Het doel van deze hoor/wederhoor is te valideren of de interpretatie van de onderzoekers bij de focusgroep past bij de interpretatie van de betrokkenen (Abma, 2005).



In de vorige fase de zijn belanghebbenden binnen de grenzen van hun eigen groep gebleven (verticale interactie); hier gaat het om de dialoog over de grenzen van de stakeholdergroepen heen (horizontale interactie). De dialooggroep lijkt op de storytelling focusgroep, maar verschilt qua doel. Waar de focusgroep zich richt op het uitdiepen van de verhalen en perspectieven, is het doel van de dialooggroep om de verschillende verhalen en perspectieven elkaar te laten ontmoeten over de groepen heen. Daarom vereist de dialooggroep een heterogene samenstelling. Door verschillende perspectieven via de verhalen bloot te leggen en hierover in oprechte dialoog te treden, ontstaat wederzijds begrip dat kan – maar niet moet – leiden tot consensus (Abma & Widdershoven, 2005).

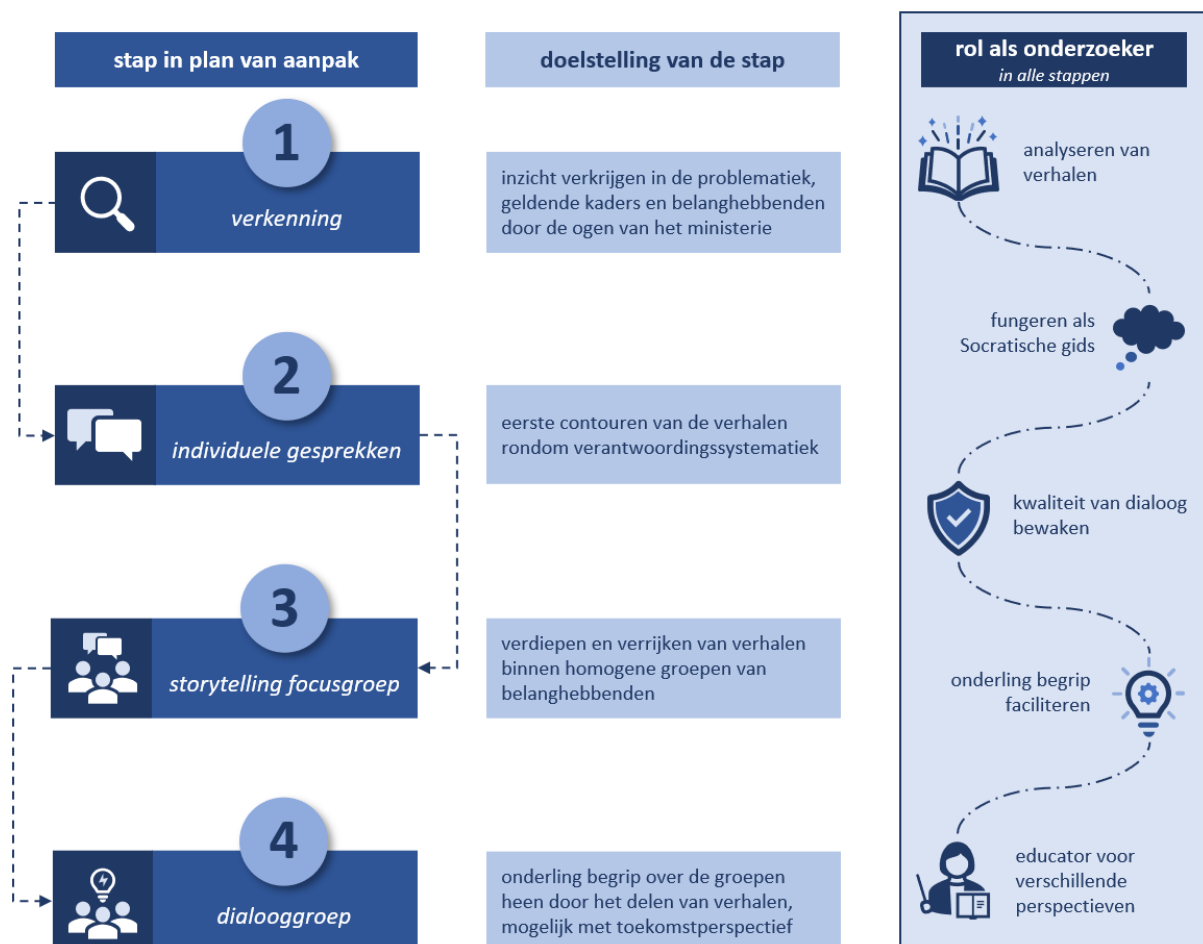
In de dialooggroep worden meerdere verhalen, aangevuld met eventuele citaten vanuit de focusgroep, besproken. Dezelfde selectiecriteria als bij de focusgroep gelden. In deze dialoog gaat het niet alleen om de verhalen *per se*, maar ook om de normen, waarden en aannames die daarachter schuilgaan. In die optiek wordt derde-orde-leren nagestreefd (vgl. Boonstra, 2000;

Werkman, 2006). Vanuit dat wederzijdse begrip kan vervolgens de dialoog over de toekomst van verantwoording bij culturele subsidies worden gestimuleerd. Dit is wat Moerkerken (2020) aanstipt met haar beschrijving van een functioneel conflict: “[W]e zijn het niet eens, maar laten onze verschillen naast elkaar bestaan en onderzoeken welke ideeën er dan ontstaan die recht doen aan allebei de opvattingen.” (p. 102).

Aangezien de dialooggroep met alle *groepen* van belanghebbenden plaatsvindt, en het omwille van de kwaliteit van de dialoog niet bevorderlijk is meer dan tien personen uit te nodigen, kunnen belanghebbenden aan het einde van iedere homogene focusgroep aangeven wie hen wil vertegenwoordigen in de heterogene dialooggroep. Van de dialooggroep wordt op dezelfde manier een verslag gemaakt en teruggeleefd voor hoor/wederhoor.

Het ministerie heeft de wens geuit geen lijvig rapport te hoeven ontvangen, omdat de beoogde interventie eerder plaatsvindt in de focusgroepen en dialooggroep, dan in geschrift. Hier wijken we af van het voorgeschreven rapport met *zgn. thick descriptions* (Abma, 2005; Geertz, 1973). Om echter ervoor te zorgen dat inzichten, de verschillende verhalen en andere uitkomsten niet alleen blijven circuleren onder deelnemers, wordt samen met de contactpersonen een samenvattend verslag geschreven van het co-creatieproces, dat ook ter validatie aan de belanghebbende partijen wordt voorgelegd.

Samenvattend kan het plan als volgt schematisch worden weergegeven:



Figuur 2. Samenvattend beeld onderzoeksplan door auteur.

3

Reflectie op de (on)mogelijkheid van Responsief Auditen

Dit hoofdstuk vormt het inhoudelijke slotstuk van deze bijdrage. In dit laatste hoofdstuk zet ik een reflectie uiteen over de (on)mogelijkheid van Responsief Auditen. In onderstaande paragrafen geef ik aan wanneer ik steun op inzichten uit focusgroepen en gesprekken; en wanneer ik teruggrijp op inzichten vanuit (wetenschappelijke) literatuur.

Dat brengt me bij het slotstuk van deze zoektocht. Waar de vorige gedeelten zich richtten op de mogelijkheden van een andere manier van auditen – en een concreet plan van aanpak als resultaat – vormt dit hoofdstuk het sluitstuk. Daarbij staat centraal of deze aanpak geschikt is voor een IAF, en specifiek voor de ADR.

Pragmatische zijde van onderzoeksvaardigheden en -houding

Pragmatisch gezien werd vanuit de focusgroep en individuele gesprekken benoemd dat er allereerst onderzoekers nodig zijn die een specifieke (en wellicht voor sommigen een nog onbekende) set aan onderzoeksvaardigheden bezitten. Het vereist technieken die behavioural auditors bekend zijn (vgl. Van der Meulen & Otten, 2013a, 2013b). Het vereist vraagtechnieken zoals gebruikelijk in Socratische gespreksvoering en -begeleiding; analysevaardigheden waarbij inductieve analyse een kernrol vervult; en een andere grondhouding die niet onbekend is voor degenen die als *trusted advisors* worden gezien (Abma & Widdershoven, 2011). Het omvat een besef dat je als onderzoeker niet in de rol kruipt van technisch expert – met een inhoudelijke boventoon – maar die van een procesbegeleider of -adviseur om kennis en inzichten te laten ontstaan. Met gerichte selectie hebben wij ervoor gezorgd dat het team beschikt over deze onderzoeksvaardigheden en grondhouding. Allen beschikken over een sociaalwetenschappelijke achtergrond; enkelen hebben ervaring met leergeschiedenisonderzoeken; en enkelen hebben trainingen gevolgd in Socratische gespreksvaardigheidstechnieken.

Responsieve onderzoeks aanpak als auditaanpak?

Los van deze onderzoeksvaardigheden, komt een meer principiële vraag opborrelen of deze aanpak een als auditaanpak geldt. Binnen de focusgroep werd geponeerd, zoals een goed auditor betaamt, dat dit afhankelijk is van de definitie van audit. Zoals tabel 1 (p. 8) tentatief toont, kent de Responsieve aanpak de drie elementen van toetsend onderzoek, hetzij met andere grondslagen. Naast deze drie elementen, werd ook de definitie van internal auditing van het IIA (n.d.) aangehaald: “De internal auditfunctie helpt een organisatie haar doelstellingen te realiseren door op basis van een *systematische* en *gedisciplineerde* aanpak de doelmatigheid van de processen van governance, risicomanagement en beheersing te *evalueren* en te verbeteren.” De drie cursief gemaakte elementen laten zien dat de Responsieve aanpak past binnen deze systematische, gedisciplineerde wijze van evalueren. Toch biedt dit geen soelaas om de aanpak als auditaanpak te bestempelen. Onderstaand schets ik de contouren van een reflectie.

Zoals eerder geschetst, stoelt de responsieve aanpak op andere (lees: subjectivistische) beginselen van onderzoek. Ik heb geprobeerd om deze responsieve beginselen te verenigen met de componenten van een auditaanpak (namelijk: norm, situatie en oordeel). Immers: deze kerncomponenten definiëren toetsend onderzoek; en audit is praktijkgericht toetsend onderzoek.

Toch blijkt dat deze componenten *an sich* niet doorslaggevend zijn voor het bepalen wat tot de categorie audit behoort. Deze regels horen thuis binnen een logica; je kunt het zien als grammatica.

Binnen de ADR zijn er een tweetal elementen die duiden op de dominantie van een objectivistische logica, wanneer we de regels rondom het bestempelen van een auditaanpak willen bezien. Dat zijn onafhankelijkheid en objectiviteit. Deze kernprincipes blijven regievoeren op de regels van het bestempelen van een auditaanpak. Betrokkenen delen onafhankelijkheid een plaatselijke dimensie toe: het gaat om afstand tussen de onderzoeker (lees: auditor) en degene of datgene die object is van onderzoek.² Oftewel: het onderscheid tussen object en subject is eenduidig. Vervolgens kan de onafhankelijke onderzoeker met gepaste objectiviteit het onderzoek uitvoeren én een *eigen* oordeel kan geven. Daarin moet deze, zo werd gesteld, onderscheid kunnen maken tussen feiten, percepties en meningen. Daarmee wordt de objectieve werkelijkheid van feiten onderscheiden van de subjectieve werkelijkheden van percepties en meningen.

Deze grondbeginselen zijn bepalend voor de objectivistische logica. De grammaticaregels van de logica bepalen dat een toetsende aanpak die niet past binnen het vocabulaire van onafhankelijkheid en objectiviteit c.q. objectiviteit, niet als een auditaanpak kan worden bestempeld. Daarmee is de invulling van deze beginselen *toegepast* op de drie elementen van toetsend onderzoek noodzakelijk zijn voor het bestempelen van een onderzoeks aanpak tot een auditaanpak; niet de drie elementen van toetsend onderzoek *per se*. Daarom kan ik ook niet anders concluderen dan dat het – binnen de huidige logica binnen de ADR – onmogelijk is om te spreken over een responsieve audit.

Responsieve aanpak binnen de ADR als IAF?

Open staat nog of deze aanpak wel binnen het palet van IAF-activiteiten past. Binnen de focusgroep en gesprekken werd teruggerepen naar de elementen van “meerwaarde” en “[de] organisatie te verbeteren” uit de definitie van internal auditing. Uiteindelijk is dat het ultieme doel van iedere IAF-activiteit. Dit is geen nieuwe stelling. Verschillende publicaties hebben hier reeds aandacht aan besteed: een internal auditor kan meer doen dan alleen auditen, in de zin van toetsend onderzoek (Bos, De Korte & Otten, 2017; De Korte, Otten & Schuiten, 2021; IIA, 2009; Van Twist & De Korte, 2022). Daarentegen kunnen ook gereede twijfels geplaatst worden bij het labelen van activiteiten als ‘audit’. Reeds in de jaren ‘90 waarschuwde Powell (1999) voor de auditexplosie en het auditbaar maken van steeds meer elementen in een organisatie.

De IIA-waaier (2009) illustreert bovendien welke typen activiteiten een interne auditfunctie kan leveren om meerwaarde te stimuleren. Daarbinnen past deze aanpak – volgens betrokkenen – onder andere vanwege de veronderstelling van een meervoudige realiteit. Zo gaf één van de geïnterviewden aan: “Er doen zich momenteel nog weinig kwesties voor, waarin verschillende betrokkenen voor eenzelfde realiteit staan; vraagstukken zijn complex en ambigu.” Het is precies voor deze vraagstukken dat een responsieve methode met storytelling geschikt is, zo stelt ook Abma (2006). In die context herhaal ik nogmaals dat het doel is te komen tot wederzijds begrip. Vanuit dat wederzijdse begrip kan (eventueel) worden gezocht naar toekomstrichtingen, recht doende aan ieders werkelijkheid. De term functioneel conflict van Moerkerken (2020) herhaal ik in dit kader nogmaals.

Er is dus wel degelijk ruimte voor deze, op subjectivistische aannames gestoelde aanpak. Opvallend is dat de – allicht rigide – grammaticaregels van de objectivistische logica niet in volledigheid doorklinken voor alle activiteiten binnen de IAF. Deze stelling laat zien dat er niet alleen ruimte is voor een andere dan de traditionele auditaanpak om van meerwaarde te zijn, maar dat deze aanpak

² Opvallend is dat deze constructie afwijkt van de definitie van het IIA, waarbij onafhankelijkheid wordt toebedeeld aan de ophanging van de IAF binnen de eigen organisatie. Het laat wederom zien dat het concept ‘onafhankelijkheid’ een – binnen auditland – ambigu concept is. Voor een nadere analyse hiervan verwijs ik graag naar Nordin (2022).

ook gebaseerd kan zijn op andere dan – en dus los kan staan van – de fundamentele principes die regeren binnen de hiervoor genoemde objectivistische auditlogica. Er is de ruimte om weg te stappen van het ideaal van een eenduidige werkelijkheid, die objectiveerbaar kan worden gemeten of vastgesteld. Er is de ruimte om in nabijheid en interactie met de organisatie tot inzichten te komen en daarin meer een faciliterende, educatieve, Socratische rol te vervullen in plaats van een inhoudelijke, objectieve expertrol.

Einde in zicht?

Aanvankelijk ben ik gestart met de vraag hoe een andere wijze van auditen die meervoudige werkelijkheden veronderstelt, de gereedheidskist van de internal auditor kan verrijken. Uit mijn literatuuronderzoek bleek dat een andere manier van onderzoeken mogelijk is en dat deze binnen de evaluatiegemeenschap wordt toegepast met de Responsieve Evaluatie. Reflecterend blijkt echter dat het niet mogelijk is om een wijze van *auditen* toe te passen die een meervoudige werkelijkheden veronderstelt. De objectivistische auditlogica waarin de beginselen van onafhankelijkheid en objectiviteit regievoeren, botsen met de grondbeginselen van een responsieve aanpak – en in het verlengde daarvan met iedere aanpak die een meervoudige werkelijkheden veronderstelt. Daarentegen is er wel ruimte om deze responsieve aanpak toe te passen *binnen de variëteit van activiteiten van de IAF*. Deze omvatten immers meer dan alleen auditen. Deze aanpak vereist wel specifieke onderzoeksvaardigheden die een *behavioural auditor* toepast en een fundamenteel andere grondhouding die de *trusted advisor* als muziek in de oren klinkt.

Dit antwoord gaat gepaard met een academische korrel zout en een kritische noot. In de methodologische verantwoording pak ik de korrel zout beet. De kritische noot uit ik in de vorm van een aantal vragen, waarin ik in zekere zin ook de kritiek van Eulrich en Lenz (2020) herken. In hoeverre is het nog mogelijk (en raadzaam) om verschillende activiteiten toe te voegen aan die van de IAF? Floreren we dan in verscheidenheid of verliezen we focus? In hoeverre is het de taak van de internal auditor nog om die aanvullende activiteiten uit te voeren? Of nog sterker: zijn we internal auditors – en daarmee gehouden aan beroepsregels – die ook andersoortige onderzoeken uitvoeren en consultancytaken kunnen vervullen? In hoeverre moeten internal auditors daarin rolvast blijven vasthouden aan de principes van onafhankelijkheid en objectiviteit? Of zijn we meer breed georiënteerde onderzoekers en adviseurs die een veelheid aan activiteiten uitvoeren, waaronder internal audits, en ook alleen bij die producten als internal auditor fungeren? Ik heb hier (nog) geen antwoord op; hiervoor is verder onderzoek nodig.

Afsluitend: deze zoektocht stopt hier. Gestart met een nieuwsgierigheid naar fundamenteel anders auditen, laat deze bijdrage zien dat dit, op dit moment, binnen deze specifieke context, (nog) niet mogelijk is. Desalniettemin is deze aanpak uitermate geschikt om te gebruiken voor ambigue vraagstukken, waar internal auditfuncties voor kunnen worden gevraagd. Met het schrijven van deze bijdrage kijk ik terug op een intensieve, maar leerzame periode met reflectieve cycli en relevante uitkomsten. Bovenal hoop ik jou meegenomen te hebben in deze zoektocht en misschien te hebben geïnspireerd voor het uitproberen van deze aanpak. Als je geïnteresseerd bent in de achterliggende theorie en methodologie, kun je door gaan met lezen in het tweede deel van deze bijdrage: de methodologische verantwoording.

TWEEDE DEEL

Methodologische

Verantwoording

Methodologische verantwoording met vijf reflectieve cycli

Je bent aangekomen bij de methodologische verantwoording van de zoektocht. Waar het andere onderdeel van het referaat zich richtte op de uitkomsten en overkoepelende reflectie, beschouw ik in deze verantwoording hoe ik daartoe ben gekomen. Op deze manier zet ik uiteen hoe ik ervoor heb gezorgd dat dit kwalitatieve onderzoek ook van kwaliteit is en dat ik me terdege bewust ben van de beperkingen die dit onderzoek inherent met zich meebrengt.

Reflecterend op dit onderzoekstraject werd me iets helder. Waar de Responsieve Evaluatie wordt gekenmerkt door een organisch ontwerp, was dat bij mijn verkenning ook het geval. Na de reflectie zag ik namelijk in dat ik telkens een onderzoeksvraag heb gesteld en die heb uitgediept. Vervolgens leidde dat tot een antwoord, met daaraan direct een nieuwe onderzoeksvraag gekoppeld. Zo heb ik het proces achteraf kunnen reconstrueren tot vijf cycli. In deze methodologische verantwoording neem ik je dan ook mee aan de hand van deze vijf cycli (zie figuur 3).



Figuur 3. Weergave reflectieve cycli, met bijbehorende onderzoeksvragen en methoden door auteur.

Ik werk in deze verantwoording met een zgn. twekolommenrapportage. Aan de linkerkant geef ik daarin inhoudelijk antwoord op de **vragen** die ik me in de cyclus heb gesteld. In de rechterkolom neem ik jou als lezer mee in het onderzoeksproces, zodat je je een beeld kan vormen van mijn ervaringen tijdens deze verkenning. Zo probeer ik een uitgebreide beschrijving te geven van het onderzoeksproces. Het is daarmee tevens een logboek waarin ik beschrijf hoe bepaalde gesprekken, lezingen of andere onderzoeken mij hebben geïnspireerd en beïnvloed. Dit is één van de manieren waarop ik **invulling geef aan de kwaliteitscriteria van deugdelijk kwalitatief onderzoek** (Bakker & Boer, 2021; Lincoln & Guba, 1985). Daarnaast gebruik ik deze rechterkolom om **concepten** nader te duiden in kaders. De tekst zal de kleuren hebben die ik hier ook heb gebruikt.

Eerste cyclus: het plan (op hoofdlijnen) voor een afstudeeronderzoek

Deze eerste cyclus was ik bijna vergeten toen ik deze methodologische verantwoording aan het schrijven was. Maar het is wel het startpunt geweest: deze verkenning die ik ben gestart in het kader van mijn afstudeerproces aan de studie *Internal Audit & Advisory* aan de Erasmus Universiteit Rotterdam. Vanuit dat startpunt heb ik een probleemstelling geschetst.

Van oorsprong is het uitvoeren van een audit dominant positivistisch ingeschoten (Dirsmith, e.a., 1985). De dominante aanpak is om aan de hand van een vooraf opgesteld normenkader een situatie waar te nemen en vervolgens deductief te analyseren, om zo te komen tot conclusies, dan wel oordelen. Bij positivistisch onderzoek wordt een sociale waarheid verondersteld die objectief kan worden waargenomen met gecontroleerde experimenten.

Uit mijn eigen praktijk herken ik dat auditors en onderzoekers werkend in het auditvak voornamelijk naturalistische methoden van dataverzameling gebruiken, zoals interviews en documentenanalyse (vgl. Lincoln en Guba, 1985; Beuving & de Vries, 2015). Op methodisch niveau worden dus al naturalistische (versus positivistische) methoden toegepast. Op methodologisch niveau lijkt het of audit dominant positivistisch blijft ingeschoten in het vaststellen van een objectieve sociale waarheid. Kijkend naar sociaalwetenschappelijke onderzoeksmethodologie zijn er ook andere wereldbeelden.

Centraal binnen onderzoeksmethodologie is de vraag wat kennis legitiem maakt. Onderzoek (en audit als een specifiek soort onderzoek) heeft tot doel het genereren, bevestigen, dan wel ontkrachten van kennis, informatie en inzichten (Verschuren & Doorewaard, 2021). Ik zie audit als een specifiek soort onderzoek dat binnen een bepaalde, organisatorische context kennis genereert. Immers: met een klassieke assurance-audit (als vorm van toetsend onderzoek) kan aanvullende zekerheid worden geleverd, waarmee een bepaalde situatie kan worden bevestigd, dan wel worden ontkracht voor wat betreft het varen van een juiste organisatorische koers.

De vraag die bij mij naar boven kwam, is wat ervoor zorgt dat de inzichten vanuit een audit legitiem zijn. Het zou de bredere onderzoeksgemeenschap teniet doen deze vraag te beantwoorden dat dit volgt vanuit toetsend onderzoek. Er zijn immers andere soorten onderzoek naast de toetsende vorm, die eveneens legitieme inzichten kunnen verschaffen. Denk hierbij bijvoorbeeld aan onderzoek dat een inductieve onderzoekslogica volgt. Hierbij wordt geen analyse- of normenkader opgesteld waaraan de werkelijkheid wordt getoetst, maar wordt er gewerkt met een meer open blik ten opzichte van een onderzoeksobject. Zelfs als auditors die eerst vooronderzoek doen en vervolgens zelf een normenkader maken en deze afstemmen met een opdrachtgever, is er nog steeds sprake van

Van huis uit ben ik geen bedrijfskundige, auditor of accountant; ik heb studies gevolgd in internationale betrekkingen, bestuurs- en organisatiewetenschap en Europese publieke zaken. Alle studies met een achtergrond in de sociale wetenschappen. In deze studies ben ik geconfronteerd met wetenschapsfilosofie en de notie van meerdere werkelijkheden. Vanuit die notie heb ik zowel mijn bachelor als masterscriptie geschreven, met daarbij discoursanalyse als centrale pijler.

Toen ik vervolgens bij de Europese Rekenkamer en later bij de ADR ben gaan werken, werd ik blootgesteld aan de wereld van auditen, die van toetsend onderzoek. Ik trok gelijk parallellen met positivistisch onderzoek. Een manier van onderzoek waar ik me altijd tegen heb verzet in mijn studietijd.

En plots had ik een interessant onderwerp te pakken voor een verkenning in de context van het afstudeeronderzoek: is er een fundamenteel andere manier van auditen mogelijk, die de principes van positivistisch auditen loslaat?

Na een korte literatuurstudie ben ik toen terecht gekomen op het magnum opus van Guba en Lincoln (1989). Ik herkende de beginselen van sociaal constructivistisch onderzoek en wilde m'n tanden daar verder in gaan zetten. Zo is het initiële plan voor mijn afstudeeronderzoek ontstaan.

Daarbij moet gelijk gezegd worden dat het altijd een plan is geweest, waarvan afgeweken kan worden. Dat heeft mijn begeleider me ook op het hart gedrukt. Een voorbeeld: de scherpe lezer ziet ongetwijfeld al dat de vraag die ik hier aan het einde van de eerste cyclus heb gesteld, anders is dan de vraag

een onderzoek dat een deductieve logica volgt. Immers: met een vooraf opgesteld c.q. afgestemd normenkader (lees :een gesloten blik) wordt een onderzoek verricht.

die ik uiteindelijk centraal heb gesteld in de vorige bijdrage. Daarover meer in de tweede cyclus.

Er zijn dus – redenerend vanuit sociaalwetenschappelijke onderzoeksmethodologie – ook andere manieren van onderzoek doen, die legitieme inzichten op kunnen leveren. Manieren waarbij we de traditionele manier van het werken met een normenkader achterwege laten. Toch kijken we daar als auditors tot op heden nog onvoldoende naar. Daarmee doen we niet alleen onszelf, maar bovenal de opdrachtgevers tekort. Want andere manieren van onderzoek, kunnen andere – allicht beter passende – inzichten opleveren.

Hieruit volgend heb ik de volgende, initiële onderzoeksvraag gedefinieerd: **“Hoe kan de constructivistische onderzoeksmethode van de vierde evaluatiegeneratie bijdragen aan de onderzoeksmethoden in de gereedschapskist van de internal auditor?”**

Tweede cyclus: Audit bezien vanuit wetenschappelijke literatuur

Langer stilstaand bij de definitie van auditing is het enigszins frappant dat in de literatuur, noch in praxis een eenduidige, breed gedragen definitie van auditing bestaat; het is niet een vastomlijnde activiteit. Daarmee bedoel ik de enorme verscheidenheid aan werkzaamheden te duiden. Er zijn interne auditors die audits uitvoeren voor het management en bestuur en externe auditors die audits uitvoeren voor het maatschappelijke verkeer (Bos, De Korte & Otten, 2017). Internal auditing omvat zowel audit- als advies- of consultancydiensten (IIA, n.d.). Naast dit onderscheid in type audits, is er ook een verschil in auditdisciplines. Grofweg biedt een organisatie een drietal verschillende smaken aan: financial, IT- en/of operational audits. De definitie van audit die ik hier hanteer, ontleen ik aan Bos, De Korte en Otten (2017), die stellen dat audit toetsend onderzoek is.

Audit omvat een te onderzoeken realiteit met een normenkader leidend tot een oordeel

Waar deze verschillende typering met name het verschil tussen de audits duiden, wil ik in deze bijdrage me voornamelijk richten op het middel dat alle audits bindt. Nauwkeuriger gezegd: de drie middelen – ingrediënten – die in alle audits naar voren dienen te komen. Deze drie ingrediënten zijn als een robuust raamwerk te zien: een waar te nemen realiteit, een vooraf opgesteld normenkader, en een oordeel (vgl. Bos, De Korte & Otten, 2017; Van Twist & De Korte, 2022). De vergelijking van de realiteit en het normenkader leidt uiteindelijk tot het oordeel van de auditor. Iedere audit omvat deze drie ingrediënten.

Bos, De Korte en Otten (2017) stellen dat een audit louter de eerste twee aspecten hoeft te omvatten, en dat een toevoeging van het component 'oordeel' leidt tot de typering van een assurance-onderzoek, zoals dat vereist is conform de regels die gelden voor financial auditors (vgl. De Korte, Otten & Schuiten, 2021). Daarentegen wordt er nog steeds een vergelijking gemaakt van de te auditen situatie en het te hanteren normenkader, en worden verschillen geduid. Dit wordt dan een bevinding genoemd.

Auditing als een objectieve activiteit– waarheid, ideaal, of utopie?

Ik poneer een stelling: 'Auditing is een objectieve activiteit'. Deze stelling ga ik met drie brillen bekijken, elk met een andere blik qua wetenschapsfilosofische aannames.

*Deze tweede cyclus was de eerste échte activiteit die ik ondernam voor de verkenning. Ik vond het belangrijk om bij het begin te beginnen: **wat is audit?** Zoals het hiernaast beschrevene laat zien, is dat geen eenvoudige vraag.*

Ik had op dit moment een aantal auditvakken gevolgd en werd geïntroduceerd met de definitie van audit als toetsend onderzoek, waarbij een norm met een situatie wordt vergeleken, hetgeen leidt tot bevindingen. Dat gaf houvast in de verdere verkenning.

Daar werd ik ook op gewezen in een gesprek met Mark van Twist. Hoewel dat niet het onderwerp van gesprek was (daar kom ik nog op terug), gaf hij toen de suggestie om het literatuuronderzoek vorm te geven aan de hand van de drie ingrediënten van een audit: een realiteit, een norm en een oordeel. Later kwam ik deze driedeling ook tegen in een artikel van Mark van Twist en Ron de Korte (2022).

*De volgende vraag kwam opborrelen: **Wat is de invulling van verschillende wetenschapsfilosofische aannames op de drie componenten van een audit?** Toen ben ik gaan 'spelen' met theorie; en zoals een kleuter ook tig keer omvalt in de zandbak, ben ik ook vaak gestruikeld over literatuur. Maar uiteindelijk ben ik gekomen tot een vergelijking – op hoofdlijnen en niet recht doende aan de complexiteit – van auditing als activiteit, gelieerd aan wetenschapsfilosofische aannames. In het andere onderdeel heb ik dit samengevoegd als 'wereldbeeld'.*

Om dit weer goed voor de geest te halen, ben ik teruggegaan naar de eerste keer toen ik met deze begrippen werd geconfronteerd. Ik heb teruggerepen naar het artikel van Marsh en Furlong (2002) waar zij dit haarfijn uitleggen. Onderstaand kader is op hun werk gebaseerd.

Auditing is per definitie objectief. Objectiviteit zit in alles en iedereen. Subjectiviteit is een taboe.

Een klassieke audit is gestoeld op de kernwaarden die in de sociaalwetenschappelijke literatuur wordt omschreven als positivisme: een paradigma gestoeld op objectivistische **ontologische, epistemologische** en methodologische grondslagen (Van Grinsven, 2021). Met andere woorden: onafhankelijke onderzoekers kunnen een fenomeen op afstand waarnemen en op basis van objectief onderzoek komen tot wetmatigheden m.b.t. het specifieke fenomeen. Het paradigma ziet de onderzoeker als een objectieve expert, die – op afstand – kan onderzoeken (Moses & Knutsen, 2012). De aanpak is veelal deductief: op basis van vooraf opgestelde hypothesen (lees: normen en verwachtingen). Er wordt hier ook veelal gesproken over twee werelden: de werkelijkheid van de onderzoeker die op afstand staat van de werkelijkheid die deze wil onderzoeken. Alle drie ingrediënten van een audit kunnen dus als objectief kunnen worden gezien: er bestaat een objectieve realiteit die we als onderzoekers met een objectief normenkader kunnen waarnemen, om te komen tot objectieve conclusies. Maar is dit wel zo realistisch?

We maken audit zo objectief mogelijk. We zijn niet volledig objectief, maar laten we ons best doen objectiviteit na te streven.

De kritisch realist kijkt hier genuanceerder naar (Johnson & Duberley, 2000; Van Grinsven, 2021; Van Twist & De Korte, 2022). Deze veronderstelt dat het onmogelijk is om de wereld om ons heen honderd procent objectief waar te nemen. We blijven zelf ook mensen, die gevormd zijn met bepaalde opvoeding, culturen, scholing, die ons zichtbaar en onzichtbaar beïnvloeden in onderzoek. De sociale wereld – de wereld van menselijke interactie – is en blijft echter waar te nemen. Hier komt de metafoer van de olifant en de blindeman om de hoek kijken: we kennen het fenomeen van de olifant, maar kunnen deze slechts in beperkte mate waarnemen (Daigneault, 2013). In dat opzicht kunnen we spreken van een subjectieve epistemologie, gepaard met een objectieve ontologie: de kritisch realist weet dat er een werkelijkheid is om waar te nemen en weet tegelijkertijd dat de toegang om tot kennis te komen over die werkelijkheid inherent beperkt is.

Vertaald naar de auditpraktijk, zien we meer en meer kritisch realisten ten tonele komen, als we deze in vergelijking stellen tot de drie kerningrediënten van een audit. De kritisch realist beargumenteert dat het niet mogelijk is om tot een objectief oordeel te komen aan de hand van het onderzoek dat deze heeft uitgevoerd. Dat wil zeggen: het auditoordeel en oordeelsvorming zijn subjectief. Het oordeel en de conclusies komen subjectief tot stand, omdat je met verschillende onderzoekers werkt; allen met eigen denkbeelden, overtuigingen, vooronderstellingen etc. In de Engelse versie van de IIA-standaarden zien we deze gedachtegang ook terug: er wordt daar gesproken van een *auditor opinion*. Daarmee wil ik niet stellen dat de auditor een

Kader 1. Ontologie en epistemologie.

Ontologie betreft aannames over de werkelijkheid. Het wordt ook wel de “theory of being” genoemd (p. 18). Grofweg zijn daar twee keuzes in: objectivistisch (“foundationalist”) en subjectivistisch (“anti-foundationalist”) (p. 22-32). **Epistemologie** betreft de aannames over kennis en wat je als legitieme kennis beschouwt. Daarom wordt het ook een “theory of knowledge” genoemd (p. 18-19).

Daarnaast werd het artikel van Johnson en Duberley (2000) door mijn begeleider werd aanbevolen. Zijn advies weerklonk in het schrijfproces telkens in mijn achterhoofd: ‘Maak het begrijpelijk.’ Ik hoop dat ook voldoende te hebben gedaan met de samenvattende tabellen aan het einde. Ze hebben mij persoonlijk erg geholpen om overzicht te houden in het oerwoud van wetenschapsfilosofische literatuur.

Later in mijn studie aan de EUR maakte ik kennis met de behavioural audit (zie kader 2), die ook een subjectivistische ontologie veronderstelt. En toen raakte ik in lichte paniek. ‘De aannames van de behavioural audit lijken verdacht veel op die van het sociaal constructivisme,’ ging er toen door mijn hoofd. ‘Al dat uitzoekwerk voor niets.’ Ik was toen namelijk al in de derde cyclus beland.

Kader 2. Behavioural audit.

Behavioural audit volgt de leergeschiedenismethode, zoals vormgegeven aan het MIT (Van der Meulen & Otten, 2013a, 2013b). Het vindt grondslagen in de cognitieve psychologie en veronderstelt subjectieve, individuele werkelijkheden. Die komen ten tonele via mentale modellen. Ik zie de behavioural audit dan ook als een onderdeel van de postmoderne paraplu (zie hieronder) en zie louter enkele accentverschillen met de voorgestelde aanpak in dit afstudeeronderzoek.

Na meer literatuuronderzoek ontdekte ik de stroming van het sociaal constructionisme (Burr & Dick, 2017; Cunliffe, 2008; Galbin,

nietszeggende mening projecteert in de beoefening van het beroep, maar dat een oordeel inherent subjectief is. Daarnaast ziet de kritisch realist dat het te hanteren normenkader inherent subjectief is. Dit zit immers al verweven in de terminologie zelf: een normenvrij normenkader is – behalve taalkundig onzuiver – volstrekt onmogelijk. Verder zijn er in het onderzoek nog tal van andere punten die de subjectiviteit onderstrepen: de gesprekken met de geïnterviewden, de analyse van resultaten, het schrijven van een rapportage, enzovoorts.

Bewustwording van deze subjectiviteit betekent niet dat de kritisch-realist deze ook accepteert. Het ideaal van objectiviteit moet worden nagestreefd. Daarom moeten maatregelen worden genomen om subjectiviteit te mitigeren en het onderzoek te objectiveren met triangulatie (vgl. Bos, De Korte & Otten, 2017): zorg voor meerdere onderzoekers – bij voorkeur met verschillende achtergronden en senioriteit – verschillende bronnen en verschillende methoden. Immers: ook tijdens bijvoorbeeld het veldwerk is subjectiviteit niet weg te denken. Met triangulatie kan dit grotendeels worden gemitigeerd. Hoewel oordelen nog steeds inherent subjectief zijn, zijn deze gestoeld op deugdelijke procedures die het onderzoeksproces zo veel mogelijk objectiveren. Kritisch realisten benadrukken de noodzaak van het van tevoren afstemmen van het normenkader met de opdrachtgever, met uitzondering van extern opgelegde normen – zoals voorgeschreven in wet en regelgeving (Bos, De Korte & Otten, 2017). Dan is er in ieder geval overeenstemming bereikt over de te hanteren subjectiviteit in de audit. Zo ontstaat dus in zekere zin geen omarming van subjectiviteit, maar eerder een acceptatie van intersubjectiviteit, omdat het ideaal van objectiviteit onhaalbaar is.

Auditing is subjectief en dat geeft inzicht. Objectiviteit is nooit haalbaar. Ook niet als je het nastreeft. Omarm subjectiviteit.

Ten slotte beschrijf ik een meer subjectieve bril. De bril die aangeeft dat objectiviteit in onderzoek niet bestaat – en dus een utopie is – op alle vlakken die hiervoor zijn benoemd. In de wetenschapsfilosofische literatuur zijn er allerlei verschillende **postmoderne** stromingen die zich in zekere mate aansluiten bij deze stellingname (Van Twist & De Korte, 2022). Ik ben me er terdege van bewust dat deze stromingen elk hun eigen kenmerken en accentpunten kennen. Ik beperk me tot twee specifieke stromingen: het **sociaal constructivisme** en het **sociaal constructionisme** (zie kader 3 en 5). Een vorm van audit die sociaal constructionistisch gedachtegoed volgt, is er nog niet. De eerste stroming zien we binnen het auditvak terugkomen in de vorm van de **behavioural audit** (zie kader 2; Van der Meulen & Otten, 2013a, 2013b).

Bovenstaand heb ik kort een drieluik geduid waarin auditing – gedefinieerd met de componenten realiteit, normenkader en oordeel – wordt gekoppeld aan grondslagen vanuit wetenschapsfilosofische literatuur. De samenvatting heb ik hieronder in een tabel weergegeven.

2014; Hosking, 2011; Shotter, 1995; Shotter, 2010). De verschillen met het sociaal constructivisme zijn niet groot (zie kader 3 voor de overeenkomsten en kader 5 voor de verschillen), maar leiden toch tot een ietwat andere aanpak.

Kader 3. Overeenkomsten sociaal constructivisme en -constructionisme.

Zowel **sociaal constructivisme** en **sociaal constructionisme** verwerpen het idee van een enkele realiteit die we onafhankelijk en objectief kunnen waarnemen vanuit de sociale wetenschappen.¹ Ze veronderstellen het idee van meervoudige werkelijkheden. Beide stromingen veronderstellen dat onderzoekers inherent subjectief zijn en dat onderzoekers het ideaal van objectiviteit moeten loslaten. En in dat opzicht kunnen we beide stromingen ook onder de noemer van het postmodernisme plaatsen.

Kader 5. Verschillen sociaal constructivisme en -constructionisme.

Het grootste verschil bevindt zich in hun blik op de aard van de verschillende werkelijkheden. Sociaal constructivisme heeft een meer individualistische (actor-georiënteerde) blik hierop: werkelijkheden zijn cognitief en worden per individu geconstrueerd. Sociaal constructionisme heeft een interactieve en dialogische blik van het fenomeen: werkelijkheden zijn relationeel en worden in interactie geconstrueerd. Boonstra (2000) vat sociaal constructionisme goed samen: het gaat om “sociaal construeren van betekenissen in directe interactie met anderen, waarbij meerdere stemmen ruimte krijgen in dialoog” (p. 22).

Kader 4. Postmodernisme.

Het **postmodernisme** is onderhand een paraplu geworden van verschillende stromingen die deze subjectivistische grondbeginselen delen, maar accentverschillen leggen in de ontologische, epistemologische en methodologische veronderstellingen.

	Auditing is per definitie objectief.	We maken audit zo objectief mogelijk.	Auditing is subjectief en dat geeft inzicht.
Realiteit	<i>objectief</i>	<i>objectief</i>	<i>subjectief</i>
Normenkader	<i>objectief</i>	<i>subjectief</i>	<i>subjectief</i>
Oordeel	<i>objectief</i>	<i>intersubjectief bij auditors</i>	<i>subjectief</i>
Stroming in de wetenschap	<i>positivist</i>	<i>kritisch-realist</i>	<i>sociaal constructivisme</i>

Tabel 3. Vergelijking van de auteur van drie stellingen, ingedeeld naar de ingrediënten van een audit.

Eveneens laat bovenstaand zien dat er kritisch-realistische auditors zijn die onderkennen dat ze niet een volledige beschrijving van de werkelijkheid kunnen geven en zich bewust zijn van hun limieten, maar toch hun best doen een zo objectief mogelijk product op te leveren (Bos, De Korte & Otten, 2017). Dit type audit en auditor heb ik onderstaand samengevat.

Audit zoals we dat doorgaans kennen...	
Ontologie	Audit wordt gekenmerkt door een objectivistische ontologie: er is één werkelijkheid die we objectief kunnen waarnemen. De behavioural audit is hierop de uitzondering. Deze veronderstelt dat realiteiten zich bevinden in mentale modellen (Van der Meulen & Otten, 2013a).
Epistemologie	Binnen auditland zijn er twee smaken qua epistemologische veronderstellingen. Objectivistisch: we kunnen objectief tot een oordeel komen op basis van deugdelijk onderzoek. Subjectivistisch: we zijn niet honderd procent objectief, maar doen ons best zo veel mogelijk te objectiveren om te komen tot betrouwbare auditbevindingen.
Methodologie	Veelal deductief, met een vooraf opgesteld normenkader. De behavioural audit als inductief onderzoek is hierop de grootste uitzondering.
Auditor	Afhankelijk van de gekozen epistemologie een onafhankelijke, objectieve onderzoeker (objectivistisch) of een expert bewust van de grenzen van diens kennis en vaardigheden (subjectivistisch).
Doel audit	Een audit geeft hoe dan ook inzicht in een bepaald fenomeen. Aan een audit wordt zekerheid ontleend, hetzij expliciet door het afgeven van assurance, hetzij impliciet door het laten uitvoeren van een onderzoek door een auditprofessional (Driessen & Molenkamp, 2012).

Tabel 4. Beschrijving van de auteur van een audit zoals we doorgaans kennen, ingedeeld op wetenschapsfilosofische kenmerken

Derde cyclus: Hoe audit en evaluatie dichter bij elkaar liggen dan je denkt

Audit en evaluatie als verwante vakgebieden

Audit en evaluatie zijn verwant; het zijn zeer aan elkaar gelieerde vakgebieden. Eerst duid ik echter een aantal verschillen. Ten eerste is de nadruk verschillend. Bij auditing ligt de focus meer op bedrijfsvoering, waar evaluatie zich meer richt op de daadwerkelijk opgeleverde producten c.q. beleidsvoering (Mayne, 2006). Daar wil ik wel bij opmerken dat dit verschil zeer miniem is, omdat beide activiteiten elkaar nodig hebben. Daarnaast geeft Clarke (1999) aan dat het verschil meer ligt in het soort vragen dat leidend is (Chelimsky, 1985). Bij auditing zijn de vragen nadrukkelijk toetsend en normatief, waar deze bij een evaluatie ook verklarend en beschrijvend kunnen zijn (Mayne, 2006; vgl. De Korte, Otten & Schuiten, 2021). Dit komt tevens overeen met de eerdere definitie met de drie componenten van een audit, waar een feitelijke situatie aan een normenkader wordt getoetst. Wel dient daarbij de opmerking geplaatst te worden dat een evaluatie nagenoeg altijd tot doel heeft om de (toegevoegde) waarde van een bepaalde activiteit dan wel product te duiden c.q. te beoordelen (Guba & Lincoln, 1989).

Afgezien van deze relatief kleine verschillen, zijn er ook overeenkomsten tussen de vakgebieden van evaluatie en auditing. Allereerst kunnen beide vakgebieden worden geclassificeerd als een soort toegepast onderzoek. Dat wil zeggen dat het geen onderzoek is dat een hoger wetenschappelijk doel dient, maar onderzoek ter bevordering van een specifiek beleidsdomein, of bedrijfsvoeringsaspect. Ze hebben tot doel het bevorderen van lokale kennis en lokale theorie (vgl. Geertz, 1983; Cox, 1981). Oftewel: internal auditing en evaluatie hebben niet het ontwikkelen van metatheoretische veronderstellingen tot doel, maar het produceren van kennis die specifiek genoeg is voor een bepaalde situatie. In dat opzicht werken beide vakgebieden dus ook vaak met gevalstudies (Francis, 1994). Ten tweede zijn zowel internal auditing als interne evaluatie gericht op het rapporteren aan het (senior) management van een organisatie (Bemelmans-Vidéc, e.a., 2017). Daarnaast zijn beide gebieden ook niet onbekend met het doen van uitspraken c.q. oordelen over de (toegevoegde) waarde van een bepaald auditobject, dan wel evaluatieobject. Deze kunnen een verbeter- of verantwoordingsdoel dienen, of een combinatie van beide (Mayne, 2006; Verwoerd, e.a., 2020).

Kortom: afgezien van verschil in focus en onderwerp, hebben beide vakgebieden tot doel het genereren van specifieke kennis binnen een specifieke context met praktijkgericht onderzoek.

Drie Generaties van Evaluatie

In hun magnum opus “*Fourth Generation Evaluation*” traceren Guba en Lincoln (1989) **drie evaluatiegeneraties** in programma-

*Het einde van de tweede cyclus was nabij. Ik wist vanuit een module beleidsevaluatie tijdens mijn masterstudie dat er verschillende stromingen waren, die getraceerd konden worden naar de hierboven geschetste aannames. Ik wilde dus over de schutting kijken bij de evaluatiegemeenschap. Eerst heb ik gekeken of dat überhaupt mogelijk was: **Wat zijn de overeenkomsten en verschillen tussen audit en evaluatie?** Immers: als de verschillen te groot zouden zijn, lag het niet voor de hand om inzichten van de evaluatiegemeenschap te gebruiken. Dat heeft geleid tot een nadere literatuurstudie.*

*Ik heb toen teruggegrepen op het magnum opus van Lincoln en Guba. Ik had alleen nog maar over het boek gelezen, maar ben het in deze cyclus van A tot Z gaan lezen. Ik wilde antwoord op de vraag: **Wat is de Vierde Evaluatiegeneratie en wat zijn voor- en nadelen (in de toepassing) ervan?***

evaluaties, voornamelijk in het onderwijs. De eerste generatie is de meetgeneratie en wordt gekenmerkt door het adagium 'meten is weten'. In de tweede generatie komt daar een component bij: beschrijvingen. Maar ook dit blijkt niet voldoende: er is behoefte aan een oordeel van de evaluator over de toegevoegde waarde van het evaluatieobject in relatie tot de gestelde doelstellingen en gebruikte middelen. Wederom zijn de gelijkenissen met het auditvak overduidelijk.

In navolging van Stakes (1975) responsieve evaluatiegedachtegoed, stellen Guba en Lincoln (1989) dat de tijd is aangebroken voor een paradigmawisseling naar een vierde evaluatiegeneratie, omdat eerdere generaties lijden aan een diepgewortelde kwaal: het blijven vasthouden aan het natuurwetenschappelijk ideaal van dé waarheid (lees: positivisme) en daarmee de uitsluiting van wat zij waardenpluralisme noemen. Daarom stellen zij een nieuwe evaluatiegeneratie voor, gegrond in de postmoderne onderzoekstraditie (Guba & Lincoln, 1989, p. 83).

Vierde evaluatiegeneratie gegrond in postmodernisme

De vierde evaluatiegeneratie kort samenvattend: met verdiepende, open gesprekken destilleren onderzoekers thema's die volgens betrokkenen belangrijk zijn. Onderzoekers spreken net zo veel mensen tot het punt van saturatie is bereikt; een kernpunt van inductief onderzoek (zie bijv. Saunders e.a., 2018; Sebele-Mpofu, 2020). Daarna vinden groepsgesprekken plaats over de constructies op de geïnventariseerde thema's. Die kunnen uiteenlopen en zo ontstaat er discussie. Het hoogst haalbare doel is een onderhandelde consensus over de constructies, zodat een gedeelde realiteit ontstaat (Guba & Lincoln, 1985).

Ontologische en epistemologische paradigmawisseling: realiteiten ontstaan met constructies en interpretaties

De vierde evaluatiegeneratie wordt gekenmerkt door een subjectivistische ontologie en epistemologie. Guba en Lincoln (1985) betogen daarmee een **naturalistische onderzoeksbenadering**. Daarmee wilden zij een paradigmawisseling inzetten.

Daarin staat niet de zoektocht naar dé waarheid centraal, maar ontdekt de evaluator de constructies en interpretaties van de belanghebbenden. Die constructies en interpretaties zijn immers wat realiteiten zijn en maken. Dit is het moment waar de positivist en kritisch-realist afhaken. Waar zij veronderstellen dat sociale fenomenen en relaties simpelweg bestaan, stelt de vierde evaluatiegeneratie dat deze interactief ontstaan. Oftewel: deze sociale fenomenen en relaties ontstaan door interactie, via taal: wat we zeggen, zien we (Van Twist, 2018; vgl. Chia, 1996; Fairclough, 1993; Milliken, 1999). Dat impliceert vervolgens: "there are as many constructions as there are constructors" (Guba & Lincoln, 1989, p. 144). Maar niet alle constructies zijn

Kader 6. Nadere toelichting eerste en tweede evaluatiegeneratie.

Binnen de eerste evaluatiegeneratie, wordt meten als afdoende gezien om een uitspraak te kunnen doen over de waarde van een evaluatieobject. De evaluator is binnen deze generatie een "unbiased technical expert able to administer the appropriate test and pronounce with authority the result and what is meant" (Laughlin & Broadbent, 1996, p. 434). Bij de tweede evaluatiegeneratie realiseert men zich dat alleen meten niet afdoende is; de evaluator dient de sterkten en zwakten van het evaluatieobject te beschrijven (Laughlin & Broadbent, 1996).

Kader 7. Naturalistische onderzoeksbenadering

Voor hun magnum opus hebben Lincoln en Guba (1985) een andere toonaangevende bijdrage gepubliceerd: 'Naturalistic Inquiry'. Hier betogen zij een naturalistische onderzoekswijze: onderzoek wordt binnen de natuurlijke (tegenover experimentele) context uitgevoerd (bijv. 'gewoon' op kantoor). De technieken veranderen: geen experimenten en controlegroepen, maar observaties en gesprekken. Het gaat volgens hen echter een stap verder dan louter de gereedschapskist en technieken aanpassen: het betreft een wisseling van mindset, gedachten, onderliggende patronen; een paradigmawisseling (vgl. Vanderplaat, 1995, p. 86).

verschillend; constructies en interpretaties kunnen worden gedeeld. Zo ontstaat een dominant geldende constructie, die door (een groep) personen als legitiem wordt geconstrueerd, beschouwd en in stand gehouden (Van Twist, 2018). De vierde evaluatiegeneratie wil die constructies en interpretaties ontdekken, blootleggen en begrijpen.

Methodologische paradigmawisseling: hermeneutiek en dialectiek als centrale pijlers

Guba en Lincoln (1989) classificeren hun methodologie als een *“hermeneutic/dialectic process that takes full advantage, and account of the observer/observed interaction to create a constructed reality that is as informed and sophisticated as it can be made at any particular point in time”* (p. 44, nadruk weggelaten). Hier staat een aantal belangrijke componenten. Allereerst: de evaluatie is óók een constructie van een realiteit (Vanderplaat, 1995, p. 90). Met een gestructureerde aanpak, waarin betrokkenen hun spreekwoordelijke ei op veel verschillende manieren kwijt kunnen, is deze geconstrueerde realiteit geïnformeerd en verrijkt.

Dan resteren nog twee centrale pijlers: hermeneutiek en dialectiek. Hermeneutiek – zoals voorgesteld door Gadamer (1975) – richt zich op **interpretatieve cycli** (Bernstein, 1983; Schwandt, 1999). Dialectiek is de tweede centrale pijler van de methodologie. Het betreft de juxtapositie van verschillende visies, meningen, constructies en interpretaties om zo te komen tot die geïnformeerde en verrijkte realiteit die de evaluatie poogt weer te geven. Die juxtapositie vindt plaats tijdens de groepssessies van de betrokkenen. Daar bevragen zij elkaar en elkaars constructies kritisch en daarbij geldt het adagium: **‘wat standhoudt, is waar’**. Oftewel: welke constructie aan het einde van die groepssessie de toets der kritiek hebben doorstaan, wordt beschouwd als de geldende constructie.

Kritiek richting aannames over macht, actie en consensus

De mogelijke paradigmawisseling in het evaluatievakgebied gaat niet gepaard zonder kritiek. Hieronder vat ik de twee belangrijkste kritische noten samen.

Een veel geuit kritiekpunt betreft het concept ‘macht’. Zo geeft Vanderplaat (1995) bijvoorbeeld aan dat de methode onvoldoende relatieve machtsstructuren in acht neemt. Er wordt verondersteld dat men op gelijke basis staat, of hoort te staan, in de evaluatie en dat dit – naïef gezegd – verder voor geen problemen zal zorgen (vgl. Gregory, 2000). Daar komt bij – zo beargumenteren Laughlin en Broadbent (1996) – dat de opdrachtgever het vetorecht behoudt over besluiten in de laatste fase. Daarmee blijft de uitwerking van het probleem rondom machtsstructuren niet alleen onderbelicht, maar wordt deze ook – met dit zgn. vetorecht – in stand gehouden.

Kader 8. Interpretatieve cycli.

Een eerste **interpretatieve cyclus** vindt plaats op het moment dat iemand een constructie, een interpretatie maakt van de realiteit. Een tweede cyclus vindt plaats wanneer iemand (veelal een onderzoeker) die constructie probeert te begrijpen (lees: te interpreteren) (vgl. Schuiten & De Korte, 2022).

Kader 9. Adagium: ‘Wat standhoudt, is waar’

Ik heb de eenvoud van dit adagium aan Mark van Twist te danken. Shotter (1995) biedt een meer ingewikkelde versie: “Indeed returning to [Gergens] claim that “what we count as knowledge are temporary locations in dialogic space,” what [Gergen] sees as important is the other’s “position” in that space and one’s ability to confront it with one’s own claims “to know”. Where, as he points out, the capability in doing this involve a whole collection of different, essentially argumentative or rhetorical skills, that must be coordinated if one is to sustain one’s claim in the face of challenges from others.” (pp. 47-48).

Na deze beschouwing van Guba en Lincoln (1989) ben ik op zoek gegaan naar literatuur die deze bijdrage kritisch heeft beschouwd. Vanuit daar zijn een aantal belangrijke bezwaren naar voren gekomen. Na deze kritische noten bij de vierde evaluatiegeneratie, tezamen met het

Het tweede kritische geluid betreft het ideaal van consensus. Ten eerste beargumenteert Gregory (2000), dat consensus onterecht wordt gezien als automatisch leidend tot verandering, waar het gebrek aan consensus – volgens diezelfde logica – zou leiden tot een impasse. Laughlin en Broadbent (1996) wenden zich tot Habermas' communicatieve actietheorie om te betogen dat:

enlightened understanding (whether agreed in a consensus or not) does *not* lead automatically to action. There is a clear distinction between theory and praxis in Habermas's world. Action alternatives need to be debated by the discursive parties just as rigorously following similar discursive rules as in the initial understanding process. (p. 445)

Maar ook een bekender spreekwoord vat het goed samen: 'er zijn meerdere wegen die naar Rome leiden'. Kortom: men kan het met elkaar eens zijn over een probleem of bepaalde oorzaken, maar toch kiezen voor verschillende acties. Daarom is het van belang om daar explicieter aandacht aan te schenken.

De tweede kant van de consensusmedaille betreft een methodologische inconsistentie in het gedachtegoed. Nogmaals: het na te streven ideaal dat is consensus over een gedeelde realiteit. Door dit centraal te stellen – en zelfs als succes criterium voor een goede evaluatie te beschouwen – wordt er dus toch waarde gehecht aan een bepaalde waarheid die moet worden nagestreefd, namelijk de gedeelde realiteit (Kushner, 1996). Het is dan ook alleen deze realiteit die in het eindrapport wordt opgenomen. Daarmee worden andere realiteiten dus als illegitiem beschouwd vanuit het evaluatieproces.

Ondanks een poging tot paradigmawisseling zijn deze twee kritiekgeluiden te traceren. Een nadere uitwerking van de vierde evaluatiegeneratie geeft gehoor aan deze kritische geluiden: de responsieve evaluatie, zoals vormgegeven door Tineke Abma.

literatuuronderzoek over sociaal constructionisme, heb ik mijn hoofdvraag gewijzigd: "Hoe kan een sociaal constructionistische methode van onderzoek de gereedheidskist van de internal auditor verrijken?"

Vervolgens ben ik verder de literatuur in gedoken en kwam ik terecht bij een nadere toepassing van de Responsieve Evaluatie van Tineke Abma. Ik werd enthousiast: het was een concrete uitwerking, waarin uitingen van het sociaal constructionisme voorkwamen.

Vierde cyclus: Abma's Responsieve Evaluatie

Een concrete uitwerking van sociaal constructionisme zit in de methode van Responsieve Evaluatie, zoals door o.a. Tineke Abma is vormgegeven. Het is een nadere uitwerking van de vierde evaluatiegeneratie van Guba en Lincoln (1989), gegrond in het gedachtegoed van Stake (1975), maar focust zich meer op het concept van dialoog. Deze evaluatie is bij uitstek geschikt voor evaluatieobjecten waarbij óf het object zelf óf de context zich kenmerkt door ambiguïteit en complexiteit (vgl. Wadsworth, 2001). De responsieve evaluatie is “geen technisch procedé [...] [maar] een onderzoeksbenadering.” (Van Berkel, e.a., 2022, p. 8). Kort door de bocht: het doel van de responsieve evaluatie is om tot wederkerig begrip te komen door dialoog, waardoor bewustwording en (nieuwe) inzichten ontstaan (Woelders & Abma, 2019), om eventueel – maar niet noodzakelijk – tot gemeenschappelijke actie te komen. Hieronder beschrijf ik de Responsieve Evaluatie met zes pijlers, aangevuld met een kritische beschouwing per pijler. Eerst wend ik mij tot een korte samenvatting van het evaluatieproces.

Responsieve Evaluatie: wat doe je kortgezegd?

De Responsieve Evaluatie begint net als andere evaluaties met een kennisvraag: een behoefte om een evaluatieobject te beschouwen en daar een oordeel over te geven. Daarom begint het met een intake met een beoogd opdrachtgever; bij voorkeur zelfs met een kleinschalige projectgroep die meegenomen wordt in alle evaluatiestappen (Abma, 2005, p. 285). Samen met deze projectgroep worden de belanghebbende groepen geïdentificeerd: iedereen met een mogelijk belang bij de evaluatie en het evaluatieobject.

Hoewel de Responsieve Evaluatie wordt gekenmerkt door een zogenaamd verschijnend ontwerp, zijn er idealiter drietal fasen (Abma, 2001, 2003, 2005). In de eerste fase verzamelt de evaluator met open gesprekken en participatieve observatie zo veel mogelijk verschillende verhalen bij de verschillende belanghebbende groepen. Deze verhalen worden geanalyseerd en gepresenteerd in een interim-rapport. In deze tweede fase worden verhalen in homogene focusgroepen besproken. Deze dialoog waarin verhalen tegenover elkaar worden gezet – waarin verschillende betekenissen met elkaar interacteren – vormt vervolgens de input voor de laatste fase. In deze derde stap gaan namelijk de verschillende belanghebbenden in dialoog tijdens een heterogene dialoogsessie. Deze einddialoog vormt dan ook input voor het eindrapport van de evaluatie, waarin betrokkenen dus gezamenlijk zich een oordeel hebben gevormd over (lees: betekenis hebben gegeven aan) het evaluatieobject. Het evaluatierapport is in feite een “*collage of stories and conversations*” (Abma, 2003, p. 235), met een rijkdom aan context om een oordeel te vormen over het evaluatieobject. Een

*En toen begon mijn echte duik in het diepe; althans, zo heb ik het ervaren. Ik wilde antwoord krijgen op de vragen: **Wat houdt Abma's Responsieve Evaluatie in? Op welke principes is het gegrond? Hoe ziet het proces van een Responsieve Evaluatie eruit?** Dat heeft geresulteerd in een diepgaande literatuurstudie in de Responsieve Evaluatie, begrippen als 'hermeneutiek', 'dialoog', een verdieping van 'sociaal constructionisme' en narratief onderzoek.*

Toen kwam ik er ook achter dat de responsieve evaluatie meer een grondhouding is dan “een technisch procedé” (Van Berkel, e.a., 2022, p. 8) en dat er dus niet één juiste manier is van responsief evalueren. Ik kwam erachter dat Abma's Responsieve Evaluatie een specifieke vorm is die narratief onderzoek evenaart en om die reden open gesprekken en verhalenworkshops hanteert. Een andere invulling van responsief evalueren, met andere methoden is uiteraard ook mogelijk. Voor dit afstudeeronderzoek heb ik me echter gericht op het centraal stellen van verhalen. Dat heeft ook te maken met verschillende gradaties waarin belanghebbenden kunnen worden betrokken (Van Heijster, e.a., 2020). Deze vorm van de Responsieve Evaluatie zit aan het einde van het spectrum, waar de betrokkenheid het grootst is.

Op dit punt wist ik nog niet dat ik een mogelijke casestudy zou vinden voor de Responsieve Evaluatie.

scherpe lezer doorziet al dat de voornaamste leerinterventie van de Responsieve Evaluatie zich niet bevindt in dit eindrapport, maar in de focus- en dialooggroepen waarin betekenissen met elkaar interacteren, (nieuwe) inzichten zich vormen, en wederkerig begrip ontstaat.

Responsieve Evaluatie: waarvoor staat het?

In de eerste pijler plaats ik – in navolging van Abma en Widdershoven (2008) – evaluatie als onderneming binnen een sociale praktijk: het uitvoeren van een evaluatie is geen neutrale, onbevangen activiteit. Vervolgens duik ik in de onderliggende veronderstellingen van de Responsieve Evaluatie, het sociaal constructionistisch gedachtegoed van o.a. Cunliffe (2008) en Hosking (2011) volgend. Ten derde sta ik stil bij het concept van dialoog als zowel axiomatisch als methodologisch principe binnen de Responsieve Evaluatie. De vierde pijler betreft het verschijnende ontwerp, waar de vijfde pijler het belang van pluralisme en variëteit benadrukt. Ten slotte beschouwt de zesde pijler de verschillende rollen van de evaluator.

1. Relatieve & Sociaal: Evalueren als onderneming is een sociale praktijk

De grondpijler van de Responsieve Evaluatie onderkent dat evalueren inherent een sociale en relationele handeling is (Abma & Widdershoven, 2008). Het is allesbehalve een objectieve, onafhankelijke exercitie in een vacuüm. Het is een manier van praktijkgericht onderzoek waarin sociale verhoudingen tussen evaluator en belanghebbenden aan continue verandering onderhevig zijn tijdens het onderzoeksproces; niets staat in steen geschreven. Daar moet de evaluator zich van bewust zijn.

Abma en Widdershoven (2008) classificeren vier perspectieven op sociale verhoudingen in een evaluatie. Vanuit hun betoog moge het duidelijk zijn dat de responsieve evaluatie sociale verhoudingen ziet als dé centrale component van de evaluatie. Immers: in navolging van een subjectivistisch paradigma ontstaan werkelijkheden en onze kennis daarover vanuit interactie die men met elkaar heeft. De evaluator is niet langer de onafhankelijke, technische expert, maar fungeert als een provocateur tot begrip (Abma & Stake, 2001; Schwandt, 2001). De evaluator is niet langer een externe actor, maar draagt actief bij aan de creatie van onderling begrip bij en tussen belanghebbenden (Visse, e.a., 2011). In die context schrijft Karlsson (2001) ook dat het uiteindelijke doel van een dergelijke evaluatie niet ligt bij een waardeoordeel van een evaluatieobject, maar dat het onderzoeksproces *per se*, waarin dat wederkerig begrip ontstaat, het belangrijkste doel is (vgl. Patton, 2018). De evaluatie en het -proces zijn dus ook een leerinterventie (zie o.a. Abma 2003; Smit & Verdonschot, 2010; Snoeren, e.a., 2015).

Zoals eerder beschreven was één van de meest vocale kritieken op de vierde evaluatiegeneratie gericht op de omgang met asymmetrische machtsverhoudingen (vgl. Vanderplaat, 1995), hetgeen zich vertaalt naar relatief gewicht van bepaalde

In mijn beschouwing van de Responsieve Evaluatie ben ik terechtgekomen bij een artikel van Cunliffe (2008). Zij verwoordde precies waar ik op dat moment mijn vinger niet op kon leggen: sociaal constructionisme. Aan de hand van bijvoeglijke naamwoorden beschreef zij deze stroming in de sociale wetenschappen. Ik probeer haar in deze bijdrage recht te doen, door een soortgelijke indeling te hanteren om de Responsieve Evaluatie als concept, als onderzoeksbenadering te duiden (vgl. Van Berkel, e.a., 2022).

Die bijvoeglijke naamwoorden geven nadere duiding aan de pijlers als overkoepelende en richtinggevende betekenissen. Aanvankelijk – achteraf bezien vrij naïef – heb ik op verschillende manieren geprobeerd de verschillende verhoudingen tussen de pijlers weer te geven: als treden en zuilen van een Griekse tempel, als spinnenweb met de overal aan elkaar verbonden draden, enzovoorts. Maar alle verschillende vormen konden de onderlinge interactie en verbinding van de pijlers niet goed genoeg verbeelden. Het moge daarom duidelijk zijn dat de ene pijler niet zonder de andere pijler kan en dat de verschillende duidingen ook in deze kritische beschouwing in continue interactie met elkaar staan.

betekenissen. Bedenk een setting waarin men wordt gevraagd naar ervaringen omtrent een programmadoel, waarbij de programmamanager aangeeft dat dit doel lang en breed gehaald is en dat verdere discussie niet nodig is. Het behoeft geen lange uiteenzetting om aan te tonen dat een programmamedewerker wel twee keer nadenkt om een andere ervaring te delen... Dit – karikaturaal beschreven – fenomeen wordt binnen institutioneel-theoretische literatuur, en met name binnen het discursief institutionalisme, ook wel de **macht over ideeën** genoemd (Carstensen & Schmidt, 2016).

Deze kritiek is bij de Responsieve Evaluatie grotendeels weggenomen met de onderzoeksfasering; m.n. de tweede fase van homogene focusgroepen is hierbij belangrijk. Zo worden namelijk in de te onderscheiden belanghebbende groepen de verschillende geïdentificeerde verhalen onafhankelijk van elkaar behandeld (Baur e.a., 2010). Men kan eerst ‘onder elkaar’ deze verhalen en de verschillende betekenissen daarvan bespreken, voordat deze exercitie in een heterogene dialooggroep wordt herhaald. Het volledig wegnemen van formele machtsrelaties is niet mogelijk, noch zou het wenselijk zijn. Juist de interactie tussen verschillende verhalen uit verschillende groepen draagt bij aan onderlinge betekenisgeving en bewustwording – het ultieme doel van de Responsieve Evaluatie. Hierover meer in onderstaand.

2. Samenwerkend, Inclusief & Interactief: Stimuleren van omgang met belanghebbenden, alsook onder belanghebbenden onderling

Voordat ik verderga met de andere pijlers, is het goed om stil te staan bij deze tweede pijler. De kernwoorden zijn ‘interactief’, ‘samenwerkend’ en ‘inclusief’. Er ligt immers in de Responsieve Evaluatie een nadruk op het betrekken van een grote variëteit aan belanghebbenden. Het pleit voor een actieve inclusie van alle groepen die een belang hebben bij de evaluatie en/of het evaluatieobject (Abma, e.a., 2009). Ieders perspectief, veronderstellingen en onderliggende logica’s spelen een rol.

Daarnaast ziet de responsieve evaluator betrokkenen niet als passieve actoren; ze zijn niet louter respondenten of informatieverstrekkers, maar betrokkenen zijn actieve partners in de evaluatie (Abma, 2006, p. 39). In dat opzicht zijn ze co-creators van de (kennis voortkomend uit de) evaluatie en werkt de evaluator *samen met* hen. In dezen is de Responsieve Evaluatie niet de enige vorm van evaluatie die een dergelijk ideaal nastreeft. We zien dit nastreven van volledige samenwerking ook terug in de **Democratische Evaluatie** van Greene (1997, 2001) en de **Transformatieve Evaluatie** van Mertens (2010, 2017).

Al deze evaluatievormen hebben gemeen dat ze vooropstellen dat er geen eenduidige werkelijkheid is. Ze onderkennen wél dat er een *dominante* realiteit kan zijn. Deze dominante realiteit is contextafhankelijk; immers: deze wordt sociaal geconstrueerd,

De notie van asymmetrische machtsrelaties was mij niet onbekend. Voor mijn masterscriptie heb ik gewerkt met de theorie van discursief institutionalisme van Vivien Schmidt (2008). Ik heb hierop teruggegrepen, om enigszins weerklank te kunnen bieden aan bezwaren omtrent solipsisme en de befaamde uitspraak “anything goes” bij subjectivistische benaderingen (vgl. Edwards e.a., 1995). Het concept van ideële macht (hieronder verder uitgewerkt) biedt namelijk inzicht in de verschillende discursieve krachten die aan het werk zijn en (mede) bepalen waarom sommige ideeën beter resoneren dan andere.

Kader 9. Ideële macht.

In hun conceptualisatie van “ideële macht”, beschrijven Carstensen en Schmidt (2016) deze macht over ideeën als het vermogen om te bepalen wat als legitiem of (collectief) geaccepteerd wordt gezien. Dit heeft niets te maken met de overtuigingskracht van het idee zelf, noch met de resonantie ervan met geldende, contextuele normen en waarden; het heeft louter iets te maken met de persoon of groep die het idee oppert. Degene die de boodschap brengt, heeft dusdanige autoriteit, legitimiteit, een vorm van macht, dat diens ideeën automatisch meer gewicht en legitimiteit met zich meebrengen. En zodoende hebben zij een zekere mate van controle over de productie van betekenis(geving). De andere soorten ideële macht die zij identificeren zijn: macht via ideeën en macht binnen ideeën.

Kader 10. Transformatieve Evaluatie.

*De Responsieve Evaluatie verschilt van de **Transformatieve Evaluatie (TE)**, omdat de laatste transformatie beoogt van sociale verhoudingen. Marxistische, feministische en post-kolonialistisch theoretische standpunten liggen dichtbij. Je zou kunnen beweren dat de TE een stap verder gaat dan de Responsieve Evaluatie. Waar deze democratisch is in de zin **Zie volgende pagina voor vervolg.***

bevochten en/of in standgehouden. Om dat te onderzoeken, pleiten al deze benaderingen voor een inclusie van álle belanghebbende partijen, waarin de evaluator één van die partijen is en niet daarboven staat. Het is dus van het allerhoogste belang dat je als evaluator in deze benadering jezelf, noch jouw perspectief c.q. realiteit, als belangrijker acht dan welke andere dan ook.

Bovenstaand lezende, komt het ook niet als verrassing dat de Responsieve Evaluatie een plek kent binnen de bredere beweging van participatief (actie)onderzoek (Woelders & Abma, 2019). Zoals ook nader zal blijken uit de volgende pijler, staan dialoog en interactie hoog in het vaandel bij de responsieve evaluatie. Interactie is immers het vehikel waarmee betekenis, begrip en inzicht tot stand komen en werkelijkheden worden geconstrueerd (Hawkesworth, 1989; Vanderplaat, 2009; Van Twist, 2018). Deze vorm van evaluatie stimuleert zowel interactie met belanghebbenden, alsook bij belanghebbenden onderling, om zo wederkerig begrip (zie ook volgende pijler) te bevorderen. In de rest van dit hoofdstuk geef ik een nadere beschrijving over de verschillende vormen van interactie die een rol spelen binnen de responsieve evaluatie.

Het is eveneens niet verrassend om Van der Haar & Hosking (2004) te volgen in de stelling dat de Responsieve Evaluatie een sociaal constructionistische benadering is. Het sociaal constructionisme onderschrijft een intersubjectieve ontologie: realiteit ontstaat tussen agenten. Daarmee volgt het nadrukkelijk geen objectieve ontologie ('er is één sociale waarheid'), maar een variatie op de subjectieve ontologie ('er zijn meerdere sociale werkelijkheden'). En dat maakt de Responsieve Evaluatie dus ook verschillend van een klassieke audit – die een objectieve ontologie volgt – maar ook van een behavioural audit – die een ontologie volgt waar sociale werkelijkheden zijn belegd in mentale modellen, en dus in een individu's cognitie.

3. Dialogisch: dialoog staat ontologisch, epistemologisch én methodologisch centraal

Wie Responsieve Evaluatie zegt, moet in dezelfde ademteug het woord 'dialoog' noemen. De Responsieve Evaluatie volgt namelijk een dialogisch proces. Dat houdt in dat het voeren van een oprechte dialoog centraal staat. In een oprechte dialoog is men op zoek naar betekenissen door daar met elkaar over te spreken, gedachten te ordenen, kritisch te bevragen, etc. Je zou het in zekere zin kunnen typeren als een Socratisch gesprek (Wiss, 2020; Karlsson, 2001). De evaluator en de belanghebbenden gaan op zoek naar elkaars veronderstellingen, perspectieven en gedachten met als doel deze te willen begrijpen (Karlsson, 2001).

In tegenstelling tot de vierde evaluatiegeneratie, stelt de Responsieve Evaluatie dat consensus niet het uiteindelijke doel hoeft te zijn; het doel is creatie van onderling wederkerig begrip tussen belanghebbenden, op basis van een mate van

Vervolg kader 10. Transformatieve Evaluatie.

deze democratisch is in de zin van Greene door alle mogelijke betrokkenen een stem te geven (zie kader 11), geeft de TE bepaalde stemmen – en met name stemmen vanuit een gedefinieerde minderheid – extra weging (Mertens, 2017). Waar de Responsieve Evaluatie oog heeft voor de verschillende machtsverhoudingen, zijn deze in de TE het centrale punt; het streeft sociale rechtvaardigheid na (Mertens, 2010). Dat gaat – nogmaals – een stap verder dan de Responsieve Evaluatie, waar de inclusie van de betrokken en belanghebbende partijen voorop staat.

Kader 11. Democratische Evaluatie.

De Democratische Evaluatie staat tussen de Responsieve en Transformatieve Evaluatie. Deze vorm van evaluatie streeft eveneens belangenbehartiging van bepaalde doelgroepen na, omdat deze in reguliere sociale verhoudingen worden buitengesloten. De evaluator wordt dus ook een expliciete behartiger voor deze belangen. Daarin verschilt de benadering nog steeds van de Responsieve Evaluatie. Waar deze twee vormen van evaluatie dichter tot elkaar treden, is omtrent het concept van dialoog. De Democratische Evaluatie ziet – net als de Responsieve Evaluatie – dat de dialoog tussen de betrokken en belanghebbende partijen de constituerende component is van de verhouding. En in die dialoog komt dus ook betekenis tot stand en kun je eventuele verandering stimuleren en bewerkstelligen.

bewustwording en verworven inzichten in het evaluatieproces (Widdershoven, 2001). Wederkerig begrip kan leiden tot overeenstemming over een bepaalde realiteit (lees: consensus), maar dat hoeft niet. Een Responsieve Evaluatie is succesvol wanneer *“it results in better understandings [and will] lead to an enhanced insight among all participants, so that they are better equipped to interact creatively with their surroundings”* (Abma & Widdershoven, 2011, p. 678).

Dialoog wordt zowel een **ontologisch als methodologisch** karakter toebedeeld. De nadruk op dialoog doorwoekert alle fasen van het onderzoek in verschillende verschijningsvormen. In eerste instantie vindt de dialoog plaats tussen de evaluator en betrokkenen. Dit is tijdens de eerste fase met verkennende, onderzoekende gesprekken. Omdat het gesprek een dialogisch karakter heeft, is het allesbehalve een semigestructureerd interview. Immers: dat is meer een vragend gesprek; een monologisch gesprek, waarin vraag en antwoord als een eenrichtingsverkeer dienen: je wilt antwoord hebben op een van tevoren bedachte vraag (Schwandt, 1997). Hoewel dit in een semigestructureerd interview minder aan de orde is dan in een gestructureerd interview, behoudt deze vorm van bevragen nog steeds het monologische karakter. Je bent minder vrij in je onderzoekende houding; je hebt een richting die je op wilt. En die richting is eenzijdig bepaald. Een dialogisch gesprek laat zich meer kennen door het onderzoekende karakter van beide partijen; een tweerichtingsverkeer (Abma, 2001). Samen gaan de evaluator en de betrokkene op zoek naar verhalen, onderliggende veronderstellingen (Hudaib & Haniffa, 2009; Schwandt, 1994). Een tweede instantie waar de dialoog zich voordoet, is in de homogene focusgroep. Deze focusgroep dient ertoe om bij geanalyseerde verhalen nader de verdieping op te zoeken; nader veronderstellingen, normen en waarden en paradigma's uit te kristalliseren (Abma, 2004). Het is dus wederom een onderzoekend gesprek, maar nu met meerdere betrokkenen. Deze betrokkenen hebben – mede gelet op de uiteenzetting vanuit de eerste pijler – wel hetzelfde belang. Dit is weer anders waar – in de derde instantie – betrokkenen met verschillende belangen bij elkaar komen in de heterogene dialooggroep. Hierin zit de dubbele werking van de dialoog besloten: enerzijds onderzoekend door elkaars realiteiten te begrijpen; anderzijds wordt door dit begrip een nieuwe realiteit gecreëerd, dan wel een bestaande realiteit verrijkt (Abma, 2005; Camargo-Borges & Rasera, 2013; Gadamer, 1975).

In laatste instantie komt de dialoog ook naar voren in de heterogene dialooggroep. Waar de vorige dialooggroep leidt tot verdieping van verhalen per belanghebbende groep, is het doel van deze dialoog om de verhalen (lees: realiteiten) bij monde van de verschillende belanghebbende groepen met elkaar in interactie te laten treden. Het doel is om hier *“increased personal and mutual understanding or enhanced understanding of differences”* te creëren (Abma & Heijman, 2013, pp. 464-465; vgl. Nierse & Abma, 2011). Dit is dan ook de factor waaraan de

Kader 12. Dialoog als ontologie én methodologie.

Het is behulpzaam onderscheid te maken tussen een dialogische ontmoeting als “ontological condition and the hermeneutic mode of inquiry” (Schwandt, 1997, p. 76). In dat opzicht laat het ook de overgang zien van de vorige twee pijlers naar deze pijler. Waar ik in de vorige pijler kritisch uiteen heb gezet dat dialoog dé plek is waar in interactie realiteiten ontstaan, bevestigd en bevestigd worden – en kennis dus ook op die manier ontstaat – heeft dialoog ook een meer methodologisch karakter. Door de dialoog met elkaar aan te gaan, ontstaat een onderzoekend gesprek, waarin – binnen de Responsieve Evaluatie – verhalen naar boven komen die staan voor veronderstellingen, aannames, normen, waarden, paradigma's, etc. In die zin is dialoog – naast een ontologisch axioma – ook een methodiek van onderzoeken.

Wel moet onderstreept worden dat de dialoog als onderzoeksmethode nadrukkelijk moet passen binnen het postmoderne paradigma. Dat houdt in dat de dialoog een “narrative rationality” volgt (Abma, 2001, p. 247), waarbij persoonlijke verhalen van betrokkenen, het begrip daarvoor en de zoektocht naar onderliggende waarden centraal staat. Dit staat haaks op een meer instrumentalistische inzet van de dialoog als onderzoeksmethode, waarin argumenten worden getoetst aan feitelijk bewijs (Abma & Widdershoven, 2005, p. 95). Een inzet van de dialoog op deze instrumentalistische manier zou mogelijk alleen meer bijdragen aan wat Kauko en Pajunen (2021) omschrijven met de ‘thickening modern’ binnen de evaluatiegemeenschap; het fenomeen waar methoden behorend in het postmoderne – met

kwaliteit van de dialoog kan worden ontleend. Nadrukkelijk is het doel niet om hier tot een (verslachte vorm van) consensus te komen; dit *kan* een uitkomst zijn, maar dit *moet* niet (Abma, 2004; Abma & Widdershoven, 2005). Centrale kwaliteitsaspecten van een dialoog als onderzoeksmethode zijn openheid, wederzijds respect, inclusie en betrokkenheid (zie voor een nadere uitwerking hiervan: Greene, 2001, p. 183-185; vgl. Abma, e.a., 2001).

bijbehorende onderliggende veronderstellingen en waarden – ten onrechte worden ingezet in evaluaties die doordrenkt zijn van modern-paradigmatische elementen.

4. Verschijnend: Onderzoeksontwerp en -proces ontstaan terwijl je bezig bent

Om methodologisch in lijn te blijven met de vorige pijler, betreft de vierde pijler een nadere classificatie van het onderzoeksontwerp en -proces. Een meer traditionele evaluatie (en eveneens audit) is vrij bepalend en top-down ingeschoten (lees: een deductieve logica volgend): het ontwerp van tevoren wordt vastgesteld en afgestemd met de opdrachtgever, om vervolgens exact de stappen te volgen die van tevoren zijn bedacht, zonder daar in al te grote mate van af te wijken. Een meer verschijnende vorm van evaluatie laat dit los. Wel blijven in de Responsieve Evaluatie drie fasen van onderzoek staan: verkennende gesprekken, homogene focusgroepen en heterogene dialooggroepen. Wanneer je echter van de ene fase in de andere fase overstapt, is van tevoren nog niet bekend.

Een eerste manier waarop je meer verschijnend te werk gaat, is door je als evaluator te laten leiden door de onderwerpen en perspectieven die de betrokkenen aandragen (Abma, 2005, 2006). Daarmee werk je dus niet met een deductief plan van aanpak, waarin alle variabelen, criteria, indicatoren en consorten van tevoren zijn uitgeschreven, maar hanteer je een meer open blik met een inductieve logica. Een tweede manier is verbonden met het concept van saturatie (Saunders e.a., 2018; Sebele-Mpofu, 2020). Bij een meer bepalend en richtinggevend ontwerp wordt van tevoren bepaald hoeveel en met welke functionarissen in het kader van de evaluatie wordt gesproken. Bij een meer verschijnend ontwerp blijft de evaluator initiële gesprekken voeren, totdat deze geen nieuwe inzichten meer krijgt vanuit de betrokkenen (Nierse & Abma, 2011, p. 413).

5. Verschillend & Pluralistisch: Variëteit van perspectieven, verhalen en betekenissen staat voorop

In nauwe aansluiting op de vorige pijler, zijn de kernwoorden van deze vijfde pijler 'verschillend' en 'pluralistisch'. In een responsieve evaluatie ga je op zoek naar verschillen om zo veel mogelijk verschillende perspectieven de mogelijkheid te geven ten tonele te komen. Daarmee is het ook een inherent pluralistische benadering, waarin waarde wordt gehecht aan een grote variëteit van persoonlijke verhalen (Fraser, 2004). Deze verscheidenheid is niet erg. Sterker nog: deze verscheidenheid doet ertoe. Binnen een organisatie leven veel verschillende verhalen, die constituerend zijn voor verschillende werkelijkheden en daaruit voortvloeiende handelingen. Dit wordt

ook wel de **performatieve werking van taal** genoemd (Burr & Dick, 2017; Milliken, 1999; Van Twist, 2018).

Dit komt tot uiting in de analyse van de verhalen, waar niet alleen met een thematische, interpretatieve analyse wordt gewerkt, maar ook met een meer narratieve analyse (Abma, 2001). Er wordt geanalyseerd hoe de reeks aan gebeurtenissen (lees: het plot) wordt geordend (Bussman, e.a., 2013; Czarniawska, 2004); welke personages daar een rol in spelen en hoe de verteller zich tot hen verhoudt (Soderberg, 2006; Sools, 2012; Van der Arend, 2007). Dat maakt het ook een sociaal constructionistische analyse die zich richt op *hoe* een verhaal wordt geconstrueerd en zich niet alleen focust op *wat* in het verhaal wordt verteld (Cunliffe, 2008; Fletcher, 2006; Grandy, 2018).

De performatieve werking blijkt uit het dominante verhaal dat leeft en een organisatie doorwoekert, alsook met de zgn. tegenverhalen die op een of meerdere manieren dit dominante verhaal tegenspreken (Abma, 2003). Door deze verschillen en overeenkomsten bloot te leggen en daar vervolgens over in dialoog te treden, ontstaat een zekere mate van inzicht en bewustwording over de verschillende werkelijkheden binnen een organisatie. Om deze verscheidenheid aan perspectieven te kunnen identificeren, vraag je in de initiële gesprekken de geïnterviewde een andere participant te nomineren voor het onderzoek die een andere kijk heeft op de situatie dan van diegene (Abma, 2003). Dit draagt ook bij aan het creëren van de saturatie, zoals in de vorige paragraaf besproken.

6. Rol van de evaluator: allesbehalve een inhoudelijke expert

Vanuit al het bovenstaande blijkt al dat er andere rollen liggen weggelegd voor de evaluator dan bij meer toetsend onderzoek. Het moge duidelijk zijn vanuit de bovenstaande alinea's dat de het ideaal (of voor de criticus: de illusie) van een objectieve onderzoeker niet wordt nagestreefd. De evaluator is daarmee ook niet langer een inhoudelijk expert. Als de responsieve evaluator niet objectief een feitelijke situatie vaststelt, wat doet deze dan wel? Het antwoord is veelzijdig.

Je bent een verhalenanalist: je analyseert de verschillende verhalen die je vanuit de gesprekken hebt opgehaald. Wat veelal naar voren zal komen is een dominant verhaal – een verhaal dat alle gesprekken doorwoekert – en zgn. tegenverhalen (Abma, 2003, p. 224). Dit zijn verhalen die tegen het dominante verhaal in gaan, die dit verhaal ontkrachten, of – in een minder extreem geval – nuanceren. Maar de rol van responsieve evaluator gaat verder dan alleen het analyseren van verhalen.

Je creëert en borgt de kwaliteitscriteria voor een goede dialoog (vgl. Greene, 2001). De openheid van het gesprek staat voorop, zodat het gesprek een onderzoekend karakter kan krijgen (Wisse, 2020, p. 134). In die zin ben je een facilitator van het dialogische proces (Abma, 2003). Daarnaast blijf je altijd kritisch op wat

Kader 13. Performatieve werking van taal.

Specifieker gaat Vivien Schmidt (2008) in op deze performatieve werking bij haar duiding van de 'macht in ideeën' als separate analytische categorie van ideële macht. Bepaalde ideeën en verhalen blijven beter beklijven in een organisatie en spreken meer aan dan andere ideeën. Zij stelt dat dit alles te maken heeft met de mate waarin een idee resoneert met zgn. "background ideational abilities" (p. 314), o.a. ontleend aan Bourdieu's concept van habitus. Deze vaardigheden van de betreffende actor zorgen ervoor dat je kúnt bepalen wat goed (en dus ook slecht) past in een gegeven situatie, ingegeven vanuit een specifieke (organisatie)context (Schmidt, 2010).

anderen zeggen, op hun verhalen, hun realiteiten en onderliggende ideeën. Hierin ligt een rol voor de evaluator als Socratische gids, waarbij betrokkenen geconfronteerd kunnen worden met hun eigen aannames en de effecten daarvan op de geconstrueerde realiteit(en) (Abma, 2005, p. 280). Dit kan in alle fasen van de responsieve evaluatie: in de individuele gesprekken, de homogene focusgroepen en de heterogene dialooggroep.

Je laat de belanghebbenden onderling en de verschillende belanghebbende groepen kennismaken met elkaars verschillende verhalen en realiteiten, zodat dit leidt tot een vergroot begrip onderling dat er eerst nog niet was. Je hebt in feite een rol van educator. Je geeft de stemmen die niet zo vaak voorkomen ook vooral het woord (Abma & Widdershoven, 2005). Ook deze verhalen, deze realiteiten doen ertoe. Let wel: je spreekt niet namens deze groepen, maar je draagt er zorg voor dat ook deze groep een podium krijgt (Baur, e.a., 2010). Zo draag je bij aan de kernwaarde van inclusie van de responsieve evaluatie: je ondersteunt groepen met minder formele – dan wel informele – macht (Greene, 2001).

Vijfde cyclus: Van Responsieve Evaluatie naar een aanpak voor een Responsieve Audit?

Binnen de casus, zoals beschreven in het andere onderdeel van deze bijdrage, heb ik een plan geschreven voor een mogelijke Responsieve Audit. Een eerste versie van dit plan was nagenoeg een exacte vertaling van verschillende bijdragen van Abma en anderen, toegespitst op de specifieke casus. In feite was dit het resultaat van het literatuuronderzoek, dat ik vervolgens lichtelijk empirisch heb gevalideerd. Op basis daarvan heb ik aanpassingen gedaan aan het plan, waarvan het resultaat in het andere onderdeel van de bijdrage is weergegeven. Deze **empirische validatie** betrof twee stappen, die ik hieronder toelicht.

Empirische validatie met tweede versie van plan van aanpak als resultaat

Een eerste stap heb ik gezet door met de betrokken collega een tweetal ongestructureerde gesprekken te voeren over de praktische toepasbaarheid van het plan van aanpak. Uit die twee gesprekken zijn drie aspecten naar voren gekomen, die hebben geleid tot wijzigingen in de eerste versie van het plan van aanpak. Deze wijzigingen zijn:

- De verlengde betrokkenheid wordt op een andere manier geborgd dan door participatieve observatie. Dat zou betekenen dat één van de onderzoekers intensief mee zou draaien op de beleidsafdeling. Het doel van de verlengde betrokkenheid is een dieper inzicht in het werk. De betrokken collega heeft jarenlange ervaring bij dit ministerie, is bekend met de betrokken afdelingen en kent het beleidsveld. Dat was afdoende om het doel van dieper inzicht te bereiken.
- Deugdelijk kwalitatief onderzoek wordt gekenmerkt door transcripties en dat de interpretatie daarvan vervolgens wordt teruggelegd bij de betrokkene (zie bijv.: Meadows & Morse, 2001). Het is geen geheim dat transcriberen een tijdsintensief proces is. Dit heeft geleid tot het verwijderen van deze stap in de verslaglegging; wel wordt het verhaal van de betrokkene uitgeschreven en voor hoor/wederhoor teruggelegd. Op deze manier wordt de authenticiteit van het verhaal bewaakt met een andere validatiestrategie.
- De stakeholderanalyse en probleemdefinitie zijn nog onbenutte fora voor interactie en dialoog. De uitwerking ervan wordt primair bij de onderzoeker verondersteld. Door de stakeholderanalyse en probleemdefinitie met een meer uitgebreide workshop interactief te maken, ontstaat al ruimte voor de dialoog tussen – in dit geval – de betrokken ambtenaren en wordt de eerste aanzet gedaan tot een reflectief proces over de relatieve belangen in het speelveld. Om meer duiding te kunnen geven aan het concept ‘belang’, is het model van Balane e.a. (2020) gebruikt. Zij hebben onderscheid gemaakt tussen vier assen die een belanghebbende kunnen definiëren, te weten:

*Toen kwam vervolgens enigszins opluchting: ik had een collega gevonden die net een opdracht zou gaan starten en op zoek was naar een manier van onderzoeken die recht zou doen aan het betrekken van meerdere belanghebbenden. Wij zijn toen met elkaar in contact gekomen via onze directeur onderzoek. Dat betekende dat ik een empirische validatie kon doen voor deze verkenning. De vragen die bij mij opkwamen, waren: **In hoeverre kun je de aanpak van de Responsieve Evaluatie toepasbaar maken voor deze specifieke casus? Wat zijn daarbij sterke punten? Wat zijn mogelijke belemmeringen?***

*Achteraf had ik deze eerste **empirische validatie** anders kunnen (en allicht moeten) doen. Nu heb ik gelijk na de eerste ronde – op basis van twee gesprekken met dezelfde persoon – al wijzigingen doorgevoerd in het plan van aanpak. Vervolgens is de tweede versie van het plan gevalideerd in de focusgroep en twee individuele gesprekken. Daarin had ik ook de eerste versie van het plan centraal kunnen stellen en de verschillende beelden onderling kunnen vergelijken.*

belang, bevoegd, bekwaam en bereid. Dit geeft betrokkenen meer houvast in hun kwalitatieve inschatting van het speelveld.

Beperkte methodologische fundering leidt tot noodzaak voor aanvullende empirische validatie

Daarmee is een tweede versie van het plan van aanpak ontstaan. Omwille van de geloofwaardigheid van de empirische validatie, zijn aanvullende methoden ingezet: een focusgroep en twee open gesprekken. Naast deze meer actiegerichte doelstelling, zijn de methoden nog voor een andere, meer reflectieve doelstelling ingezet: reflectie op de aanpak in relatie tot het concept 'audit'. Voor de keuze voor deelname aan de focusgroep heb ik enkele **selectiecriteria** toegepast. Een achttal collega's zijn uitgenodigd; uiteindelijk heeft de focusgroep met een viertal collega's plaatsgevonden (hieronder geduid als 'OA'). De tweede methode betrof het voeren van open gesprekken. Deze gesprekken zijn gevoerd met twee leden van het managementteam van de ADR (hieronder geduid als 'MT'), met verantwoordelijkheid voor de uitvoering van deze mogelijke opdracht.

De uitkomst hiervan is beperkt in scope. Het reflecteert louter de ADR, de interne auditor binnen de Rijksoverheid. Dat brengt inherente beperkingen met zich mee, met name ten aanzien van de overdraagbaarheid van de uitkomsten. Allereerst reflecteert het louter op de interne auditkant van het verhaal; niet op de toepasbaarheid van deze aanpak door een externe auditor, zoals de Algemene Rekenkamer. Daarnaast is het ook alleen een perspectief vanuit de interne afdeling: het perspectief van de opdrachtgever wordt bijvoorbeeld niet meegenomen. Andere interne auditdiensten binnen de Rijksoverheid (zoals die van de Dienst Justitiële Inrichtingen) of andere overheden (zoals de Gemeente Den Haag) zijn ook buiten scope gelaten. Ten slotte heeft de verkenning zich ook alleen beperkt tot de publieke sector.

Daarentegen moet daarbij de doelstelling van dit onderzoek in het achterhoofd gehouden worden. Het doel is om een verkenning uit te voeren naar de mogelijke toepasbaarheid van een andere manier van auditen. Het is daarmee geen multiple gevalstudie die vergelijkingen probeert te trekken tussen verschillende interne auditdiensten, of tussen de interne en externe auditor binnen de Rijksoverheid. Het doel is om, binnen de context van de ADR, een eerste beeld te krijgen van de toepasbaarheid van een dergelijke methodiek. Dit gezegd hebbende, zijn de gekozen methodieken van de focusgroep en de ongestructureerde interviews adequate methoden voor deze verkenning.

Om de onderlinge uitkomsten met elkaar te vergelijken, zijn in de focusgroep en de twee gesprekken dezelfde, open vragen gesteld. Voor de actiegerichte doelstelling waren dat: "Welke elementen van het plan van aanpak zie je als kracht?" en "2. Welke belemmeringen zie je bij deze aanpak?" Voor reflectieve doelstelling is de vraag gesteld: "3. In hoeverre beschouw je deze aanpak als een auditaanpak?" Van deze focusgroep en de gesprekken zijn gespreksverslagen gemaakt. Gebruikte citaten zijn voor

Eenzijds wilde ik het plan empirisch valideren om het te verbeteren. Anderzijds wilde ik ook de collectieve kennis van collega's benutten om te reflecteren op de toepasbaarheid van deze methode binnen de interne auditfunctie. Vandaar dat ik de reflectieve doelstelling heb toegevoegd aan de focusgroep en individuele gesprekken.

*Deze **selectiecriteria** waren de volgende. De collega moet in het bezit zijn van de RO-titel, omdat ik kennis over operational audit en internal audit wil benutten in de focusgroep, mede gelet op de tweede doelstelling. In aanvulling daarop moet de collega ofwel worden ingezet als opdrachtgerichte kwaliteitsbeoordelaar, ofwel lid zijn van de vakgroep operational audit. Verder heb ik het aantal jaren werkervaring binnen de ADR als criterium gekomen: minimaal vijf jaar. Vervolgens heb ik variëteit gezocht in de expertise van verschillende collega's. Deze betroffen onder andere: leergeschiedenisonderzoek, compliance-audit, assurance-audit, informatiebeveiliging, risicomanagement en duurzaamheid.*

hoor/wederhoor bij betrokkenen voorgelegd. Dat heeft niet tot aanpassingen geleid.

Uitkomsten van de aanvullende empirische validatie, qua actiegerichte doelstelling

Hieronder schets ik beknopt de uitkomsten van de focusgroep en de twee open gesprekken, gegroepeerd aan de twee vragen die zijn gesteld.

1. Welke elementen van het plan van aanpak zie je als kracht?

De “gestructureerde onderzoeks aanpak” *an sich* zagen meerderen als kracht (OA #2). Het is een “goed methodologisch onderbouwde aanpak die de verschillende fasen helder weergeeft” (OA #1). Een dergelijke aanpak helpt je als onderzoeker om te komen tot gedragen resultaten: “Je voorkomt hiermee dat je een ‘Commissie der Wijzen’ wordt die slimme dingen roept. Met de input van betrokkenen draag je actief bij aan draagvlak.” (OA #2). Hier werd ook door een MT-lid op gewezen: “Het ophalen en terugleggen van verhalen is een sterke manier [van onderzoek]. [...] Het gaat erom dat de [auditee] tot bewustwording komt.” (MT #1). Die bewustwording bij de auditee komt eveneens naar voren door het interactief betrekken van de auditee in alle fasen: “Normaliter onderzoeken wij de probleemanalyse en valideren we deze; [in de voorgestelde aanpak] schrijft de auditee ook echt mee aan de probleemdefinitie.” (OA #3). Met verhalen krijg je toegang tot de “wereld van de ervaringskennis” van de betrokkenen (MT #1). Met verschillende verhalen krijg je een breder en groter inzicht: het zijn verhalen die beeldend zijn en spreken tot de leefwereld van de betrokkenen. Zeker “voor de complexe beleidssituaties” kan deze methode helpen om de “meervoudige perspectieven op tafel te krijgen” (MT #1).

2. Welke belemmeringen zie je bij deze aanpak?

De belemmeringen kunnen geclusterd worden in twee groepen: de technieken en de selectie van belanghebbenden. Zo werd genoemd dat dit type onderzoek “vraagt om een specifieke set aan onderzoeks- en interviewvaardigheden die niet iedere [... auditor] heeft.” (MT #2). Een deelnemer gaf aan dat deze technieken binnen de ADR aanwezig zijn: “Ik herken de insteek voor de gesprekken vanuit Auditen vanuit de Bedoeling en [het leergeschiedenisonderzoek]” (OA #1). Daarnaast werd de selectie van belanghebbenden als delicaat proces geduid, waar *bias* in geslopen kan zijn. De zorg werd uitgesproken of er “bepaalde groepen niet gehoord of minder gehoord zijn, of dat er een allesoverheersend perspectief en belangen [zijn] die niet worden uitgesproken” (MT #1). In de focusgroep gaf een deelnemer vervolgens aan: “De kans is klein dat je over alle belanghebbende groepen heen geen verschillende belangen en beelden krijgt. [...] Je bent juist op zoek naar die tegengestelde punten.” (OA #1).

Uiteindelijk heeft deze aanvullende empirische validatie geleid tot een aanvulling op het plan, waarin meer gesprekstechnieken zijn geduid in de individuele gesprekken (zie hoofdstuk 3 van het andere onderdeel voor het resultaat). Voor het punt van de selectie van

En als kers op de taart kon ik meedraaien in een andere 'krent in de pap': een

belanghebbenden is tijdens de focusgroep en de open gesprekken benadrukt dat aan het einde van het gesprek aan de desbetreffende belanghebbende wordt gevraagd wie hij/zij zou nomineren om gesproken te worden. Zo voorkom je een eenzijdig perspectief vanuit het ministerie op de lijst met de te spreken belanghebbenden.

Uitkomsten van de aanvullende empirische validatie, qua reflectieve doelstelling

De uitkomsten van de derde vraag is verweven in de interne reflectie in het andere onderdeel van de bijdrage (zie hoofdstuk 3). Het is geenszins bedoeld als resultaat van een thematische noch discoursanalyse. Het is bedoeld als slotreflectie, waarin ik enkele contouren heb proberen te schetsen van twee logica's die mij opvielen. Gelet op de verkennende aard van dit onderzoek, benadruk ik dat dit louter *contouren* zijn van *mogelijke* logica's binnen de organisatie.

De academische korrels zout bij deze zoektocht

Het is belangrijk om academische kanttekeningen te plaatsen bij dit onderzoek. Allereerst betreft de conclusie een reflectie. Deze reflectie komt voort uit gesprekken louter met enkele senior collegae en managementleden binnen de ADR. Het is een construct binnen de specifieke context van een grote interne, publieke auditfunctie. Het kan daarom lastig te over te dragen zijn naar andere interne auditfuncties. Andere auditfuncties zouden met een ander construct kunnen komen. Daarnaast wil ik benadrukken dat ik dit een tentatieve analyse heb genoemd, waarin ik met enkele collegae de reflectie ben aangegaan over deze aanpak en de plek ervan binnen een interne auditfunctie. De toevoeging van andere respondenten zoals opdrachtgevers, leden van auditcommittees en minder senior collegae kan mogelijk tot andere inzichten leiden. Bovenal heeft deze bijdrage zich gericht op het ontwikkelen van een plan van aanpak en een reflectie daarop. Dat laat de uitvoering en de (gepercipieerde) toegevoegde waarde buiten scope van het onderzoek.

Desondanks biedt deze zoektocht waardevolle inzichten voor de internal auditpraktijk. Het toont dat er wel degelijk fundamenteel andere onderzoeksmethodieken mogelijk zijn om toe te passen als interne auditfunctie. En dat deze methodiek geschikt kan zijn voor organisaties bij ambigue en complexe vraagstukken; vraagstukken die zich vaker wel dan niet voordoen binnen de huidige VUCA-wereld. De methodiek zorgt voor onderlinge interactie en is dus meer dan alleen een onderzoek; het zorgt in interveniërende zin voor verbinding met de focus op het wederzijdse begrip. Deze zoektocht laat tevens zien dat het geen kwaad kan om bij aanpalende praktijkgerichte werkvelden over de schutting te kijken voor waardevolle inzichten en methodieken. Naast psychologie, antropologie en nog vele andere velden kan beleidsevaluatie voor mijn part ook worden toegevoegd aan deze rij.

leergeschiedenisevaluatie van een programma. Een participatieve evaluatie, waarin we werden ondersteund door een universitair hoofddocent uit Groningen: Rik Peters. Tijdens zijn begeleiding heeft hij ons handreikingen gegeven voor het gebruik van gesprekstechnieken, gebruikelijk in de leergeschiedenismethode. Deze bleken uitermate geschikt voor deze aanpak. Ik mocht zijn nog te publiceren teksten gebruiken voor deze scriptie.

In het andere onderdeel van deze bijdrage heb ik de conclusie van de verkennende zoektocht geschetst. Tezamen met dit onderdeel is de bijdrage voor het afstudeeronderzoek compleet. Het uitgebreide literatuuronderzoek heb ik in deze verantwoording beschreven, evenals een (over de cycli verspreide) beschouwing over de methodologie die onderliggend is geweest aan deze zoektocht. Het beschrijft de resultaten voor wat betreft de actiegerichte doelstelling van de empirische validatie. Voor het daadwerkelijke plan van aanpak – het resultaat van het literatuuronderzoek en de uitkomsten van het actiegerichte gedeelte van de empirische validatie – verwijs ik naar het andere onderdeel. Eveneens verwijs ik naar dat onderdeel voor mijn reflectie – mede gevoed door de reflectie van de focusgroep en de individuele gesprekken – op de (on)mogelijkheid van deze responsieve auditaanpak.

De weg naar toekomstige zoektochten

Naast de academische korrels zout, is het ook van belang wegen voor de toekomst te schetsen. In de hierboven beschreven cycli ben ik telkens voor nieuwe vragen komen te staan. Hieronder schets ik nog enkele vragen die een weg kunnen inslaan voor vervolgonderzoek.

Een allereerste vraag die openstaat: “Kan een ‘lichtere’ vorm van responsief evalueren wel worden beschouwd als een auditaanpak?” Deze aanpak kende immers een *empowerment* karakter. Een responsieve vorm die meer een samenwerkende of participerende karakter aanneemt, kan allicht wel de toets van de objectivistische auditlogica doorstaan (Van Heijster, e.a., 2020).

Een tweede vraag die bij mij leeft: “Is dit tentatieve beeld eveneens van toepassing binnen interne auditfuncties van zowel de publieke als private sector?” Een meervoudige gevalstudie kan een verdiepende inblik geven in verschillende interne auditfuncties hun constructies rondom een responsieve aanpak binnen (de) audit(functie). Om meer recht te doen aan de complexiteit van het vraagstuk, zouden daar ook andere respondenten dan louter internal auditors voor kunnen worden benaderd.

Een andere vraag die in deze laatste cyclus bij mij naar boven kwam, is: “Hoe interacteren de objectivistische logica en de subjectivistische logica binnen de ADR?” De reflectie had louter tot doel om de eerste contouren van deze twee logica’s te schetsen. Een verdiepende gevalstudie – met de ADR als analytische categorie – kan zich richten op de aard van de interactie tussen de discoursen: kunnen de discoursen vredig naast elkaar bestaan; of kunnen ze elkaar aanvullen; of voeren ze competitie met elkaar? Een dergelijk onderzoek zou ook eventuele verbanden kunnen proberen te leggen met de driedelige categorisatie van Eulrich en Lenz (2020) voor de ontwikkeling van de interne auditfunctie. Daarbij is het van belang om andere perspectieven dan alleen die van de interne auditor centraal te stellen.

Dit relateert zich dan ook aan een andere vraag die vervolgonderzoek zou kunnen beantwoorden: “In hoeverre komen de constructies van internal auditors en andere stakeholders overeen voor wat betreft de rol van internal audit en een responsieve aanpak?” Een mogelijke insteek daarbij binnen het Nederlandse publieke bestel zou – in navolging van Roussy (2013) – *role theory* centraal kunnen stellen en daar verschillende stakeholders bij betrekken. Te denken valt aan leden van auditcommittees, opdrachtgevers van audits, de Algemene Rekenkamer, directeurs van de afdelingen Financieel-Economische Zaken, enzovoorts.

Ten slotte zou vervolgonderzoek zich ook kunnen richten op de daadwerkelijk (gepercipieerde) toegevoegde waarde van een responsieve – of zelfs breder: postmoderne – aanpak binnen een singuliere gevalstudie of met een uitbreiding van meerdere gevallen. Een interessante insteek daarbij zou kunnen liggen in het verbinden

van bestuurskundige theorieën rondom *new public management* en *new public governance* met organisatiekundige theorieën over eerste, tweede en derde orde leren (Werkman, 2006).

Literatuurlijst

- Abma, T. A. (2003). Learning by telling: Storytelling workshops as an organizational learning intervention. *Management Learning*, 34(2), 221-240.
- Abma, T.A. (2004). Responsive evaluation: The meaning and special contribution to public administration. *Public Administration*, 82(4), 993-1012.
- Abma, T. A. (2005). Responsive evaluation: Its meaning and special contribution to health promotion. *Evaluation and Program Planning*, 28, 279-289.
- Abma, T. A. (2006). The practice and politics of responsive evaluation. *American Journal of Evaluation*, 27(1), 31-43.
- Abma, T. A., Greene, J. C., Karlsson, O., Ryan, K., Schwandt, T. A., & Widdershoven, G. A. M. (2001). Dialogue on dialogue. *Evaluation*, 7(2), 164-180.
- Abma, T. A., & Heijlsman, A. (2013). Crossing cultures: Health promotion for senior migrants in the Netherlands. *Health Promotion International*, 30(3), 1-13.
- Abma, T. A., Nierse, C. J., & Widdershoven, G. A. M. (2009). Patients as partners in responsive research: methodological notions for collaborations in mixed research teams. *Qualitative Health Research*, 19(3), 401-415.
- Abma, T. A., & Stake, R. E. (2001). Stake's responsive evaluation: Core ideas and evolution. *New Directions for Evaluation*, 92, 7-22.
- Abma, T. A., & Widdershoven, G. A. M. (2005). Sharing stories: Narrative and dialogue in responsive nursing evaluation. *Evaluation & the Health Professions*, 28(1), 90-109.
- Abma, T. A., & Widdershoven, G. A. M. (2008). Evaluation and/as social relation. *Evaluation*, 14(2), 209-225.
- Abma, T. A., & Widdershoven, G. A. M. (2011). Evaluation as a relationally responsible practice. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research* (4^e ed.). SAGE Publications Ltd.
- Bakker, M., & De Boer, F. (2021). Kwaliteitscriteria in kwalitatief onderzoek: welke termen hanteer je als kwalitatief onderzoeker? *KWALON*, 26(2), 88-96.
- Balane, M. A., Palafox, B., Palileo-Villanueva, L. M., McKee, M., & Balabanova, D. (2020). Enhancing the use of stakeholder analysis for policy implementation research: towards a novel framing and operationalized measures. *BMJ Global Health*, 5(11), 1-12.
- Baur, V. E., Van Elteren, A. H. G., Nierse, C. J., & Abma, T. A. (2010). Dealing with distrust and power dynamics: Asymmetric relations among stakeholders in responsive evaluation. *Evaluation*, 16(3), 233-248.
- Bemelmans-Videc, M., Perrin, B., & Lonsdale, J. (2017). *Making accountability work: Dilemmas for Evaluation and for Audit*. Routledge.
- Bernstein, R. J. (1983). *Beyond objectivism and relativism: Science, hermeneutics, and praxis*. University of Pennsylvania Press.
- Beuving, J., & de Vries, G. (2015). *Doing qualitative research: The craft of naturalistic inquiry*. Amsterdam University Press.

- Boonstra, J. J. (2000). *Lopen over water: Over dynamiek van organiseren, vernieuwen en leren*. Vossiuspers.
- Bos, P., De Korte, R., & Otten, J. (2017). *Management control auditing: Bijdragen aan doelrealisatie en verbetering*. Auditing.nl
- Brown, J. M. (2017). A dialogical research methodology based on Buber: intersubjectivity in the research interview. *Journal Of Family Therapy*, 39, 415-436.
- Burgess-Limerick, T., & Burgess-Limerick, R. (1998). Conversational interviews and multiple-case research in psychology. *Australian Journal of Psychology*, 50(2), 63-70.
- Burr, V., & Dick, D. (2017). Social constructionism. In B. Gough (Ed.), *The Palgrave Handbook of Critical Social Psychology* (pp. 59-80). Palgrave Macmillan.
- Bussman, M., Kuiper, C., & Maas, A. (2013). Klinkende meerstemmige verhalen. *KWALON*, 18(3), 47-54.
- Camargo-Borges, C., & Rasera, E. F. (2013). Social constructionism in the context of organization development: Dialogue, imagination, and co-creation as resources of change. *SAGE Open*, 1-7.
- Carstensen, M. B., & Schmidt, V. A. (2016). Power through, over and in ideas: conceptualizing ideational power in discursive institutionalism. *Journal of European Public Policy*, 23(3), 318-337.
- Chelimsky, E. (1985). Comparing and contrasting auditing and evaluation: Some notes on their relationship. *Evaluation Review*, 9(4), 483-503.
- Chia, R. (1996). The problem of reflexivity in organizational research: Towards a postmodern science of organization. *Organization Articles*, 3(1), 31-59.
- Clarke, A. (1999). *Evaluation Research*. SAGE Publications Ltd.
- Cox, R. W. (1981). Social forces, states and world orders: Beyond international relations theory. *Millennium: Journal of International Studies*, 10(2), 126-155.
- Cunliffe, A. L. (2008). Orientations to social constructionism : Relationally responsive social constructionism and its implications for knowledge and learning. *Management Learning*, 39(2), 123-139.
- Czarniawska, B. (2004). *Narratives in social science research*. SAGE Publications Ltd.
- Daigneault, P. (2013). The blind men and the elephant: A metaphor to illuminate the role of researchers and reviewers in social science. *Methodological Innovations*, 8(2), 82-89.
- De Korte, R., Otten, J., & Schuiten, F. (2021). Wendbaarheid met internal audit(deel)producten. *AuditMagazine*.
- Dirsmith, M. W., Covalski, M. A., & McAllister, J. P. (1985). Of paradigms and metaphors in auditing thought. *Contemporary Accounting Research*, 2(1), 46-68.
- Driessen, A. J. G., & Molenkamp, A. (2012). *Internal auditing: Een managementkundige benadering* (5^e ed.). Boom.
- Edley, N. (2001). Unravelling social constructionism. *Theory & Psychology*, 11(3), 433-441.
- Edwards, D., Ashmore, M., & Potter, J. (1995). Death and furniture: the rhetoric, politics and theology of bottom line arguments against relativism. *History of the human sciences*, 8(2), 25-49.

- Eulrich, M., & Lenz, R. (2020). Defining, measuring and communicating the value of internal audit: Best practices for the profession. *Internal Audit Foundation*.
- Fairclough, N. (1993). Critical discourse analysis and the marketization of public discourse: The universities. *Discourse & Society*, 4(2), 133-168.
- Fletcher, D. (2006). Entrepreneurial processes and the social construction of opportunity. *Entrepreneurship and Regional Development*, 18(5), 421-440.
- Francis, J. R. (1994). Auditing, hermeneutics, and subjectivity. *Accounting, Organizations and Society*, 19(3), 235-269.
- Fraser, H. (2004). Doing narrative research: Analysing personal stories line by line. *Qualitative Social Work*, 3(2), 179-201.
- Gadamer, H. (1975). Hermeneutics and social science. *Philosophy & Social Criticism*, 2(4), 307-316.
- Galbin, A. (2014). An introduction to social constructionism. *Social Research Reports*, 26, 82-92.
- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz. *The Interpretation of cultures*. Basic Books, Inc.
- Geertz, C. (1983). *Local Knowledge*. Basic Books, Inc.
- Grandy, G. (2018). An introduction to constructionism for qualitative researchers in business and management. In C. Cassell, A. L. Cunliffe & G. Grandy (Eds.), *The SAGE Handbook of qualitative business and management research methods: History and traditions* (pp. 173-184). SAGE Publications Ltd.
- Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice*, 18(1), 25-35.
- Greene, J. C. (2001). Dialogue in evaluation: A relational perspective. *Evaluation*, 7(2), 181-187.
- Gregory, A. (2000). Problematizing participation: A critical review of approaches to participation in evaluation theory. *Evaluation*, 6(2), 179-199.
- Van Grinsven, M. (2021). Organisatieverandering: 'meervoudigheid' in perspectief. In M. de Witte, M. J. Vink & M. van Grinsven (Eds.). *Essenties van verandermanagement: een kleine canon van veranderkundige benaderingen*. Boom.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. SAGE Publications Inc.
- Hawkesworth, M. E. (1989). Knowers, knowing, known: Feminist theory and claims of truth. *Signs*, 14(3), 533-557.
- Hudaib, M., & Haniffa, R. (2009). Exploring auditor independence: An interpretive approach. *Accounting, Auditing & Accountability Journal*, 22(2), 221-246.
- Hosking, D. M. (2011). Telling tales of relations: Appreciating relational constructionism. *Organization Studies*, 32(1), 47-65.
- IIA (n.d.). *Definition of Internal Auditing*. The Institute of Internal Auditors.
<https://www.theiia.org/en/standards/what-are-the-standards/definition-of-internal-audit/>
- IIA (2009). *IIA Position Paper: The role of internal auditing in enterprise-wide risk management*. The Institute of Internal Auditors.
<https://www.theiia.org/globalassets/documents/resources/the-role-of-internal-auditing-in-enterprise-wide-risk-management-january-2009/pp-the-role-of-internal-auditing-in-enterprise-risk-management.pdf>

- Johnson, P., & Duberley, J. *Understanding management research*. SAGE Publications Ltd.
- Karlsson, O. (2001). Critical dialogue: Its value and meaning. *Evaluation*, 7(2), 211-227.
- Kauko, J., & Pajunen, M. K. T. (2021). The thickening modern: developing a research agenda beyond intensifying rationalism. In P. Dahler-Larsen (Ed.). *A research agenda for Evaluation* (pp. 21-42). Edward Elgar Publishing.
- Kushner, S. (1996). The limits of constructivism in evaluation. *Evaluation*, 2(2), 189-200.
- Laughlin, R., & Broadbent, J. (1996). Redesigning fourth generation evaluation: An evaluation model for the public-sector reforms in the UK? *Evaluation*, 2(4), 431-451.
- Lenz, R., & Hahn, U. (2015). A Synthesis of Empirical Internal Audit Effectiveness Literature Pointing to New Research Opportunities. *Managerial Auditing Journal*, 30, 1-42.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE Publications Inc.
- Marsh, D., & Furlong, P. (2002). A skin, not a sweater: Ontology and epistemology in political science. In D. Marsh & G. Stoker (Eds.). *Theory and methods in political science* (pp. 17-41). Pallgrave MacMillan.
- Mayne, J. (2006). Audit and evaluation in public management: Challenges, reforms, and different roles. *The Canadian Journal of Program Evaluation*, 21(1), 11-45.
- Meadows L., Morse J. M. (2001). Constructing evidence within the qualitative project. In J. M. Morse, J. Swanson, A. Kuzel, (Eds.). *The nature of evidence in qualitative inquiry* (pp. 187–202), SAGE Publications Ltd.
- Mertens, D. M. (2010). Transformative mixed methods research. *Qualitative Inquiry*, 16(6), 469-474.
- Mertens, D. M. (2017). Transformative research: personal and societal. *International Journal for Transformative Research*, 4(1), 18-24.
- Milliken, J. (1999). The study of discourse in international relations: A critique of research and methods. *European Journal of International Relations*, 5(2), 225-254.
- Moerkerken, S. (2020). Een kwestie van verschil: Sociaal-constructionistisch interveniëren volgens Henk van Dongen. *M&O*, 5(6), 100-109.
- Moses, J. W., & Knutsen, T. (2012). *Ways of knowing: Competing methodologies in social and political research* (2e ed.). Palgrave Macmillan.
- Naidoo, I. (2020). Audit and evaluation: Working collaboratively to support accountability. *Evaluation*, 26(2), 177-189.
- Nierse, C. J., & Abma, T. A. (2011). Developing voice and empowerment: the first step towards a broad consultation in research agenda setting. *Journal of Intellectual Disability Research*, 55(4), 411-421.
- Niessen, T., Abma, T. A., Widdershoven, G. A. M., & Van der Vleuten, C. (2008). Learning-in-(inter)action: A dialogical turn to evaluation and learning. In K. Ryan & J. B. Cousins (Eds.), *The SAGE International Handbook for Educational Evaluation*. SAGE Publications Ltd.
- Nordin, I. G. (2022). Narratives of internal audit: The Sisyphean work of becoming “independent”. *Critical Perspectives on Accounting*.
- Peters, R. G. P. (2023). *Learning histories: Vat krijgen op je organisatiecultuur* [Ongepubliceerd manuscript]. Noordhoff Uitgevers.

- Powell, M. (1999). *The audit society: Rituals of verification*. Oxford University Press.
- Raad voor Cultuur (2022). *Beweging in het bestel: Verkenning op weg naar een nieuw advies over het cultuurbestel*.
- Regeling op het specifiek cultuurbeleid (2021, 1 juli). Geraadpleegd van <https://wetten.overheid.nl/BWBR0027597/2020-09-08#Hoofdstuk6>
- Roman, B., & Waterschoot, K. (2018). Het verhaal als kwaliteitsinstrument: Een toepassing van narratief onderzoek voor de ouderenzorg. *Tilburg University*. <https://www.zorgvoorbeter.nl/zorgvoorbeter/media/documents/thema/persoonsgerichte-zorg/handreiking-verhaal-als-kwaliteitsinstrument.pdf>
- Roussy, M. (2013). Internal auditors' roles: From watchdogs to helpers and protectors of the top manager. *Critical Perspectives on Accounting*, 24(7), 550-571.
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*, 52, 1893-1907.
- Schmidt, V. A. (2008). Discursive institutionalism: The explanatory power of ideas and discourse. *Annual Review of Political Science*, 11, 303-326.
- Schmidt, V. A. (2010). Taking ideas and discourse seriously: explaining change through discursive institutionalism as the fourth 'new institutionalism'. *European Political Science Review*, 2(1), 1-25.
- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118–137). Sage Publications, Inc.
- Schwandt, T. A. (1997). Evaluation as practical hermeneutics. *Evaluation*, 3(1), 69-83.
- Schwandt, T. A. (1999). On understanding understanding. *Qualitative Inquiry*, 5(4), 451-464.
- Schwandt, T. A. (2001). A postscript on thinking about dialogue. *Evaluation*, 7(2), 264-276.
- Schuiten, F., & De Korte, R. (2022). Wat is inzicht en wat moet je ermee? *AuditMagazine*.
- Sebele-Mpofu, F. Y. (2020). Saturation controversy in qualitative research: Complexities and underlying assumptions. A literature review. *Cogent Social Sciences*, 6(1), 1-17.
- Smit, M., & Verdonschot, S. (2010). *Praktijkonderzoek: Motor voor verandering in organisaties*. Boh Stafleu van Loghum.
- Snoeren, M. M. W. C., Niessen, T. J. H., & Abma, T. A. (2013). Beyond dichotomies: Towards a more encompassing view of learning. *Management Learning*, 6(2), 137-155.
- Shotter, J. (1995). In dialogue: Social constructionism and radical constructivism. In L. Steffe & J. Gale (Eds.), *Constructivism in Education* (pp. 41-56). Routledge.
- Shotter, J. (2010). Movements of feeling and moments of judgment: Towards an ontological social constructionism. *International Journal of Action Research*, 6(1), 16-42.
- Soderberg, A. (2006). Narrative interviewing and narrative analysis in a study of a cross-border merger. *Management International Review*, 46(4), 397-416.
- Sools, A. (2012). Narratief onderzoek. *KWALON*, 17(1), 27-35.
- Stake, R. E. (1975). *Evaluation the arts in education: A responsive approach*. Merrill.

- Türetken, O., Jethefer, S., & Ozkan, B. (2020). Internal audit effectiveness: operationalization and influencing factors. *Managerial Auditing Journal*, 35(2), 238-271.
- Van Berkel, J., Van Heijster, H., Abma, T. A., Boot, C., & De Vet, E. (2022). *Onderzoekershandleiding: Responsieve evaluatie van gezondheidsbevordering in de werksetting*. Wageningen University & Research.
- Van der Arend, S.H. (2007). *Pleitbezorgers, procesmanagers en participanten: Interactief beleid en de rolverdeling tussen overheid en burgers in de Nederlandse democratie*. Eburon Uitgeverij B.V.
- Van der Haar, D., & Hosking, D. M. (2004). Evaluating appreciative inquiry: a relational constructionist perspective. *Human relations*, 57(8), 1017-1036.
- Van Heijster, H., Van Berkel, J., Abma, T. A., Boot, C. R. L., & De Vet, E. (2020). Responsive evaluation of stakeholder dialogue as a worksite health promotion intervention to contribute to the reduction of SEP related health inequalities: a study protocol. *BMC Health Services Research*, 20(196), 1-9.
- Van der Meulen, I., & Otten, J. (2013a). Behavioural auditing: het onderzoeken van gedrag in organisaties. *Auditing Magazine*, 12(1), 33-35.
- Van der Meulen, I., & Otten, J. (2013b). Behavioural auditing: het onderzoeken van gedrag in organisaties. *Auditing Magazine*, 12(2), 37-39.
- Van Twist, M. (2018). *Woorden wisselen: Werken aan een hertaling van besturen, organiseren en adviseren*. Management IMPACT.
- Van Twist, M., & De Korte, R. (2022). Auditing: inspiratie voor innovatie vanuit de sociale wetenschappen. *Maandblad voor Accountancy en Bedrijfseconomie*, 96(7/8), 177-185.
- Vanderplaat, M. (1995). Beyond technique: Issues in evaluating for empowerment. *Evaluation*, 1(1), 81-96.
- Verschuren, P., & Doorewaard, H. (2021). *Het ontwerpen van een onderzoek* (6^e ed.). Boom.
- Verwoerd, L., Klaasen, P., & Regeer, B. J. (2019). How to normalize reflexive evaluation? Navigating between legitimacy and integrity. *Evaluation*, 27(2), 229-250.
- Visse, M., Abma, T. A., & Widdershoven, G. A. M. (2011). Relational responsibilities in responsive evaluation. *Evaluation and Program Planning*, 35, 97-104.
- Wadsworth, Y. (2001). Becoming responsive – and some consequences for evaluation as dialogue across distance. *New Directions for Evaluation*, 92, 45-58.
- Werkman, R. A. (2006). *Werelden van verschil: hoe actoren in organisaties vraagstukken in veranderprocessen hanteren en creëren*. Universiteit van Amsterdam.
- Widdershoven, G. A. M. (2001). Dialogue in evaluation: A hermeneutic perspective. *Evaluation*, 7(2), 253-263.
- Wiss, E. (2020). *Socrates op sneakers: Filosofische gids voor het stellen van goede vragen*. Ambo|Anthos.
- Woelders, S., & Abma, T. A. (2019). Participatory action research to enhance the collective involvement of residents in elderly care: About power, dialogue and understanding. *Action Research*, 17(4), 528-548.