



Rijksoverheid

Hoe is gedrag te onderzoeken?

Overzicht van 18 onderzoeksmethodieken voor effectiever beleid

Behavioural Insights Netwerk Nederland

bin_{-nl}

Hoe is gedrag te onderzoeken?

Overzicht van 18 onderzoeksmethodieken voor effectiever beleid

Behavioural Insights Netwerk Nederland

bin_{-nl}

Voorwoord

Gedrag veranderen is complex. Waarom vertonen mensen ongewenst gedrag? Hoe zijn ze aan te zetten tot het gewenste gedrag? Onderzoek kan helpen deze kluwen te ontwarren om zo beter, effectiever beleid te maken. Maar hoe pak je dat onderzoek aan? Deze uitgave laat je 18 mogelijke onderzoeksmethodieken zien en licht deze toe met concrete praktijkvoorbeelden en interviews.

Deze publicatie is bedoeld voor ambtenaren die gedragsinzichten (willen) toepassen, om daarmee beleid, communicatie, uitvoering en handhaving effectiever te maken. We bieden een overzicht van verschillende onderzoeksmethoden en waarvoor je ze kunt gebruiken. Zo kun je beter bepalen op welke manier je een gedragsvraagstuk uit je eigen praktijk kunt (laten) onderzoeken. Dit kan behulpzaam zijn in adviezen aan collega's of gesprekken met onderzoekers. Basale kennis van onderzoek wordt wel al verondersteld in deze publicatie.

Door goed te onderzoeken voorkom je dat beleid wordt gebaseerd op onjuiste aannames over de doelgroep. Of op wensdenken: 'mensen hebben veel interesse in mijn beleidsonderwerp', 'mensen verwerken informatie altijd op een rationele manier' of 'als mensen weten hoe het zit, gaan ze vanzelf hun gedrag aanpassen'. Onderzoek geeft je bovendien meer grip op de juiste interventies om gedrag te veranderen. En je kunt het gebruiken om de effecten van je interventies aan te tonen en daarmee je activiteiten te verantwoorden.

Tot slot, wat we anderen vaak adviseren over gedrag is onverkort van toepassing op onderzoek doen: houd het zo simpel mogelijk, begin op tijd, maak het leuk, betrek anderen erbij en leer ervan!

Joost Loef

Voorzitter werkgroep onderzoeksmethodieken BIN NL

Inhoud

Waarom juist deze onderzoeksmethodieken?	4
Hoe kies je een methode voor je eigen onderzoek?	5
Diepte-interviews	10
Focusgroepen	12
Dagboekmethode	14
Klantreizen	16
Ongeschreven regels	20
Enquêteonderzoek	22
Keuze-experimenten	24
Impliciet meten via reactietijdentaken	26
Big data analytics	28
Research community	32
Waarderend onderzoeken	34
Public design	38
Randomized Controlled Trial	40
Praktijkexperimenten	44
Propensity Score Matching	46
Regression Discontinuity	48
Difference-in-Difference	50
Voor- en nameting	54
Meer weten over BIN NL?	58

Waarom juist deze onderzoeksmethodieken?

Deze publicatie bevat 18 onderzoeksmethodieken die kunnen helpen om grip te krijgen op gedragsverandering. Daarmee ontstaat geen uitputtend overzicht van wat er onderzoekstechnisch allemaal mogelijk is of in de markt beschikbaar. Wel krijg je een breed beeld van relevante methoden voor de Rijksoverheid. Hoe zijn de methodieken geselecteerd?

Gericht op gedragsonderzoek

De geselecteerde methodieken zijn geschikt voor gedragsonderzoek. Of een onderzoeksvraag gaat over gedrag is gemakkelijk na te gaan: gedrag is iets waar je een foto van kunt maken. Zo is 'bewustwording' geen gedrag, maar 'een formulier invullen' wel. Gedrag is daarmee wezenlijk anders dan een voornemen om iets te doen of een positieve houding ten opzichte van bepaald gedrag.

Voor een specifiek onderzoeksdoel

Gegevens over specifiek gedrag zijn in de praktijk lang niet altijd aanwezig. In deze publicatie ligt de nadruk op methodieken die deze gegevens speciaal voor het onderzoeksdoel verzamelen. Daarom komen literatuurstudie en deskresearch er bijvoorbeeld niet in voor. Wel maakt bijvoorbeeld big data-analyse slim gebruik van gegevens die al voor andere doeleinden verzameld zijn.

Uit sociale wetenschappen, beproefd bij het Rijk

De onderzoeksmethoden komen vooral uit de sociale wetenschappen. We hebben methodieken gekozen die vaak toegepast worden voor gedragsonderzoek en waarmee ruime ervaring is opgedaan binnen de Rijksoverheid. Bij de planbureaus (Centraal Planbureau, Planbureau voor de Leefomgeving, Sociaal en Cultureel Planbureau), bij overheidsdiensten (Auditdienst Rijk, Dienst Publiek en Communicatie) en bij het Centraal Bureau voor de Statistiek.

Daardoor zijn andere interessante methodes afgefallen, zoals longitudinaal onderzoek (hoe gedrag zich ontwikkelt op lange termijn), *participatory narrative inquiry* (voor gedrag in complexe situaties) en psychofysiologische metingen van gedrag (zoals oogbewegingen of hersenactiviteit).

Hoe kies je een methode voor je eigen onderzoek?

Welke van deze 18 methodieken is geschikt om je onderzoeksvraag te beantwoorden? Welke passen bij de verschillende stappen in de beleidscyclus: verkennen, beleid ontwikkelen en evalueren? En op basis van welke afwegingen kies je een methode?

Welk type onderzoek in welke beleidsfase?

De rol van onderzoek verschilt per beleidsfase. In de verkennende fase kan onderzoek inzicht geven in je doelgroep en in het gedragsvraagstuk dat centraal staat. Tijdens de beleidsontwikkeling, als je interventies ontwerpt, kun je met onderzoek tot onderbouwde keuzes komen. Bij de beleidsevaluatie kun je toetsen of de gedragsinterventie het beoogde effect gaat hebben (vooraf) of heeft gehad (achteraf) bij je specifieke doelgroep.

In de tabel op de volgende pagina staan 18 methodieken. Sommige zijn wellicht al bekend, andere minder. Allemaal zijn ze relevant voor onderzoek naar gedrag en gedragsverandering. Aangegeven staat in welke beleidsfase ze van pas komen.

Welke methodiek de meest geschikte is voor je onderzoek, hangt van veel factoren af. Hieronder staan enkele overwegingen: in het algemeen en per beleidsfase.

Algemene overwegingen: wat is wenselijk en mogelijk?

Kies voor een methodiek die zo 'hard' mogelijke data oplevert

Voor gedragsonderzoek zijn dat gegevens over werkelijke gedragingen. Observeren van gedrag heeft daarmee de voorkeur boven vragen naar gedrag. Wanneer je iets wilt weten over koopgedrag, kun je beter de gegevens van een webwinkel gebruiken dan een vraag stellen op de site van de webwinkel. In de praktijk is gestructureerde observatie van gedrag (denk aan geweld tegen politie of ambulancepersoneel) lang niet altijd haalbaar of wenselijk. Dan moet je werken met een benadering of simulatie van het 'werkelijke' gedrag of ernaar vragen.

Laat onderzoeksmethoden elkaar versterken

Harde data geven slechts beperkt inzicht in onderliggende motivaties van de doelgroep of de context waarin het gedrag plaatsvindt. Methoden met een kwalitatieve component, bijvoorbeeld diepte-interviews, brengen je vaak dichterbij de doelgroep en hun beleving. Je kunt open vragen stellen en doorvragen. Schat altijd de uitkomsten op waarde: mensen weten lang niet altijd waarom ze iets doen en als ze het wel weten, dan willen ze dat in onderzoek, bijvoorbeeld in vragenlijsten, niet altijd zeggen. Andere onderzoeksmethoden, bijvoorbeeld die reactietijden meten, kunnen juist de meer onbewuste processen blootleggen. Zo kunnen diverse onderzoeksmethodieken elkaar aanvullen om een volledig beeld te geven van het gedrag en de doelgroep.

	Verkennde fase	Beleidsontwikkeling	Evaluatie
	<i>Wat is het gedragsprobleem? Wie vertonen het (on)gewenste gedrag? Hoe vaak komt het (on)gewenste gedrag voor? Waar en wanneer? Welke factoren bepalen het (on-)gewenste gedrag?</i>	<i>Hoe is het ongewenste gedrag bij de doelgroep(en) te veranderen/het gewenste gedrag te bevorderen?</i>	<i>Hoe kansrijk zijn concrete, voorgenomen gedragsinterventies? Hoe effectief was de ingezette interventie voor de gedragsverandering? Welk effect trad op bij wie? En waarom?</i>
Methodiek			
Diepte-interviews	X	X	X
Focusgroepen	X	X	
Dagboekmethode	X		
Klantreizen	X	X	
Ongeschreven regels	X	X	
Enquêteonderzoek	X	X	X
Keuze-experimenten	X		
Impliciet meten	X		X
Big data analytics	X		X
Research community	X	X	
Waarderend onderzoeken	X	X	
Public design	X	X	
Randomized Controlled Trial			X
Praktijkexperimenten	X		X
Quasi-experimenteel onderzoek*			X
Voor- en nameting			X

* Hier vallen drie methodieken onder: Propensity Score Matching, Regression Discontinuity en Difference-in-Difference.

Overweeg de ethische aspecten van het onderzoek

Bij wetenschappelijk onderzoek worden de voorgestelde methode en interventies vooraf grondig getoetst door een ethische commissie. Ook in onderzoek bij beleid is het belangrijk om vooraf goed in kaart te brengen welke ethische aspecten er spelen bij het onderzoek, en daarbij weloverwogen beslissingen te nemen. Wat zijn mogelijk negatieve gevolgen van het onderzoek voor de verschillende betrokkenen? Denk daarbij aan privacy, transparantie en gelijkwaardigheid.

Kies de juiste doelgroep voor het onderzoek

In verkennend onderzoek is het nuttig om de doelgroep ruim te formuleren, om zo te kunnen ontdekken welke relevante verschillen er tussen groepen bestaan. Bij het ontwikkelen en evalueren van interventies moet je juist goed afbakenen welke groep of groepen je in het onderzoek wilt betrekken. Wie wil je met de interventie bereiken?

Plan lang genoeg vooruit

Schat in hoe lang de looptijd van het onderzoek wordt bij een bepaalde methode, en welke dataverzameling nodig is. Om een energiebesparingsmaatregel te onderzoeken moet je bijvoorbeeld rekening houden met het stookseizoen. Elke evaluatie vraagt een goede planning. Ook als je alleen een meting achteraf gaat uitvoeren, moet je die al voorbereiden voor je de interventie inzet.

Ga na wat er in je budget past

Onderzoekskosten kunnen variëren van de prijs voor een auto tot die voor een huis. Wat voor auto of huis (goedkoop of duur) hangt af van de invulling van het onderzoek. Voor een snelle peiling in een *research community* kun je met een paar duizend euro al klaar zijn.

Een enquête is vaak duurder, maar dat varieert ook, afhankelijk van hoe uitgebreid en gedegen je het gedrag of de gedragsverandering wilt meten. Soms blijken de onderzoekskosten in de praktijk verrassend: zo kan een online experiment vaak heel snel en relatief goedkoop worden uitgevoerd, terwijl een kwalitatieve evaluatie met veel stakeholders lang en duur kan uitvallen.

Verkennend onderzoek: wat weet je al over het gedragsvraagstuk?

Start met literatuurstudie en deskresearch

Voer voordat je aan een onderzoek begint, literatuurstudie en deskresearch uit om te achterhalen welke wetenschappelijke en praktijkkennis er al is over het gedrag en de doelgroep. Zo kom je al veel te weten en kun je de onderzoeksvragen richten op wat nog niet bekend is.

Onderzoek zo open mogelijk of juist gericht

Onderzoek als er nog maar weinig bekend is, je gedragsvraagstuk op een zo open mogelijke manier, zonder te sturen in een bepaalde richting of te leunen op aannames. Dat kan in bijvoorbeeld een research community of met klantreizen. Zoek als er al meer bekend is, juist gericht naar specifieke, ontbrekende kennis. Dat kan vooral met enquêteonderzoek, dagboekmethode, impliciet meten en keuze-experimenten; methodieken die vaak al voorkennis van het onderwerp vereisen.

Beleidsontwikkeling: hoe wil je komen tot een interventie?

Inventariseer kennis over interventies

Breng eerst systematisch de kennis over interventies die in het verleden zijn ingezet, in kaart. Soms zijn er zelfs meta-analyses uit de wetenschappelijke literatuur beschikbaar, waarmee je een indicatie kunt krijgen van de effectgroottes van verschillende interventies. Zo kun je relatief snel en onderbouwd interventies selecteren die in een andere context effectief waren. Het is wel nodig om de beoogde interventie vervolgens nog op effect te testen, omdat een klein verschil in context en doelgroep al kan leiden tot een groot verschil in effect. Wat in de ene situatie werkt, hoeft dat in de andere situatie niet te doen.

Put uit verkennend onderzoek

Benut het verkennend onderzoek om op ideeën te komen over interventies, als er nog weinig kennis is over interventies voor jouw specifieke gedragsvraagstuk. De stap van inzicht in factoren die gedrag bepalen naar interventies is echter niet altijd makkelijk te zetten. Ook hier zijn verschillende methodieken behulpzaam. Je kunt via onderzoek mogelijke interventies aan de doelgroep voorleggen, bijvoorbeeld in focusgroepen of interviews. Of deze interventies samen met de doelgroep ontwikkelen, zoals bij waarderend onderzoeken en *public design*. De gekozen interventie(s) kun je vervolgens testen op verwachte effectiviteit.

Evaluatie: wat is haalbaar bij het testen van de effectiviteit van interventies?

Test interventie vooraf

Test de interventie waar mogelijk ook vooraf bij de doelgroep waar de interventie zich op richt. Je kunt dan de uitkomsten van het onderzoek nog gebruiken om te bepalen of je doorgaat met de interventie of niet, of dat je die moet bijstellen. Heel zinvol, maar praktisch niet altijd haalbaar vanwege beperkte tijd en/of budget. Een kleine test vooraf (bijvoorbeeld enkele interviews met de doelgroep) is vrijwel altijd aan te bevelen boven niet testen. Heb daarbij niet alleen oog voor de beoogde effecten, maar ook voor (gewenste en ongewenste) neveneffecten.

Experiment of effectmeting?

Kies voor een evaluatie achteraf uit een experiment of een effectmeting. In een experimentele opzet krijgt één groep de interventie (de interventiegroep) en een andere groep niet (de controlegroep). In een effectmeting vergelijk je resultaten uit een meting vóór en een meting ná de interventie. Een experiment is het meest geschikt om een oorzakelijk verband aan te tonen: treedt het effect op door de interventie, of kan het ook door iets anders komen? In de praktijk kun je niet altijd voor zo'n experiment kiezen. Sommige interventies, zoals een massamediale campagne, zijn niet te testen in een experimentele opzet, omdat de hele doelgroep potentieel de interventie heeft ontvangen zodra de campagne live gaat.

Lab of veld?

Soms wordt gedrag getest in een laboratoriumsituatie, een gecontroleerde omgeving, en soms in een praktijksituatie, een

natuurlijke omgeving. Die laatste situatie heet een veldexperiment. Effecten in het lab komen onder gecontroleerde condities tot stand en zijn daarmee slechts een benadering van het werkelijke effect. Effecten in het veld zijn realistischer, maar ook gevoeliger voor (niet te controleren) externe factoren. Verder kost een veldexperiment vaak meer capaciteit.

Overweeg quasi-experimenteel onderzoek

Bij experimenten zoals een Randomized Controlled Trial (RCT) wijs je mensen willekeurig toe aan de interventiegroep en de controlegroep. Door dit random te doen is het vrij zeker dat verschillen in gedrag tussen groepen ook echt het gevolg zijn van de interventie. Als deze toewijzing niet-willekeurig is of kan, kun je met quasi-experimenteel onderzoek soms een controlegroep construeren die zo veel mogelijk lijkt op de interventiegroep.

En nu?

Alle methodieken hebben hun eigen sterke en zwakke kanten. Deze lichten we in deze publicatie toe, met suggesties om de nadelen te ondervangen. We geven er ook steeds voorbeelden bij uit de praktijk van de Rijksoverheid. Wil je zelf een onderzoek (laten) uitvoeren? Betrek dan een onderzoeker of andere collega die hierover kan adviseren, of raadpleeg een contactpersoon bij BIN NL. Zie binnl.nl.

Ervaringen lezen van collega's?

Lees de zes interviews van collega's die een onderzoeksmethodiek hebben gebruikt in hun organisatie:

- klantreizen bij de Belastingdienst
- big data analytics bij het CBS
- waarderend onderzoek bij Rijkswaterstaat
- RCT bij de Autoriteit Consument en Markt (ACM)
- quasi-experiment bij het ministerie van Economische Zaken en Klimaat (EZK)
- voor- en nameting bij het ministerie van Infrastructuur en Waterstaat (IenW).

Zie voor de resultaten van veel (gedrags)onderzoeken ook de uitgave 'Rijk aan gedragsinzichten - editie 2019' van BIN NL.

Onderzoeksmethodiek 1:

Diepte-interviews

Diepte-interviews zijn vraaggesprekken van doorgaans een half uur tot een uur over een bepaald onderwerp met meestal één persoon, bijvoorbeeld een burger, stakeholder of inhoudsdeskundige. Ze zijn bedoeld om een onderwerp te verkennen en lenen zich goed voor het bespreken van individuele gedragingen. Ze leveren diepgaand inzicht op in onder meer het waaróm van bepaald gedrag.

Methode: open gesprekken voor inzicht

Diepte-interviews kunnen – in het kader van gedragsonderzoek – gaan over bepaald gedrag of een gedragspatroon, over de situaties en omstandigheden waarin het gedrag plaatsvindt en over achterliggende motivaties. De gesprekken zijn, anders dan bij een enquête, half of niet gestructureerd. Er zijn wel vraagpunten, maar de exacte vraagstelling en de volgorde daarvan liggen niet vast. Er is gelegenheid voor de interviewer om door te vragen op de antwoorden die de respondent geeft.

De gesprekken vinden live plaats (*face to face*) of online. Als ze worden gevoerd in de omgeving waar het gedrag zich voordoet (thuis of op de werkvloer), kan dat extra inzicht geven. Vaak wordt gekozen voor 5 tot 25 gesprekken; individueel, om te voorkomen dat de antwoorden beïnvloed worden door de antwoorden van andere deelnemers. De gesprekken worden opgenomen (en soms letterlijk uitgeschreven) en geanalyseerd door de interviewer. Ook zijn met kwalitatieve analyseprogramma's patronen in de antwoorden te ontdekken die iets zeggen over de gedragingen en houdingen van de doelgroep. Soms krijgen geïnterviewden een samenvattend gespreksverslag te zien om te checken of hun verhaal goed is weergegeven. Dit gebeurt vooral bij stakeholders of inhoudsdeskundigen.

Toepassing: van verkenning en ontwerp tot evaluatie

Diepte-interviews zijn geschikt om gedragsmechanismen te achterhalen. Welke gedragingen, ervaringen en meningen zijn er en waar worden die door veroorzaakt? Dit wil je vaak weten in een verkennende fase voor beleid. Maar de methode is ook geschikt voor de ontwerpfase, als je zoekt naar een geschikte interventie om het beleidsdoel te bereiken. Hoe reageren mensen daarop en

waarom? Tot slot komt de methode ook van pas in de evaluatiefase. Niet om de effecten te meten, maar wel om mensen te vragen naar bijvoorbeeld de motieven achter hun gedrag of om inzicht te krijgen in de uitvoering van de interventie (bijvoorbeeld bij stakeholders).

Daarnaast worden diepte-interviews ook gehouden ter voorbereiding op enquêtes (zie methodiek 6). De verkregen inzichten worden dan gebruikt als input om vragenlijsten op te stellen, waarmee je motieven en gedragingen in de doelgroep kunt kwantificeren.

Type onderzoek: kwalitatief

Diepte-interviews behoren tot de kwalitatieve onderzoeksmethoden. Je kunt nagaan welke factoren spelen rond bepaald gedrag, maar geen cijfermatige uitspraken doen zoals 'x% vindt dat...'

Sterke en minder sterke punten van diepte-interviews

Sterke punten:

- Diepte-interviews zijn zeer geschikt om een onderwerp te verkennen, om complexere gedragsmechanismen te ontdekken en om verbanden te kunnen leggen tussen motieven en ervaringen, meningen en gedragingen.
- De interviewer is – anders dan in een enquête – flexibel in de vragen die hij stelt en waar hij over doorvraagt. Ook is het mogelijk om door inzet van projectieve en associatieve interviewtechnieken rationele en emotionele belemmeringen weg te nemen die kunnen spelen bij het beantwoorden van vragen.

Minder sterke punten:

- Net als bij enquêtes is werkelijk gedrag niet goed te achterhalen met diepte-interviews. Het blijft 'zelfgerapporteerd' gedrag:

gedrag dat mensen zéggén te vertonen. Vooral als het gaat om gewoontegedrag of om onbewuste processen of emoties, zijn interviews minder geschikt. Zet dan andere methodieken in, zoals de dagboekmethode of impliciet meten (zie methodieken 3 en 8). Ook kun je aanvullend inhoudsdeskundigen, zoals gedragsexperts op een specifiek beleidsterrein, interviewen om te reflecteren op de uitkomsten van het onderzoek onder de doelgroep.

- Je kunt geen representatieve uitspraken doen over de hele doelgroep. Wel krijg je met een goede selectie van de respondenten een goed beeld van het veld.

Praktijkvoorbeelden: energiebesparing en dienstreizen

- **Evaluatie van het energiebesparingsbeleid voor de gebouwde omgeving:** het Planbureau voor de Leefomgeving (PBL) zet geregeld diepte-interviews in voor beleidsonderzoek. Wat zijn bijvoorbeeld de motieven van bedrijven om energie te besparen? Met deze kennis zijn effectievere beleidsinterventies te ontwikkelen. Ook hield PBL diepte-interviews voor de evaluatie van het energiebesparingsbeleid. Zo ontstond inzicht in motieven en gedrag van verhuurders van onroerend goed. Corporaties bleken wel aan energiebesparing te willen doen, maar krijgen hun huurders onvoldoende mee. Ook bleek dat het energielabel het voor verhuurders aantrekkelijker heeft gemaakt om te investeren in energiebesparende maatregelen.
- **Verkenning van uitstelgedrag bij het boeken van dienstreizen:** rijksambtenaren moeten minimaal 21 dagen voor vertrek hun dienstreis boeken, via het TEM-systeem van 3W. Als ze dat later doen, zijn vliegtickets vaak duurder. Waarom stellen medewerkers het boeken vaak uit? De ministeries van EZK, IenW en Buitenlandse Zaken hebben in samenwerking met 3W onderzoek

gedaan naar psychologische factoren. Er zijn twaalf diepte-interviews gehouden om een eerste beeld te krijgen. Het blijkt onder meer dat medewerkers het zekere voor het onzekere nemen: ze willen graag alle details zeker weten voordat ze boeken, en zijn 'bang' om een vlucht te moeten wijzigen of annuleren. Er heerst onduidelijkheid over wijzigingsvoorwaarden bij vluchten. Ook lijkt de norm van 21 dagen niet algemeen bekend. Vervolgens zijn deze antwoorden opgenomen in een vragenlijst, waarmee kon worden gemeten of deze factoren ook voor een representatieve groep golden. Daaruit bleek bijvoorbeeld dat slechts 16% weet wat de norm is. Deze inzichten hielpen uiteindelijk om interventies te ontwerpen.

Meer informatie

Planbureau voor de Leefomgeving (PBL):
Kees Vringer, kees.vringer@pbl.nl.

Onderzoeksmethodiek 2:

Focusgroepen

De focusgroep of groepsdiscussie is een kwalitatieve onderzoeksvorm waarbij een groep mensen gevraagd wordt naar hun percepties, ideeën en meningen over een product, dienst, concept, advertentie of idee. Voor gedragsonderzoek kun je de focusgroep inzetten om bijvoorbeeld voor- en nadelen van (voorgenomen) interventies te bediscussiëren, zoals een gekozen benadering in publiekscommunicatie.

Methode: open discussie

In een focusgroep voeren mensen een gesprek over een onderwerp in een interactieve groepssetting. Elke deelnemer krijgt de kans om aan het woord te komen en om op andere deelnemers te reageren. Dit groepsgebesprek wordt geleid door een moderator, op een locatie of online. Aan een focusgroep nemen gewoonlijk zes tot tien personen deel. Dit hangt ook af van het aantal vragen en de lengte van het gesprek. Doorgaans duren focusgroepen zo'n twee uur per gesprek, en worden er per onderzoek twee tot vier gehouden. Zo kun je bijvoorbeeld groepen samenstellen van mensen uit diverse regio's, met verschillende opleidingsniveaus of in bepaalde levensfasen. Als het de bedoeling is om gedragsinterventies te bespreken, dan nemen uiteraard mensen deel waar die interventie zich op gaat richten.

Toepassing: verkenning van gedrag en interventies

Focusgroepen worden vaak ingezet in een verkennende fase. Ze richten zich op vragen als: hoe kijken mensen aan tegen bepaald gedrag? Waarom spelen sommige gedragsmotieven wel bij de één, maar niet bij een ander? Het is een goede methodiek als je geïnteresseerd bent in de discussie, interactie en argumentatie die tot bepaalde standpunten leiden. Als er in het dagelijks leven gesproken wordt over het gedragsvraagstuk, bijvoorbeeld wel of niet vaccineren, dan kan het waardevol zijn om zo'n gesprek ook te voeren in een onderzoekssetting. Verder kun je een focusgroep gebruiken om te bepalen welke gedragsinterventies aanspreken, welke zouden kunnen werken en waarom.

Type onderzoek: kwalitatief

Focusgroepen vallen onder het kwalitatief onderzoek. Je kunt er geen vragen mee beantwoorden als: hoeveel procent van de

Nederlanders doet/vindt et cetera? Wel krijg je goed inzicht in motieven, weerstanden en gedragskeuzes die mensen maken en in hoeverre deze te beïnvloeden zijn.

Sterke en minder sterke punten van focusgroepen

Sterke punten:

- Ten opzichte van een diepte-interview (zie methodiek 1) leidt een focusgroep vaak tot een natuurlijker gesprek. Er is discussie mogelijk; meningen en standpunten worden gelijk besproken in de groep. Het zijn vaak eyeopeners als je ziet hoe mensen op elkaars gedrag en motieven reageren, en argumenten voor of tegen interventies wikkelen en wegen. Ook krijg je zo gevoel voor de mate waarin gedrag te beïnvloeden is door voorgenomen interventies.
- De methode is geschikt om een onderwerp te verkennen, zonder dat er al veel over bekend is. Door er met verschillende mensen over te praten die op elkaar reageren, komen de verschillende perspectieven snel naar voren.

Minder sterke punten:

- Werkelijk gedrag is niet goed te achterhalen met groepsdiscussies. Bij de bespreking van gedragsmotieven en voorkeuren voor interventies geldt verder – net als bij diepte-interviews – als kanttekening: mensen zijn zich niet bewust van al hun afwegingen en willen ook niet alles delen.
- Als het gaat om gedrag waar mensen van nature weinig met elkaar over praten of om onderwerpen die gevoelig liggen, zoals het voorkomen van SOA's, dan is een focusgroep minder geschikt dan een individueel interview. Bovendien verzamel je via focusgroepen minder goed individuele reacties. Niet iedereen komt over elk gesprekspunt aan het woord en de deelnemers

kunnen elkaar beïnvloeden. Dit is te ondervangen door voor een bepaald gespreksonderdeel de deelnemers eerst individueel schriftelijk te laten reageren, en pas daarna te gaan discussiëren.

- Een focusgroep heeft het risico dat één of enkele deelnemers de groep gaan domineren. Een goede gespreksleider weet dit te voorkomen. Ook lijkt de toon die aan het begin van het gesprek gezet wordt, geregeld het verdere verloop te bepalen. Dit effect is beperkt als je meerdere focusgroepen houdt. Verder kunnen bijvoorbeeld laagopgeleiden stil worden van hoogopgeleiden, maar ook andersom. Om dat te voorkomen kun je ervoor kiezen twee aparte groepen te organiseren.

Praktijkvoorbeelden: beeld van Nederland en publiekscommunicatie

- **Continu Onderzoek Burgerperspectieven (COB):** voor de Voorlichtingsraad organiseert het Sociaal en Cultureel Planbureau ieder half jaar focusgroepen op twee plekken in het land; één met lageropgeleiden en één met hogeropgeleiden. Aan bod komen: hoe gaat het met Nederland? Waarom vindt u het de goede of verkeerde kant opgaan? En wat zou Den Haag daaraan kunnen doen? Verder staat er steeds een ander thema centraal, zoals zorg of onderwijs.
- **Ontwikkeling van publiekscommunicatie:** voor grootschalige publiekscampagnes van de overheid ontwikkelen reclamebureaus een creatief concept. Met dit creatieve idee willen ze de doelgroep aanzetten tot gewenst gedrag, bijvoorbeeld op het gebied van veiligheid of duurzaamheid. Focusgroepen worden dan ingezet om de ideeën van verschillende bureaus met de doelgroep te bespreken. Dit geeft inzicht in welke elementen potentie hebben voor gedragsverandering, wat stimuleert en wat weerstand oproept.

Meer informatie

Sociaal en Cultureel Planbureau (SCP):

Pepijn van Houwelingen, p.van.houwelingen@scp.nl.

Onderzoeksmethodiek 3:

Dagboekmethode

In deze methode krijgen respondenten de vraag om gedurende een bepaalde periode in een dagboek nauwkeurig te registreren wat ze doen, en soms ook met wie en waar. Op deze manier krijg je zicht op feitelijk gedrag. Een voorbeeld is tijdsbestedingsonderzoek dat laat zien hoeveel tijd mensen besteden aan media, werk, huishouden, vrije tijd et cetera.

Methode: geregistreerde gegevens over gedrag

De dagboekmethode wordt veel gebruikt voor budgetonderzoek (waarin supermarktbezoekers bijvoorbeeld hun aankopen noteren), maar is ook zeer gebruikelijk voor metingen van de dagelijkse tijdsbesteding. Dit laatste gebeurt in het tijdsbestedingsonderzoek dat het Sociaal en Cultureel Planbureau (SCP) al sinds 1975 uitvoert onder het Nederlands publiek. Respondenten noteren hun dagelijkse activiteiten in een dagboek per tijdseenheid van tien minuten. Hiermee wordt in beeld gebracht wat er gebeurt in het leven van burgers en welke ontwikkelingen daarin te zien zijn. Zijn Nederlanders bijvoorbeeld meer of minder tijd gaan besteden aan betaald werk of vrije tijd? De antwoorden op dergelijke vragen geven inzicht in collectieve ritmes en trends.

Naast het dagboek bevat tijdsbestedingsonderzoek vaak ook enquêtes om aanvullende vragen te stellen. Bijvoorbeeld over achtergrondkenmerken (zoals leeftijd, geslacht), zodat de verschillen in tijdsbesteding per groep in kaart gebracht kunnen worden. Daarnaast worden vaak vragen gesteld over persoonlijke ervaringen en attitudes. Bijvoorbeeld: zijn mensen tevreden over hun tijdsbesteding en hoe groot is de beleefde tijdsdruk?

Toepassing: in de fase van verkenning

De inzichten uit de dagboekmethode zijn vooral van nut in de verkennende beleidsfase. Wat gebeurt er, op welk moment en bij welke groepen? Hoe beweegt tijdsbesteding (direct of indirect) mee met maatschappelijke ontwikkelingen of beleidsmaatregelen? Ook kan de methode inspiratie geven voor mogelijke aangrijpingspunten voor beleid.

Type onderzoek: kwantitatieve dataverzameling

De dagboekmethode is hoofdzakelijk kwantitatief van aard. De activiteiten worden ingevuld via codelijsten (of achteraf gecodeerd) en de gegevens worden vervolgens geaggregeerd en geanalyseerd. Het is een vorm van dataverzameling. In klantreizen (zie methodiek 4) worden dagboeken (samen met andere informatiebronnen) ook wel kwalitatief ingezet om inzicht te krijgen in het feitelijke gedrag van mensen.

Sterke en minder sterke punten van de dagboekmethode

Sterke punten:

- Onderzoek heeft laten zien dat de dagboekmethode een betrouwbaar inzicht geeft in het gedrag van mensen. Als mensen worden gevraagd naar hun tijdsbesteding, schatten ze de tijd vaak verkeerd in. Sociaal wenselijke activiteiten worden overschat, en minder wenselijke onderschat. Ook vergeten ze activiteiten of laten ze activiteiten weg die gelijktijdig plaatsvinden. De dagboekmethode is eveneens gebaseerd op zelfrapportage, maar door de standaardisering van de meeteenheid (tien minuten) vergeten mensen minder snel om activiteiten in te vullen.
- Er ontstaat een gedetailleerd beeld van het gedrag. De dagboeken bieden niet alleen zicht op hoeveel tijd er aan verschillende activiteiten wordt besteed, maar ook op welke momenten en in welk ritme (aaneengesloten of onderbroken, en wel of niet in combinatie met andere activiteiten).

Minder sterke punten:

- De methode is relatief kostbaar en vraagt veel van respondenten. Responspercentages (hoeveel procent van de benaderde mensen doen ook daadwerkelijk mee aan het onderzoek) zijn in tijdsbe-

- stedingsonderzoeken over het algemeen dan ook niet erg hoog. Dit hangt ook af van hoeveel dagen mensen mee moeten doen.
- Er kunnen geheugenproblemen optreden als respondenten het dagboek niet regelmatig bijhouden. Om dat te ondervangen worden reminders gestuurd, bijvoorbeeld via apps. Ook zijn er experimenten met andere registratievormen (*reality mining*) zoals GPS-tracking voor locatiebepaling of het loggen van communicatieverkeer via smartphones. Die hebben weer hun eigen nadelen, zoals privacykwesties.
 - Er zijn activiteiten die met dagboeken niet goed geregistreerd worden. Bijvoorbeeld omdat ze in de onderzoeksperiode net niet voorkomen (denk aan museumbezoek). Andere activiteiten duren zo kort dat ze niet in de intervallen van tien minuten terechtkomen (bijvoorbeeld het checken van nieuws via websites). Deze informatie is beter te achterhalen in een aanvullende enquête of door middel van *push notifications* van een app.

Praktijkvoorbeelden: ontwikkelingen in algemene tijdsbesteding en mediagebruik

Op basis van gegevens uit het tijdsbestedingsonderzoek van 2006, 2011 en 2016 verkent het SCP in 'Alle ballen in de lucht' (2018) de relatie tussen de tijdsbesteding van mensen en hun kwaliteit van leven. Uit dit onderzoek blijkt dat een substantieel deel van de Nederlanders tijdsdruk ervaart, vooral ouders met thuiswonende kinderen, hogeropgeleiden en vrouwen. Het rapport laat zien dat tijdsdruk niet enkel een individueel, maar ook een sociaal probleem is. Daarmee onderstreept het de urgentie van het ontwikkelen van beleid dat niet méér tijd van mensen vraagt, maar mensen juist ontlast.

De dagboekmethode wordt ook gebruikt om specifiek het mediagebruik van Nederlanders in kaart te brengen. Op basis van het Media:Tijd-onderzoek zijn verschillende SCP-publicaties verschenen. 'Nederlanders en nieuws' analyseert het gebruik van nieuwsmedia via oude en nieuwe kanalen. Het onderzoek tempert de zorgen over schadelijke filterbubbels (algoritmes die een beperkte nieuwsselectie voorschotelen) enigszins. Wel wijst het op de noodzaak om in beleid aandacht te blijven besteden aan digitale vaardigheden. Het rapport 'Lees:Tijd – Lezen in Nederland' laat een zorgelijkere trend zien. De daling in het aandeel lezers, vooral onder de jongere helft van de bevolking, zet door. Omdat veel onderzoeken de positieve bijdrage van lezen en leesvaardigheid hebben aangetoond, vraagt deze daling om actie op het gebied van leesbevorderingsbeleid.

Meer informatie

Sociaal en Cultureel Planbureau (SCP):
Annemarie Wennekers, a.wennekers@scp.nl
Anne Roeters, a.roeters@scp.nl.

Onderzoeksmethodiek 4:

Klantreizen

Een klantreis (*customer journey*) beschrijft hoe mensen in verloop van tijd omgaan met een vraag, behoefte of probleem: welke route legt iemand af en hoe past een product of dienst daarin? De term klantreis komt oorspronkelijk uit de commerciële sector, maar het gedachtegoed wordt ook toegepast op overheidsdiensten. Zo kan een gemeente bijvoorbeeld de klantreis ‘bouwvergunning aanvragen’ onderzoeken.

Een klantreis omvat alles wat een gebruiker doet, voelt en ervaart vanaf het moment van de behoefte tot en met het daadwerkelijke gebruik van een dienst of product. Die reis breng je stap voor stap in kaart, vanuit het perspectief van de gebruiker, niet van de aanbieder. Ook kun je de contactmomenten (zogenoemde *touchpoints*) opnemen: op welke momenten in hun reis hebben ze contact met een dienst of product? Bijvoorbeeld wanneer bezoeken ze een website van de overheid?

Methodie: stappen in kaart gebracht vanuit doelgroep

In een klantreis wordt beschreven waar mensen tegenaan lopen, waar frustraties en positieve ervaringen zitten, welke contactmogelijkheden zij gebruiken, waar informatie ontbreekt of extra ondersteuning nodig is om het gewenste gedrag te vertonen. Dit gebeurt op basis van verschillende bronnen. Een klantreis wordt vaak visueel weergegeven als een tijdlijn met daarin alle stappen en de bijbehorende acties en contactmomenten.

Omdat de volledige klantreis erg lang en/of complex kan zijn, bijvoorbeeld het regelen van een scheiding, kun je ook deelreizen in kaart brengen, zoals het doorgeven van deze wijziging aan officiële instanties. Ook zijn er vaak meerdere typen klanten en dus meerdere klantreizen per product of dienst. Baken vooraf goed af waar je je op gaat richten.

De methodiek werkt het best als je de reis beschrijft vanuit een persona. Een persona is geen stereotype, maar een fictief persoon die staat voor een bepaalde doelgroep met specifieke behoeften, doelen, houding en gedragspatronen. Input voor zo'n persona komt uit onderzoek. Beschrijf hem of haar alsof het een levend

persoon is. Bijvoorbeeld: ‘Anja de hulpverlener, is een echte doener. Zij wil alleen weten wat ze moet doen en wanneer.’

Toepassing: in verkennende en ontwikkelende fase

Klantreizen gebruik je in de verkennende fase van beleid om inzichtelijk te maken waar gedragsproblemen zich voordoen en hoe dat komt. Daarnaast kun je klantreizen gebruiken om te bepalen waar kansen liggen voor interventies om het gewenste gedrag te stimuleren. Bijvoorbeeld door de eigen dienstverlening beter op behoeften van gebruikers te laten aansluiten, op het juiste moment de juiste ondersteuning te bieden en frustratie en pijnpunten te minimaliseren.

Type onderzoek: analyse van kwalitatieve en kwantitatieve gegevens

De input om klantreizen te maken is uit verschillende bronnen te halen; vaak een combinatie van bestaande gegevens, aangevuld met nieuw onderzoek. Voorbeelden van kwalitatieve data zijn gegevens uit etnografisch onderzoek (observaties van de gebruikers in hun eigen omgeving), diepte-interviews, groepsdiscussies of dagboekopdrachten. Voorbeelden van kwantitatieve data zijn gegevens uit enquêteonderzoek, bezoekcijfers van websites, data van een callcenter van de organisatie of openbare data zoals van het CBS.

Sterke en minder sterke punten van klantreizen

Sterke punten:

- Bij klantreizen wordt naar de gehele context gekeken, niet alleen naar je eigen product, dienst of communicatie.
- De behoeften, beleving en ervaringen van gebruikers staan centraal, waardoor je meer dan bij andere onderzoeksvormen wordt gedwongen te redeneren vanuit de gebruiker.

Minder sterke punten:

- Soms wordt een klantreis vooral vanuit de contactmomenten van de eigen organisatie bekeken. Hierdoor ontstaat een eenzijdig beeld en mis je de echte problemen en oplossingen voor gebruikers.
- Belangrijke onderdelen van de klantreis, zoals mensen observeren in hun omgeving of mensen een dagboek laten bijhouden, zijn tijdsintensieve methodes. Wel geven ze meer inzicht in het feitelijke gedrag dan als mensen achteraf een vragenlijst krijgen over hun gedrag.

Praktijkvoorbeelden: klantreizen IND en consulaten

- **Aanvraag verblijfsvergunning:** de Immigratie- en Naturalisatiedienst (IND) behandelt aanvragen voor een verblijfsvergunning van mensen die als kennismigrant in Nederland willen werken. De IND heeft via klantreizen meer inzicht verkregen in de vier fasen die de doelgroep onderscheidt tijdens deze aanvragen: verzamelen/indienen, bevestiging/beslissing, komst migrant en verlenging. Zo is er een groep die meer tussentijds inzicht wil in de status van de aanvraag maar dat niet krijgt. Dit leidt tot extra contact met de IND. Voor een andere groep is het te weinig transparant waarom de IND een termijn heeft van drie maanden voor de beslissing. De klantreizen zijn in kaart gebracht op basis van input uit twee onderzoeken. Ten eerste achttien individuele interviews met erkende referenten, bijvoorbeeld ondernemingen die aanvragen bij de IND indienen omdat ze personeel uit het buitenland willen aantrekken. Daarnaast zijn er acht interviews geweest met vreemdelingen die een verlenging hebben aangevraagd. Op basis daarvan kon de IND haar dienstverlening optimaliseren.
- **Consulaire dienstverlening:** voor Buitenlandse Zaken (BZ) zijn klantreizen gemaakt van mensen die contact hadden met het ministerie om bijvoorbeeld een paspoort of visum aan te vragen. Er zijn interviews gevoerd met klanten die onlangs zo'n consulaire dienst hadden afgenomen. Dat waren Nederlanders in Nederland, Nederlanders in het buitenland en niet-Nederlanders in het buitenland. Door de interviews werd inzichtelijk tegen welke knelpunten zij aanlopen. Daarnaast is een intermediaire partij geïnterviewd en zijn diverse onderzoeken en databestanden doorgenomen. Op basis van het onderzoek zijn persona's gemaakt. Dit was heel bruikbaar voor BZ en gaf ideeën om de dienstverlening te moderniseren, bijvoorbeeld door chat te gebruiken.

Meer informatie

Dienst Publiek en Communicatie (DPC):
Marit Groeneveld, m.groeneveld@minaz.nl
Margot Lagendijk, m.lagendijk@minaz.nl.

“Kennis moet niet uit je buik komen”

Jo van Kruchten

kennisadviseur bij Belastingdienst, over klantreizen

“Wat maakt een zzp’er in een jaar mee met de Belastingdienst? Dat hebben we in kaart gebracht in onze Customer Journey ZZP. We deden al veel onderzoeken naar onderdelen van onze dienstverlening: wat vind je van deze brief, hoe verliep dit jaar je aangifte, et cetera? Maar dit gaf niet een **overkoepelend beeld** van de ervaren knelpunten in de interactie met de Belastingdienst. Daarom hebben we voor een klantreis gekozen vanuit de zzp’ers: hun fiscale jaar in beeld.

Eerst keken we door onze **oogharen** naar alle marktonderzoeksrapporten die er al lagen, voor een eerste beeld van de zzp’er. Toen hebben we interne gesprekken gevoerd met iedereen die contact heeft met zzp’ers: van mensen die vragen beantwoorden tot mensen die bezwaren afhandelen. Daarna gingen we langs bij fiscaal adviseurs: waarom komen zzp’ers bij jullie? Waar worstelen ze mee? En tot slot hebben we 20-30 interviews en een aantal groepsgesprekken met zzp’ers zelf gehouden.

Daarmee hebben we de hele **stroom** in beeld kunnen brengen: administratie bijhouden, aangifte doen, aanslag krijgen, bezwaar maken, klacht indienen et cetera. Zoals Ikea doet van parkeren tot afrekenen. Daar kwam een lijst uit met 20-25 ‘lastigheden’. Zoals: de regelgeving is ingewikkeld en verandert zo vaak. Die lijst hebben we vervolgens kwantitatief getoetst: welke knelpunten komen vaak voor, en welke vinden ze het ergst? In de laatste fase zijn we met zzp’ers en interne mensen om tafel gegaan om mogelijke oplossingen te bedenken.

Deze klantreis heeft concreet bijvoorbeeld een zetje gegeven voor de automatisch ingevulde winstaangifte, een pilot die nu loopt. Daarin zijn administratie en aangifte aan elkaar gekoppeld. Zzp’ers willen namelijk wel aan hun verplichtingen voldoen, maar vinden dat lastig te organiseren en blijken er ook onzeker over. ‘Vul ik de aangifte niet verkeerd in zodat ik een boete krijg?’ ‘Heb ik misschien een aftrekpost gemist?’ Die onzekerheid maakt passief. Het **doenvermogen** zoals het heet, neemt af. Waardoor ze taken gaan uitstellen.

De meerwaarde van klantreizen? Dat je de knelpunten met elkaar in verband gaat zien. Dat inzicht over passiviteit hadden we nooit boven tafel gekregen met een formulier-test, of met ‘hebben we je goed te woord gestaan?’ Kennis is de basis voor elke vernieuwing, en die kennis moet niet uit je buik komen. We zijn in **de schoenen van de zzp’er** gaan staan, en hebben zo’n jaar met hen meebeleefd. Daardoor is meer begrip ontstaan.”

! Tip van Jo voor collega's

Stem van tevoren met de opdrachtgever (dienst/ beleidsafdeling) goed af wat de vervolgstappen zijn na het onderzoek. In welke gremia worden de resultaten besproken, wie gaan er actie op ondernemen? Betrek al vroeg veel mensen bij je onderzoek: communicatiedeskundigen, dienstverleners, juristen, procesontwerpers, IT'ers. Zo maak je ze 'medeplichtig' en zo'n breed draagvlak is nuttig voor het vervolg.

Onderzoeksmethodiek 5:

Ongeschreven regels

Ongeschreven regels komen in elke groep en iedere organisatie voor. Denk alleen al aan hoe er koffie wordt gehaald: loopt één persoon voor iedereen of loop je samen? Maar het kan ook gaan over hoe er besluiten worden genomen. Het zijn sociale normen voor gedrag. De ongeschreven regels zijn wat je uitlegt aan een nieuwe medewerker: 'zo doen we dat hier'. Ze zijn niet goed of slecht, maar hebben wel allerlei (wenselijke of onwenselijke) effecten. Als je de effecten in beeld hebt en de onderliggende regels doorziet, is het voor een organisatie makkelijker om te bepalen aan welke knoppen gedraaid kan worden om de juiste kant op te veranderen.

Met Ongeschreven Regels-onderzoek kun je gedrag bespreekbaar maken en grip krijgen op 'zachte' cultuuraspecten. Het maakt het mogelijk om allerlei soorten weerstand bij verandering te voorspellen en het geeft versneld inzicht in wie en wat er toe doet in de organisatie en wat de sociale dynamiek is rondom een thema. De methodiek maakt gebruik van diepte-interviews (zie methodiek 1) in verschillende rondes.

Methodie: in interviewrondes regels herleiden

De methode om ongeschreven regels naar boven te halen is ontwikkeld door Peter Scott Morgan (The Unwritten Rules of the Game, 1994). Het onderzoek kan zich richten op de hele organisatie of op een vraagstuk of thema binnen de organisatie (zoals samenwerken, vernieuwen, dereguleren). De drie basiselementen van ongeschreven regels zijn drijfveren, machtgevers en hefboomen. Met diepte-interviews haal je die naar boven. Dat gaat op basis van drie vragen:

- Wat vind je écht belangrijk (drijfveren)?
- Wie zijn hier écht belangrijk (machtgevers)?
- Hoe krijg je wat je wilt hebben (hefboomen)?

Belangrijke interviewvaardigheden daarbij zijn: doorvragen tot het schuurt, vragen om concrete gedragsvoorbeelden en durven stiltes te laten vallen. De interviews worden uitgeschreven voor een kwantitatieve telling en een kwalitatieve analyse van de belangrijkste quotes. Op basis daarvan worden conceptregels opgesteld. In een tweede interviewronde worden de regels voorgelegd aan nieuwe respondenten. Welke regels herkennen ze, kunnen ze oorzaken en verbanden noemen? Welke effecten van die regels zien ze in de organisatie? En hebben ze voorbeelden van gedrag waaruit dat effect blijkt?

Na een tweede analyseronde blijven er gevalideerde ongeschreven regels over. Per regel worden de effecten inzichtelijk gemaakt. Daarnaast kan in een workshop met de groep worden uitgediept wat de samenhang is tussen de regels: welke oorzaak-gevolgrelaties er zijn, en hoe ze elkaar versterken.

Toepassing: in fase van verkenning en beleidsontwikkeling

Dit type gedragsonderzoek richt zich voornamelijk op gedrag binnen organisaties en minder op gedragsvraagstukken op een bepaald beleidsterrein. Het biedt naast inzicht in gedragsbepalers ook houvast en handelingsperspectief voor verandering. Het maakt het mogelijk om passende interventies te ontwikkelen en te onderbouwen, en helpt om weerstand te voorspellen. Ook kunnen uit het onderzoek de randvoorwaarden voor succes naar voren komen.

Type onderzoek: kwalitatief

Ongeschreven Regels-onderzoek is uitgebreid kwalitatief onderzoek. Basis zijn diepte-interviews; zo'n twintig stuks afhankelijk van het vraagstuk en de homogeniteit van de respondentengroep. Deze interviews duren gemiddeld anderhalf uur. Een team van onderzoekers voert het onderzoek uit en analyseert met elkaar de resultaten. De opdrachtgever kan bij sommige onderdelen betrokken worden.

Sterke en minder sterke punten van Ongeschreven Regels-onderzoek

Sterke punten:

- Als organisatieproblemen hardnekkig zijn en al verschillende opties zijn geprobeerd, kan Ongeschreven Regels-onderzoek uitkomst bieden. Het houdt rekening met 'de onderstroom' en

geeft handvatten om die bespreekbaar en ‘van iedereen’ te maken.

- Er ontstaat een gedragen beeld omdat medewerkers worden betrokken, zowel als onderwerp van onderzoek als bij de duiding van de onderzoeksresultaten.
- De onderzoeksresultaten zijn niet alleen aanleiding voor verandering, maar dit type onderzoek is vaak al een interventie op zich.

Minder sterke punten:

- De methode kan relatief intensief en kostbaar zijn als je meer dan twintig diepte-interviews wilt houden en als alle gesprekken letterlijk worden uitgeschreven in plaats van alleen de hoofdpunten. Er zijn bedrijven die transcripties maken van opnames. Als alternatief kun je kiezen voor groepsessies, maar dan is de kans op ‘groepsdenken’ groter en zijn de uitkomsten minder herleidbaar.
- Objectiviteit is onmogelijk. Wel kun je streven naar intersubjectiviteit: als de leden van het onderzoeksteam gezamenlijk betekenis geven aan de resultaten, benader je de objectiviteit het best.

Praktijkvoorbeelden: JenV verandert en interactief beleid

- **Ministerie van Justitie en Veiligheid (JenV):** de Auditdienst Rijk (ADR) heeft de afgelopen jaren onderzoeken naar ongeschreven regels uitgevoerd bij vijf departementen. Eén daarvan was het ministerie van JenV. Onderzocht zijn de ongeschreven regels die spelen rond samenwerking. Bijvoorbeeld: medewerkers komen pas in actie nadat ze hebben vastgesteld of ze erover gaan. Het programma ‘JenV verandert’ heeft de uitkomsten van het onderzoek verwerkt in de veranderaanpak. Een voorgestelde

interventie was bijvoorbeeld om de top te stimuleren de organisatie te laten zien dat ze samenwerkt.

- **Rijksoverheid:** het boek ‘Omgaan met Ongeschreven Regels’ van Max Herold bevat veel voorbeelden over ongeschreven regels rondom interactief beleid maken bij de Rijksoverheid. Herold laat zien hoe harde ongeschreven regels bij de overheid openheid in beleid belemmeren. Hij komt vervolgens met ‘copingstrategieën voor ambtenaren’.

Meer informatie

Auditdienst Rijk (ADR):

Marlies Ypma, m.e.ypma@minfin.nl

Roos Kalker, r.kalker@minfin.nl.

Onderzoeksmethodiek 6:

Enquêteonderzoek

Een enquête, ook wel *survey* of vragenlijstonderzoek genoemd, is de grootschalige afname van een vragenlijst bij een doelgroep. Dat kan op papier, mondeling of digitaal. De methodiek is geschikt als je uitspraken wilt doen over het (zelfgerapporteerd) gedrag van een bepaalde groep mensen of de werking van een interventie. Een representatieve steekproef uit de doelgroep vult de vragenlijst in. De verkregen data worden vervolgens met statistische methoden geanalyseerd.

Methodie: vragenlijstonderzoek om onder meer opinies te kwantificeren

Met enquêtes zijn ervaringen en meningen van een groep mensen te kwantificeren. Je kunt uitspraken doen als 'x% van de ondervraagden vindt...'. Ook naar gedrag kun je vragen, maar dan gaat het over 'zelfgerapporteerd gedrag' (gedrag dat mensen zeggen te vertonen) of gedragsintentie (gedrag dat mensen van plan zijn om te doen). Deze antwoorden wijken vaak af van het werkelijke gedrag, vooral bij gewoontegedrag of als emoties en onbewuste processen een rol spelen.

Het is belangrijk om een goede steekproef te trekken uit de groep die je wilt onderzoeken. De trekking moet zorgvuldig gebeuren (bijvoorbeeld aselect of gestratificeerd) en de uiteindelijke respons moet voldoende groot zijn om representatieve uitspraken te kunnen doen over de doelgroep als geheel (populatie). Als je binnen de doelgroep ook uitspraken wilt doen over subgroepen, bijvoorbeeld verschillende leeftijdsgroepen, dan is een grotere steekproef nodig. Met statistische toetsen zijn groepen mensen of metingen te vergelijken.

Toepassing: van verkenning en ontwerp tot evaluatie

Enquêtes zijn geschikt om in de verkennende fase inzichten over gedrag en achterliggende motieven van de doelgroep(en) te kwantificeren. Verder is ermee te bepalen wat de gedragsmechanismen zijn bij een voorgenomen interventie (ex-ante-evaluatie). Ook kun je de methode gebruiken voor ex-postevaluatie door groepen mensen te vragen naar hun gedrag, mening of ervaring. Enquêtes worden vaak periodiek herhaald, bijvoorbeeld om de ontwikkeling van gedrag op langere termijn te volgen. Dan heet de methodiek een monitor.

Type onderzoek: kwantitatief

Enquêtes behoren tot de kwantitatieve methoden van dataverzameling. Het is mogelijk om zowel gesloten als open vragen te stellen. In vergelijking met diepte-interviews (zie methodiek 1) hebben de antwoorden op open vragen in een vragenlijst meestal minder diepgang. Wel zijn deze 'open antwoorden' weer kwantitatief te analyseren door ze te laten coderen.

Sterke en minder sterke punten van enquêtes

Sterke punten:

- Enquêtes zijn zeer geschikt om representatieve uitspraken te doen over het (zelfgerapporteerd) gedrag, de mening en ervaringen van de doelgroep of de werking van een interventie.
- Het onderzoek is vaak snel en relatief goedkoop uitvoerbaar, vooral via online panels. Echter, vaak zijn niet alle groepen in de samenleving even goed vertegenwoordigd in deze panels (zoals laaggeletterden) en dan moet de steekproef op een andere manier getrokken worden. Het onderzoek kan dan meer tijd en geld kosten, afhankelijk van hoe hoog de eisen zijn aan de representativiteit.

Minder sterke punten:

- Enquêtes zijn meestal niet geschikt om werkelijk gedrag te achterhalen. Als er geen andere bronnen beschikbaar zijn, kun je een enquête wel gebruiken om een beeld te krijgen van het gedrag. Bijvoorbeeld met een vraag als: bent u op vakantie geweest? Door zelfrapportage kan er altijd een vertekening in de data optreden. Dit kan veroorzaakt worden doordat mensen zich het feitelijke gedrag niet goed meer kunnen herinneren of doordat het om onbewust gedrag gaat. Ook kunnen ze er, bewust of onbewust, minder eerlijk over zijn omdat bepaald gedrag niet

sociaal wenselijk is. De combinatie met een RCT (zie methodiek 13) geeft dan een goed beeld van werkelijke gedragingen.

- Enquêtes worden soms te gemakkelijk ingezet, zonder gedegen kennis over de manier van vragen stellen, over de volgorde van de vragen en de invloed van de vragen onderling. Dan zijn de resultaten minder valide. Om goed te kunnen interpreteren wat de gegeven antwoorden betekenen, mogen ondervraagden niet gestuurd worden in hun antwoorden. Voor een zo zuiver mogelijk beeld van de gedragsintentie is het bijvoorbeeld verstandig om daar eerst naar te vragen voordat het gaat over achterliggende factoren zoals de houding ten opzichte van het gedrag.

Praktijkvoorbeelden: enquêtes vanuit het PBL

- **Evaluatie van het energiebesparingsbeleid voor de gebouwde omgeving:** hiervoor heeft het Planbureau voor de Leefomgeving (PBL) een survey uitgevoerd onder huurders en woningeigenaren. Doel was te beschrijven hoe de doelgroepen beslissen over wel of niet investeren in energiebesparende maatregelen en wat de rol van het beleid daarin is. Hieruit bleek dat veel woningeigenaren meestal maar één energiebesparende maatregel tegelijk nemen. Het ministerie van Binnenlandse Zaken (BZK) heeft dit resultaat gebruikt om zijn subsidieregelingen vorm te geven. In de nieuwe regeling konden woningeigenaren alleen subsidie krijgen als zij twee of meer maatregelen tegelijkertijd nemen. Daarmee hoopt BZK woningeigenaren te verleiden meer te doen.
- **Duurzaamheidsverkenning:** wat is de milieudruk van private consumptie? Hieronder vallen alle producten en diensten die een huishouden aanschaft zoals eten, kleding, woning en vliegreizen. Om die milieudruk te achterhalen zijn drie enquêtes afgenomen onder eenzelfde groep consumenten. De eerste had als doel een

beeld te krijgen van het totale consumptiepatroon, de tweede om de waardepatronen te bepalen en een derde om de probleemperceptie van de consumenten te bepalen. Zo kwam veel informatie op tafel over consumenten zonder dat de vragenlijsten elkaar konden beïnvloeden. Uitkomst: waardepatronen en probleemperceptie bleken de milieudruk van de totale consumptie niet te beïnvloeden.

Meer informatie

Planbureau voor de Leefomgeving (PBL):
Kees Vringer, kees.vringer@pbl.nl.

Onderzoeksmethodiek 7:

Keuze-experimenten

Keuze-experimenten bieden inzicht in de voorkeuren van mensen en het mogelijke gedrag dat ze kunnen vertonen in alledaagse situaties. In deze methodiek laat je mensen kiezen uit een groot aantal opties die van elkaar verschillen op een vaste set van kenmerken. Op die manier is het relatieve belang van kenmerken vast te stellen: welke bepalen hun voorkeuren en gedragingen?

Methodie: relatieve belang vaststellen

In keuze-experimenten krijgen respondenten opties voorgelegd in willekeurige combinaties en volgordes. Ze moeten bijvoorbeeld kiezen tussen (verzonnen) sollicitanten met verschillende kenmerken, of tussen twee hypothetische pensioenstelsels. Deze vorm van keuze-experimenten heet een vignettenstudie. Uit de keuzes die mensen maken, is af te leiden wat hun voorkeur heeft en welke aspecten die voorkeur beïnvloeden. Dit verschilt van bijvoorbeeld een enquêteonderzoek (zie methodiek 6) waarin rechtstreeks wordt gevraagd naar voorkeuren, gedrag en het belang van aspecten.

Toepassing: in fase van verkenning

De methode is geschikt als je het relatieve gewicht wilt verkennen van factoren bij het keuzegedrag van burgers (bijvoorbeeld: hoe belangrijk is een duurzaamheidslabel als ze producten kiezen in de supermarkt?) of voor het draagvlak onder verschillende beleidsopties (bijvoorbeeld: welk pensioenstelsel heeft de voorkeur?). Voor een keuze-experiment is enige voorkennis nodig om een goede set van kenmerken en alternatieven te kunnen samenstellen.

Type onderzoek: kwantitatief

Keuzes kunnen het best in een vragenlijst voorgelegd worden aan een willekeurige steekproef van respondenten. Dit kan online, of door met een laptop op huisbezoek te gaan als de doelgroep anders minder goed te bereiken is. De methode van dataverzameling en analyse is kwantitatief van aard.

Sterke en minder sterke punten van keuze-experimenten

Sterke punten:

- Met keuze-experimenten kun je grootschalig toetsen wat het relatieve belang is van aspecten op iemands voorkeur of gedrag.

- Een keuze-experiment benadert goed het werkelijke gedrag van mensen waarin ze telkens keuzes moeten maken, van de dagelijkse boodschappen tot het afsluiten van een hypotheek.
- Keuze-experimenten dwingen respondenten om afwegingen te maken tussen de voorgelegde alternatieven met hun verschillende kenmerken. Als je mensen vraagt naar het belang, kunnen ze alle kenmerken belangrijk vinden. Nu moeten ze ze tegen elkaar afwegen.

Minder sterke punten:

- Keuze-experimenten zijn belastend voor respondenten. In recente studies stopten relatief veel mensen met de vragenlijst omdat ze het zwaar vonden. Ze vulden vaak 'weet ik niet' in in plaats van dat ze een keuze maakten. De consistentie van de resultaten nam na vijf keuzes duidelijk af. Bepaalde respondenten (in termen van opleidingsniveau, leeftijd, migratieachtergrond) haakten vaker of sneller af, waardoor er ook nog een vertekening optrad in de data. Houd daarom het aantal keuzes per respondent beperkt tot maximaal vijf, en het aantal kenmerken van een keuze beperkt tot maximaal zes.
- De (individuele) keuzes die mensen maken in een keuze-experiment, komen niet altijd overeen met het gedrag dat je in de maatschappij terugziet. Zo kunnen mensen in een keuze-experiment aangeven geen voorkeur te hebben voor verplicht pensioen sparen, terwijl die variant het in de praktijk goed doet. Dergelijke afwijkingen kunnen het gevolg zijn van andere factoren die een rol spelen, bijvoorbeeld doordat mensen als het erop aan komt ontzorgd willen worden.

Praktijkvoorbeelden: voorkeuren voor sollicitanten en pensioenstelsels

Het Sociaal en Cultureel Planbureau (SCP) heeft verschillende keuze-experimenten uitgevoerd:

- **Keuze tussen sollicitanten:** in deze vignettenstudie heeft het SCP respondenten gevraagd de rol aan te nemen van eigenaar van een (kantoor)boekhandel op zoek naar een nieuwe verkoopmedewerker. Ze moesten kiezen wie van de voorgelegde sollicitanten zij zouden uitnodigen voor een tweede gesprek. De respondenten kregen vijf vignetten te zien met een portretfoto van de kandidaat en een aantal kenmerken: naam, leeftijd, etniciteit, veel/weinig zelfvertrouwen, wel/geen lichamelijke beperkingen, opleidingsniveau, aantal jaren relevante werkervaring en beroep. De kenmerken werden grotendeels random toegekend aan de vignetten. De vignetten werden vervolgens ook random toegekend aan de respondenten. Zo was bijvoorbeeld de relatieve invloed van werkervaring vast te stellen, of van gezondheid; en de relatie tussen het uiterlijk van de sollicitant en de bereidheid van mensen om hem of haar aan te nemen. Zie: Van Campen & Versantvoort, 2014.
- **Keuze tussen hypothetische pensioenstelsels:** deze vignettenstudie heeft zicht gegeven op pensioenvoorkeuren van zelfstandigen en werknemers. Ze moesten kiezen tussen pensioenstelsels met verschillende kenmerken, zoals het vroeger met pensioen kunnen in ruil voor een lagere uitkering, of het flexibel kunnen opnemen en/of inleggen. Zo was vast te stellen hoe mensen bepaalde kenmerken wegen. Zie: Olsthoorn, 2019.

Meer informatie

Sociaal en Cultureel Planbureau (SCP):
Martin Olsthoorn, m.olsthoorn@scp.nl
Crétien van Campen, c.van.campen@scp.nl.

Onderzoeksmethodiek 8:

Impliciet meten via reactietijdentaken

Veel gedrag voeren we onbewust uit of op de automatische piloot. Denk aan eten of autorijden. Als je mensen er rechtstreeks naar vraagt, zoals bij diepte-interviews of enquêtes, krijg je onvoldoende inzicht in deze onbewuste processen. Dit expliciete meten ('waarom deed u dit?') werkt dan niet. Simpelweg omdat we niet in staat zijn om deze vragen te beantwoorden. Of we gaan er achteraf argumenten voor verzinnen of sociaal wenselijk antwoorden. Impliciet meten biedt dan uitkomst. Dit type onderzoek voorspelt feitelijk gedrag beter dan onderzoek naar zelfgerapporteerd gedrag.

Met impliciete testen is het mogelijk om inzicht te verkrijgen in de onbewuste processen die van invloed zijn op houding en gedrag. Hierbij hoeft een respondent niet bewust na te denken over zijn gedragsmotieven en het effect dat een interventie op hem of haar heeft. Eén van die impliciete testen is reactietijdentaken.

Methode: onbewuste gedragsprocessen achterhalen

Bij reactietijdentaken wordt de reactietijd gemeten van mensen die een bepaalde taak uitvoeren. Van de snelheid waarmee ze dat doen, is af te leiden wat het effect is van de interventie. Het doel van de test wordt daarbij verhuld.

Er zijn verschillende soorten reactietijdentaken. Daarvan is de impliciete associatietest een goed gevalideerde. In zo'n test krijgen mensen een reeks van stimuli aangeboden, zoals verschillende handelingen, in een willekeurige volgorde. Ze moeten die categoriseren als positief of negatief. Wanneer respondenten een stimulus, bijvoorbeeld het gewenste gedrag, sneller onder 'positief' scharen dan onder 'negatief', betekent dit dat ze onbewust positief tegenover dit gedrag staan. Zo is de houding ten opzichte van verschillende stimuli, zoals meerdere handelingsperspectieven, met elkaar te vergelijken. Ook kun je het effect vaststellen van blootstelling aan een interventie, bijvoorbeeld communicatie over het gewenste gedrag via een tv-spotje. Daarna volgen eventueel alsnog expliciete vragen over houding en gedrag om te kijken van welke effecten ze zich bewust zijn.

Toepassing: voor verkennen en evalueren van beleid en communicatie

Met reactietijdentaken kom je meer te weten over de huidige onbewuste houding van de doelgroep ten opzichte van (on)gewenst gedrag of de mate waarin ze geneigd zijn bepaald gedrag uit te voeren. Denk aan het melden van een vermoeden van huiselijk geweld. Het draagt zo in een verkennende fase bij aan meer inzicht in het gedragsvraagstuk. Daarnaast kun je ermee vaststellen in hoeverre een interventie (bijvoorbeeld een campagne-uiting) leidt tot gedragsverandering of de onbewuste houding beïnvloedt. Dit helpt bij het evalueren van beleid en communicatie.

Soort onderzoek: kwantitatief

Impliciet meten via reactietijden is een vorm van (online) kwantitatief onderzoek. Net als bij een enquête is het belangrijk om een representatieve steekproef te trekken uit de doelgroep. De grootte van de steekproef hangt mede af van het aantal stimuli – bijvoorbeeld gedragsalternatieven – dat je wilt voorleggen. Een impliciete test duurt bij voorkeur niet langer dan zes à zeven minuten om respondenten gemotiveerd te houden. Zo blijft er ook nog tijd over om expliciete vragen te stellen en de totale vragenlijst niet langer te maken dan twaalf minuten – een gangbaar maximum. Stel de expliciete vragen na de impliciete test, om beïnvloeding van de test te voorkomen. Wel kan vooraf gevraagd worden naar bijvoorbeeld eerdere ervaringen met het onderwerp, om het onderwerp in het geheugen te activeren. De formuleringen mogen dan niet sturend zijn.

Sterke en minder sterke punten van impliciet meten via reactietijdentaken

Sterke punten:

- Impliciete maten blijken vaak een betere voorspeller te zijn van feitelijk gedrag dan expliciete maten. Voor de impliciete associatietest is dit aangetoond op het gebied van onder meer discriminatie, gezondheidscommunicatie en stemgedrag. Oftewel: de uitkomst van zo'n test komt beter overeen met het feitelijke stemgedrag dan met wat mensen zégggen te gaan doen.
- De methode sluit goed aan bij hoe mensen zich in de dagelijkse praktijk gedragen en informatie verwerken, namelijk met beperkte aandacht, zonder veel bewuste overwegingen of diepgaande reflectie.
- In tegenstelling tot expliciete vragen is impliciet meten niet afhankelijk van de correcte herinnering van mensen of de gekozen formulering van vragen. Ook is het niet gevoelig voor versturende processen, zoals sociaal wenselijk antwoorden of het achteraf rationaliseren van gedrag.

Minder sterke punten:

- De samenhang tussen reactietijdentaken en gedrag wordt soms verondersteld zonder verdere (al dan niet wetenschappelijke) bewijsvoering.
- Reactietijdentaken worden afgenomen in een onderzoekssetting of thuis achter de computer. In de praktijk kunnen externe factoren meespelen. De effecten van een interventie op onbewuste houding of gedragsneiging die in zulke testen gevonden

worden, kunnen daardoor in werkelijkheid kleiner zijn of niet aantoonbaar.

Praktijkvoorbeelden: pretesten en campagne-effectonderzoek

Dienst Publiek en Communicatie zet reactietijdentaken in om te meten welk effect overheidscommunicatie heeft op houding en gedrag. Dit gebeurt in de evaluatiefase bij het pretesten van communicatie-uitingen en via campagne-effectonderzoek. De methodiek is ingezet voor diverse campagnes, zoals Preventie Woninginbraak (ministerie van Justitie en Veiligheid), Geweld in Huiselijke Kring (ministerie van Volksgezondheid, Welzijn en Sport) en Overal Dichtbij (ministerie van Buitenlandse Zaken). Impliciet meten heeft hier inzicht gegeven in de werking en effecten van campagnes. Zo bleek dat er na het zien van de campagne Geweld in Huiselijke Kring een duidelijk positief effect was op de ónbewuste houding om in actie te komen (meld een vermoeden), ook al verandert er niets in de bewuste houding van mensen. Sinds april 2019 is de onderzoeksmodule reactietijdentaken toegevoegd als optie binnen het campagne-effectonderzoek.

Meer informatie

Dienst Publiek en Communicatie (DPC):
Joost Loef, j.loef@minaz.nl.

Onderzoeksmethodiek 9:

Big data analytics

Big data analytics is gelijk opgekomen met de dataficering van de samenleving. De hoeveelheid gearchiveerde digitale gegevens blijft jaarlijks exponentieel groeien. Denk aan gegevens die mensen vastleggen: online conversaties, foto's, video's en documenten. Maar ook aan gegevens die door sensoren en apparaten worden geregistreerd: transacties, processen en statussen. Analyse van die data kan gedragspatronen blootleggen.

Big data-analyse krijgt steeds meer aandacht binnen de overheid. Met deze techniek kan de besluitvorming goedkoper en sneller worden, en ook beter: gebaseerd op feitelijke en actuele informatie, zoals over gedrag. De overheid past dit toe voor onder meer beleidsondersteuning, procesoptimalisatie, fraudesignalering en handhaving. Een vraag die je er bijvoorbeeld mee kunt beantwoorden is: hoe gaat de armoede in een gemeente zich ontwikkelen?

Methode: patronen vinden in gegevens en daarmee (gedrag) voorspellen

Big data-analyse maakt gebruik van statistische methodes om relaties bloot te leggen en modellen te bouwen om trends te kunnen voorspellen. Econometristen werken daar al veel langer mee. Alleen analyseren die nauwkeurig geselecteerde, geprepareerde, opgeschoonde en afgekaderde datasets. Terwijl big data-analyse werkt met de enorme hoeveelheden ongeschoonde, veelsoortige, *real time* gegevens die tegenwoordig beschikbaar zijn. Een ander verschil is dat de gebruikte modellen en algoritmen bij big data-analyses vaak een zelflerend karakter hebben. Ze zijn in staat om relaties en patronen in de data in beeld te brengen die door de complexiteit anders voor mensen onzichtbaar zouden blijven. Denk aan alle op elkaar inwerkende factoren die de verkeersdrukte op de snelwegen beïnvloeden.

Toepassing: beleidsvoorbereiding en evaluatie

Big data analytics kun je in de fase van beleidsvoorbereiding inzetten om het werkelijke gedrag van mensen of bedrijven in kaart te brengen door uiteenlopende datasets en registers te combineren. Dit kan zowel op individueel als geaggregeerd niveau. Daarnaast kan ook een verandering van gedrag na toegepaste interventies inzichtelijk worden gemaakt.

Type onderzoek: kwantitatieve analyse

Big data-analyse is een vorm van kwantitatief onderzoek waarbij verschillende soorten gegevens worden geanalyseerd. Het Centraal Bureau voor de Statistiek (CBS) heeft diverse registers en onderzoeksdatabases zoals de actuele demografiecijfers van Nederland. Het Center for Big Data Statistics van het CBS combineert die met andere beschikbare databronnen, bijvoorbeeld 3G/4G-telefoonmastdata van telecomproviders. Dit geeft input voor veel beleidsterreinen: woningmarkt, energietransitie, mobiliteit, cybercrime et cetera.

Sterke en minder sterke punten van big data analytics

Sterke punten:

- Analyses worden uitgevoerd op feitelijke data over gedrag en omgevingsfactoren, dus bijvoorbeeld niet op wat mensen zégen te doen.
- De gebruikte algoritmen kunnen onverwachte, complexe verbanden aantonen in gegevens.
- Big data-analyse is een relatief goedkope methode ten opzichte van arbeidsintensief enquêteonderzoek. Je kunt veel doen met al eerder verzamelde gegevens.
- De methode biedt de mogelijkheid om effecten van interventies of optimalisaties te monitoren over een langere periode.

Minder sterke punten:

- Big data-analyses vinden correlaties (samenhang) en geven geen inzicht in causaliteit (oorzakelijk verband). Bepaald gedrag zou bijvoorbeeld meer voor kunnen komen bij mensen met kinderen, maar in het midden blijft dan of dat komt dóór de aanwezigheid van de kinderen.

- De modellen kunnen door hun complexiteit een black box vormen. Als in bepaalde wijken de kans op woninginbraken groter blijkt, is niet precies bekend op welk onderdeel uit de data dat gebaseerd is. Daardoor is de besluitvorming die hieruit volgt, niet transparant voor burgers.
- De gebruikte algoritmen kunnen onbedoeld ongewenste effecten hebben, zoals etnisch profileren bij fraudeopsporing. Selectiviteit of *bias* in de datasets kunnen uitsluiting of discriminatie tot gevolg hebben. De universiteit van Utrecht ontwikkelde hiervoor samen met lokale overheden de Data Ethics Decision Aid (DEDA). Deze methode helpt projecteigenaren om bij voorgenomen big data-projecten ethische afwegingen te maken over onder meer impact op individuen en transparantie.

Praktijkvoorbeelden: verkeersverwachting, fraudeopsporing en politieplanning

- **Procesoptimalisatie Rijkswaterstaat:** Rijkswaterstaat gebruikt modellen die verwachtingen van verkeersdrukte leveren. Die verwachtingen zijn gebaseerd op onder meer historische verkeersmetingen, KNMI-gegevens over weersomstandigheden, trends in economische bedrijvigheid, vakantiegegevens en gegevens over griepgolven in Nederland. Op basis van deze modellen worden bijvoorbeeld de wegwerkzaamheden zo ingepland dat er zo min mogelijk files door ontstaan.
- **Fraudesignalering Belastingdienst:** op welke groepen belastingplichtigen moet de nadruk komen te liggen voor grondige controles van de inkomstenbelasting en de btw-aangifte? Risicomodellen geven de Belastingdienst daar inzicht in. Complexe, *machine learning*-technieken geven signalen af over bij wie mogelijk nader onderzoek door inspecteurs nodig is. Zo wordt ook de controlecapaciteit efficiënter ingezet.

- **Handhaving gemeente Amsterdam:** de gemeente Amsterdam gebruikt het Criminaliteits- Anticipatie Systeem (CAS) om misdaad te voorspellen zoals woninginbraken, overvallen en straatroven. Zo kan de politie gericht worden ingezet in gebieden waar een verhoogde kans is op deze zogeheten *high impact crimes*. De algoritmen van CAS gebruiken hiervoor onder meer historische criminaliteitsdata in combinatie met andere data zoals inkomensgegevens en het aantal slijterijen in een wijk.

Meer informatie

Centraal Bureau voor de Statistiek (CBS):

Joost Hurman, jwf.hurman@cbs.nl.

Dienst Publiek en Communicatie (DPC):

Martijn Warmoeskerken, m.warmoeskerken@minaz.nl.

“In databestanden kun je verdrinken”

Joost Hurman

managing director CBDS bij het CBS, over big data analytics

“Big data is een **containerbegrip**. Er vallen grote registers onder, zoals btw-gegevens van de Belastingdienst, maar ook complex te interpreteren data van sociale media. Die databronnen kun je analyseren en soms ook aan elkaar ‘knopen’ om verbanden te leggen. Dat doen we bij het Center for Big Data Statistics (CBDS), voor onze externe partners en voor onze eigen statistieken.

Voor veel statistieken verzamelen we de gegevens met enquêtes. Tijdrovend voor ons en belastend voor de ondervraagden. Het voordeel van een enquête is wel dat je precies kunt vragen wat je wilt weten. Maar je krijgt maar weinig gegevens. Bij big data is dat precies andersom: je hebt veel meer gegevens, maar niet altijd precies wat je wilt weten. Het vraagt veel studie om daar goede statistieken van te maken.

Met die big data kun je **terugkijken**: hoe heeft iets zich ontwikkeld in de afgelopen jaren? Denk aan de uitstoot van emissies door de Nederlandse zeevaart. Beleidsmakers willen dat weten voor de effectiviteit van hun maatregelen. Wij kunnen daarbij overigens alleen verbanden aangeven; we kunnen over het algemeen niet zeggen of die verklaard worden door het beleid.

Behalve terugkijken kun je op basis van die data ook **voorkijken**. Wij noemen dat geen voorspellingen – het CBS voorspelt niet. We maken in sommige gevallen wel kansberekeningen. Wat is bijvoorbeeld de kans dat mensen gaan verhuizen? We pakken dan onder meer de Polisadministratie (een register van Nederlandse inkomstgegevens) en gegevens van het Kadaster en laten daar een *machine learning*-algoritme op los. Dat geeft inzicht in verhuiskansen, en dat is bijvoorbeeld handig voor het woningmarktbeleid van BZK.

Verder gaan we steeds meer samenwerken met andere partijen zoals universiteiten en consultancybureaus om **datagedreven** werken bij de overheid te versterken. In deze samenwerking heeft elke partij zijn eigen expertise. Wij hebben veel kennis over data, anderen geven beleidsadviezen of maken toekomstscenario's.”

! Tip van Joost voor collega's

Stel van tevoren een duidelijke vraag die je beantwoord wilt krijgen: wat wil je uit de data halen? Anders verdrink je in je bestanden. En ga na of die data er in de toekomst ook nog zijn. De stabiliteit van de bron is niet altijd geborgd. Een voorbeeld is Weigh in Motion waarmee de belading van vrachtwagens op snelwegen wordt gemonitord. Dat was een belangrijke databron voor ons onderzoek. Nu is gebleken dat de weegpunten van Rijkswaterstaat er mogelijk op termijn uit gaan. Dus dan krijgen we die data niet meer.

Onderzoeksmethodiek 10:

Research community

Een research community is een online platform waarop je een vooraf geselecteerde groep mensen vragen kunt laten beantwoorden en verschillende taken kunt laten uitvoeren. Het platform is ook een afgeschermd discussieforum: alleen de respondenten (leden), de onderzoekers (moderatoren) en de opdrachtgever (meekijkers) hebben toegang.

Methodie: online platform voor vragen en discussie

In een research community komen vraagstukken aan bod die de opdrachtgever aanlevert, maar ook discussieonderwerpen die de leden zelf starten. Het is daarmee een instrument dat zowel antwoord geeft op gerichte vragen als waarmee je kunt signaleren welke (maatschappelijke) thema's leven onder de doelgroep. De inhoud en duur van een research community kunnen variëren:

Een continue community...	Een kortlopende community...
<ul style="list-style-type: none">• loopt meerdere maanden of jaren• is geschikt voor verschillende thema's en onderzoeksvragen• biedt de mogelijkheid om mensen in de tijd te volgen	<ul style="list-style-type: none">• heeft een looptijd van enkele weken• kent één duidelijk omschreven thema

Toepassing: van verkenningen tot tests

Een research community is te gebruiken voor veel verschillende typen vragen:

- Je kunt achterhalen welk gedrag mensen nu vertonen en dit in de tijd volgen. Dit gebeurt bijvoorbeeld aan de hand van huiswerk- of dagboekopdrachten. Zo worden in een verkennende fase van beleid de bestaande gedragsmechanismen inzichtelijk.
- In de ontwerpfase van beleid kun je zoeken naar mogelijke interventies om een beleidsdoel te bereiken. Dit kan cocreatief, dus samen met de leden van de community, en in meerdere rondes. Bijvoorbeeld door in het begin de potentie van mogelijke interventies te verkennen en op een later moment een meer uitgewerkt idee nog een keer te laten beoordelen.
- Je kunt online interventies testen. Denk aan zaken zoals de keuzearchitectuur bij formulieren, informatievoorziening via een

website of het laten uitproberen van een app. Het is mogelijk om varianten voor te leggen aan subgroepen binnen de community, bijvoorbeeld om de reacties te vergelijken.

Soort onderzoek: (grootschalig) online kwalitatief onderzoek

De research community is een vorm van online onderzoek waarin je verschillende onderzoeksinstrumenten kunt combineren, zoals: een korte vragenlijst, een (online) groepsdiscussie of forum, dagboekopdrachten, chatsessies en cocreatiesessies. Hiermee kun je individuele reacties van leden krijgen, maar ook zien hoe een groep gevraagd of ongevraagd ergens over denkt. De methodiek is bedoeld om inzichten te verkrijgen en niet als representatief onderzoek. Voor een langlopende community zijn rond de 150 leden gebruikelijk; voor een kortlopende community is 50 leden vaak voldoende.

Sterke en minder sterke punten van de research community

Sterke punten:

- Een research community stimuleert omgevingsgericht werken. Opdrachtgevers kunnen op ieder gewenst moment meekijken en luisteren naar wat hun doelgroep zegt en inzicht krijgen in wat mensen beweegt en wat voor hen relevante thema's zijn.
- Een research community is veelzijdig. Je kunt verschillende vraagstellingen en taken inzetten en combineren: dagboekonderzoek, ideegeneratie, et cetera.
- Een research community werkt snel en flexibel als deze eenmaal is opgezet. Waar een traditioneel onderzoek enkele weken kan duren, kun je met een community binnen een aantal dagen inzichten ophalen en soms zelfs binnen 24 uur. Dit maakt de

methode ook geschikt om in te zetten bij 'agile' werk- en ontwikkelprocessen.

Minder sterke punten:

- Het aantal respondenten is doorgaans te klein om representatieve uitspraken te doen. Ook kunnen de deelnemers meer dan gemiddeld betrokken zijn bij het onderzoeksthema. De uitkomsten geven daarom vooral een indicatie. Wel zijn de inzichten uit een community te onderbouwen met aanvullend kwantitatief onderzoek onder een representatieve steekproef.
- Leden kunnen na verloop van tijd beïnvloed raken als ze door de gestelde vragen steeds meer over een onderwerp te weten komen. Dit speelt vooral bij een continue community. Je kunt de objectiviteit verbeteren door mensen maximaal een half jaar te laten deelnemen en met regelmatig nieuwe leden aan de community toe te voegen.
- Je kunt in een community het gedrag van mensen niet rechtstreeks observeren. Wel kun je via huiswerk- of dagboekopdrachten dichterbij het werkelijke gedrag komen dan in traditioneel kwalitatief onderzoek.

Praktijkvoorbeelden: community 'Nederland denkt mee'

Research communities worden momenteel vooral ingezet om gedrag te verkennen en om communicatieve interventies te toetsen bij de doelgroep. Dienst Publiek en Communicatie doet dit in de continue research community 'Nederland denkt mee' voor onder meer het ministerie van Volksgezondheid Welzijn en Sport. Deze community telt 150 leden, een afspiegeling van het Nederlands publiek.

- **Campagne NIX18:** in de community is de slogan 'Iedereen heeft het over NIX' getest. Verder zijn vraagstukken behandeld met

ouders van kinderen tussen de 12 en 17 jaar. Daarbij is het huidige gedrag in kaart gebracht: wat zijn typische momenten waarop ouders het gesprek met hun kinderen voeren over drank? Maken zij hier wel of niet afspraken over? En waarom wel of niet?

- **Life-events:** Rijksoverheid.nl biedt informatie steeds vaker aan in de vorm van levensgebeurtenissen, zoals 'een kind krijgen', 'trouwen' of 'overlijden'. Gebruikers krijgen een overzicht van zaken die ze hiervoor moeten weten of regelen en instanties waarbij ze terecht kunnen. In de community zijn verschillende onderzoeken uitgevoerd rondom deze life-events om beter zicht te krijgen op de informatiebehoeften en het huidige informatiezoekgedrag. Met de resultaten laat de redactie van Rijksoverheid.nl de overzichten zo goed mogelijk aansluiten bij de behoeften.

Meer informatie

Dienst Publiek en Communicatie (DPC):
Marit Groeneveld, m.groeneveld@minaz.nl
Lizet van Tilburg, l.vantilburg@minaz.nl

Onderzoeksmethodiek 11:

Waarderend onderzoeken

Waarderend onderzoek (*appreciative inquiry*) is onderzoek met een positieve invalshoek om verandering in organisaties, groepen of gemeenschappen te ondersteunen. Het is gericht op het identificeren van kwaliteiten en sterktes: welk gedrag wil je versterken? Dit gebeurt in interviews en werksessies.

Methode: blik op best practices

In waarderend onderzoek worden organisaties of afdelingen begeleid om na te denken over hun eigen successen en verbetermogelijkheden. Dit activeert de intrinsieke motivatie van medewerkers om te werken aan verbetering. Met waarderend onderzoek bekrachtig je het positieve in organisaties en stimuleer je mensen om hiermee verder te gaan en van elkaars successen te leren. Dit gebeurt met alle stakeholders. Met een open, waarderende blik ga je op zoek naar best practices, die je vervolgens kunt delen met mensen binnen en buiten de organisatie. Zo kunnen ze gaan dienen als referentiekader voor gedrag.

Er is een uitgebreide methode voor waarderend onderzoek, gebaseerd op het werk van David Cooperrider. De Auditdienst Rijk heeft hier een verkorte versie van gemaakt. Het bestaat uit drie onderdelen:

- **Denken** – denken over wat in het verleden goed ging en dat waarderen;
- **Dromen** – verbeelden hoe het in de toekomst nog beter zou kunnen gaan;
- **Doen** – in een dialoog de organisatie helpen om het toekomstbeeld specifiek te maken en acties te formuleren om naar dit toekomstbeeld toe te werken.

Toepassing: in fase van verkenning en beleidsontwikkeling

Waarderend onderzoek zet je in als je interventies wilt ontwikkelen die op veel steun kunnen rekenen. Je stimuleert gewenst gedrag binnen organisaties en in de samenwerking met stakeholders om zo tot verandering te komen. Dit gebeurt op basis van goede voorbeelden die je hebt verkend.

Type onderzoek: kwalitatief

Het betreft veranderingsondersteunend, kwalitatief onderzoek. Mensen worden eerst in interviews gevraagd naar positieve ervaringen en verhalen rond een bepaald thema. Deze vormen de input voor een werksessie waarin de succesfactoren voor gewenst gedrag worden bepaald om daarmee samen een gewenste toekomst te verbeelden. Hieraan worden vervolgens concrete acties gekoppeld door en voor de deelnemers.

Sterke en minder sterke punten van waarderend onderzoek

Sterke punten:

- Door waarderend onderzoek worden mensen zich bewust van wat er goed gaat in de organisatie. Stilstaan bij gewenst gedrag is helemaal niet zo vanzelfsprekend. Mensen zijn vaak meer gericht op wat er misgaat. Wanneer bewust aandacht wordt geschonken aan wat goed gaat, kunnen ze er vervolgens ook bewust voor kiezen om deze sterke punten te behouden of uit te breiden.
- Omdat alle relevante stakeholders betrokken worden, komen interventies voor gedragsverandering tot stand die breed gedragen zijn. Door de wijze van onderzoek wordt de intrinsieke motivatie aangesproken die mensen stimuleert om te veranderen. Tegelijk is die betrokkenheid ook een randvoorwaarde om het onderzoek te laten slagen.

Minder sterke punten:

- Waarderend onderzoek biedt geen volledige analyse van een vraagstuk: het richt zich alleen op kwaliteiten en sterktes. Soms kan tijdens een waarderend onderzoek ook de behoefte ontstaan om aandacht te besteden aan wat niet goed gaat. Dit is apart te

onderzoeken, maar strikt genomen is dit geen onderdeel van waarderend onderzoek.

Praktijkvoorbeelden: Rijkswaterstaat en VWS

De Auditdienst Rijk (ADR) heeft onder meer deze twee waarderende onderzoeken uitgevoerd:

- **Rijkswaterstaat:** ADR heeft een inventarisatie gehouden van de meest motiverende omstandigheden om te werken aan kennis(opgaven). Rijkswaterstaat wil in de toekomst kunnen blijven beschikken over voldoende kennis. Er was al veel gedaan om het risico van onvoldoende kennis af te dekken. Maar er was nog weinig zicht op of hiermee de optimale omstandigheden gecreëerd waren om de kennisopgave te realiseren. Met het waarderend onderzoek zijn voor de kennisvelden assetmanagement en bouw- en onderhoudstechnologie drie zaken geïnventariseerd:
 1. condities die medewerkers stimuleren om kennis op te bouwen en te delen;
 2. motiverende condities om aan kennis te werken die zijn gemist;
 3. groeistappen voor een verdere ontwikkeling van het kennismanagement.
- **Ministerie van Volksgezondheid Welzijn en Sport (VWS):** er is een waarderend onderzoek uitgevoerd naar succesvolle werkwijzen bij praktijkteams in de zorg. In deze teams wil VWS samen met de praktijk problemen analyseren en oplossen. Voor het onderzoek zijn vijf voorbeelden geselecteerd waarbij VWS op een succesvolle manier samen met de praktijk een oplossing heeft gevonden voor een vraagstuk. Eén voorbeeld betrof verpleegkundigen en hospicemedewerkers die aan de bel trokken omdat ze steeds vastliepen als ze hulpmiddelen wilden aanvragen. Vanuit de vijf vraagstukken is verkend welke factoren hebben bijgedra-

gen aan de oplossing. Elk onderzocht voorbeeld is vervolgens beschreven in de vorm van een verhaal. De leerpunten en verhalen zijn gedeeld binnen VWS, ter inspiratie voor alle medewerkers bij VWS.

Meer informatie

Auditdienst Rijk (ADR):

Martine de Zeeuw, m.zeeuw@minfin.nl.

“Van de achten en negens kun je leren”

Astrid Paulussen

*programmamanager Kennismanagement bij Rijkswaterstaat,
over waarderend onderzoek*

“Bij Rijkswaterstaat steken we geregeld een **thermometer** in de organisatie: wat zijn de ontwikkelingen, wat loopt goed en kunnen we continueren, wat kunnen we bijsturen? Zo kijken we ook naar kennis bij RWS. Hoe houden we onze kennis op peil, bijvoorbeeld op het gebied van bouw- en onderhoudstechnologie?”

De Auditdienst adviseerde ons waarderend onderzoek, omdat die methode de mogelijkheid geeft om te ontdekken welke dingen goed werken en waarom. Veel evaluaties richten zich op wat mis gaat. Dat doet iets met de energie van de mensen die zich met hart en ziel ergens voor inzetten. Het is net als met een **schoolrapport**. Voor je het weet heb je het alleen over dat magere zesje, terwijl je de achten en negens over het hoofd ziet. Waardeer eens de factoren die het mogelijk maken om zo goed te presteren.

Zo'n benadering is zeker ook belangrijk bij de kennisontwikkeling, -deling en -borging bij RWS. Het zit in de mens gebakken om te **focussen** op het negatieve, maar dat is een valkuil voor verbetering. Het gaat ons erom dat we over de juiste kennis blijven beschikken voor de taken die mensen en politiek van RWS vragen.

We hebben bewust ADR gevraagd hiervoor externe onderzoekers met een andere **bril** en een andere aanpak. Hun sessies waren creatief en leverden veel op. Ze gebruikten onder meer plaatjes om mensen uit te nodigen hun ervaringen te vertellen. Bijvoorbeeld over stimulerende of belemmerende factoren zoals ICT of sociale factoren.

Wat we leerden? Je kunt kennisdelen niet afdwingen. Je kunt alleen de omgeving zo inrichten dat je kennisdeling faciliteert. Bijvoorbeeld door wat we wel noemen **slack in time** te creëren, te zorgen dat er regelmatig tijd is om kennis te delen. Om vakbijeenkomsten bij te wonen, om ervaringen digitaal beschikbaar te maken, om bij te praten bij de koffieautomaat. Bij grotere projecten weten mensen elkaar vaak goed te vinden. Omdat ze op dezelfde locatie werken, is er makkelijker gelegenheid voor ongeorganiseerde kennisuitwisseling. Medewerkers bij kleinere projecten op verschillende locaties zien elkaar vooral tijdens officiële overleggen. Daar missen we de spontane mogelijkheid tot uitwisseling. Dat is één van de dingen die we mee hebben kunnen geven aan lijnmanagers en kennistrekkers.”

! Tip van Astrid voor collega's

Vertel van tevoren de deelnemers goed wat de bedoeling is. Waarderend onderzoek is anders dan een reguliere enquête of een audit. Daar zijn er al veel van; mensen raken er soms 'evaluatiemoe' van. Onze ervaring is dat deelnemers heel enthousiast en gedreven zijn om ideeën toe te passen in de praktijk. Ze gingen gelijk met elkaars succesverhalen aan de slag. Die inspiratie voor vervolgstappen, dat is de kracht van deze methodiek.

Onderzoeksmethodiek 12:

Public design

Public design is een methode om maatschappelijke vraagstukken op te lossen met alle partijen die erbij betrokken zijn. De partijen komen bij elkaar in een cocreatief proces van opeenvolgende stappen. Je kunt de methode inzetten wanneer je samen een gedragsprobleem wilt verkennen en daar gezamenlijk een oplossing voor wilt ontwikkelen. Bijvoorbeeld: hoe kan de overheid de dalende vaccinatiegraad in Nederland tegengaan? Of: wie kan/moet wat doen in de klimaattransitie?

Methode: met elkaar op zoek naar nieuwe oplossingen

Public design volgt de principes van design thinking en past ze toe in het publieke domein. *Design thinking* kenmerkt zich door actieve betrokkenheid van alle partijen en een cocreatief en iteratief proces. Dit houdt in: een samenwerking waarin alle deelnemers invloed hebben op het proces en het resultaat, en een proces dat op zichzelf voortborduurdt.

Het proces bestaat uit zes stappen: ambitie, empathie, (her)definitie, ideeëngeneratie, prototype en test en evaluatie. In de eerste drie stappen ligt de nadruk op het begrijpen van de context, de betrokkenen en het probleem ('divergeren'). Pas als dat duidelijk is, kan er gewerkt worden aan de meest kansrijke inzichten en ideeën ('convergeren'). Er worden in rondes oplossingsrichtingen bedacht die aangescherpt worden op basis van feedback van betrokkenen. Zo'n oplossing (ontwerp) wordt tastbaar gemaakt en beproefd in de praktijk.

De waarde is vaak al bekend (hogere vaccinatiegraad of klimaatbestendigheid), maar onduidelijk is nog wie wat moet gaan doen en op welke manier. Om tot een goede aanpak te komen, worden gedurende het proces de intrinsieke drijfveren en onderliggende behoeftes van betrokkenen in kaart gebracht die relevant zijn voor een gedragsinterventie.

Toepassing: verkennend en als input voor beleidsontwikkeling

Public design gebruik je in de verkennende fase van beleid om inzichtelijk te maken waar (gedrags)problemen zich voordoen en hoe dat komt. Ook is het een werkwijze om met verschillende betrokkenen samen te werken aan een oplossing. Public design kan

een aanvulling zijn op tools voor gedragsverandering zoals de DOE MEE-tool van het ministerie van Infrastructuur en Waterstaat en CASI (Communicatie Activatie Strategie Instrument) van het ministerie van Algemene Zaken. Daarin is ook veel aandacht voor de doelgroep en mogelijke interventies, maar is het interactief betrekken van de doelgroep en gezamenlijk uitwerken en uitproberen van die oplossingen niet altijd onderdeel van het proces.

Type onderzoek: kwalitatief

Public design is een kwalitatieve methode. In verschillende stappen wordt onderzoek uitgevoerd. Denk aan etnografisch/locatieonderzoek, diepte-interviews en gebruikerstests.

Sterke en minder sterke punten van Public design

Sterke punten:

- Public design is geschikt voor complexe en vastgelopen vraagstukken.
- De nadruk ligt op intrinsieke waarden/behoeften van alle betrokkenen. De methodiek verhoogt het empathisch vermogen van een organisatie. Het met elkaar doorlopen van het design proces is een opbrengst op zich: de samenwerking verbetert en er ontstaat meer vertrouwen en energie.
- Er ontstaat een gedeeld beeld, breder gedragen oplossingsrichtingen en nieuwe ideeën.
- Je hebt snel resultaat.

Minder sterke punten:

- De methodiek stelt hoge eisen aan de procesbegeleiding, inclusief duidelijkheid over rollen, taken en verwachtingen. De opdrachtgever moet tegenover de betrokkenen eerlijk zijn over de beschikbare designruimte.

- Public design-trajecten kun je als opdrachtgever/organisatie niet uitbesteden. Je moet zelf aanwezig zijn en actief deelnemen.

Praktijkvoorbeelden: EZK en BZK

- **Informatiemarkt gaswinning:** EZK organiseert ieder jaar een groot aantal informatiemarkten over gaswinningsbesluiten. Zowel omwonenden als EZK waren niet gelukkig met het verloop van deze avonden. Bewoners voelden zich vaak niet gehoord in hun zorgen over gasvelden. Hierdoor werd het draagvlak niet vergroot en werden er zienswijzen ingediend die niet gingen over het winningsbesluit. Gezamenlijk hebben alle betrokken partijen (EZK, Rijksdienst voor Ondernemend Nederland, Vermilion, Staatstoezicht op de Mijnen, TNO, de gemeente Westerveld en de omwonenden) tijdens een designtraject gewerkt aan een verbeterde opzet van de lokale informatievoorziening. De bewoners zijn actief uitgenodigd en betrokken. Daarnaast zijn er thematafels die aansluiten op de zorgen en vragen van omwonenden. Ook actiegroepen hebben voortaan een eigen plek op de informatiemarkt.
- **Beleidslabs:** BZK werkt sinds 2014 met beleidslabs waarbij ook de principes van *design thinking* worden toegepast. In één van de labs was de vraag: hoe kunnen mensen van 55 tot 65 jaar met een koopwoning zo lang mogelijk comfortabel en veilig thuis blijven wonen? De eerste uitdaging: je inleven in de doelgroep en denken vanuit de gebruiker. Teamleden gingen daarvoor bij mensen thuis praten en kregen vaak ook een rondleiding door de woning. Dit leverde interessante (en soms aangrijpende) gesprekken op. Alle ervaringen kwamen terecht in een 'rodedradenverslag'. Aan de hand van dit verslag en een literatuurstudie haalde het team vervolgens ideeën op bij een groep externe betrokkenen. Drie ideeën zijn verder uitgewerkt en getest bij de

doelgroep. Een van de ideeën bleef over: de WOONBOX. Dit is een box met gratis artikelen en informatie over woningaanpassingen en innovaties.

Meer informatie

Dienst Publiek en Communicatie (DPC):
Lizet van Tilburg, l.vantilburg@minaz.nl
Marit Groeneveld, m.groeneveld@minaz.nl.

Onderzoeksmethodiek 13:

Randomized Controlled Trial

Met een Randomized Controlled Trial (RCT) kun je de effecten van een interventie op gedrag bepalen. RCT's zijn bekend van onderzoek naar de effectiviteit van medicijnen. Hierin wordt een groep patiënten willekeurig opgesplitst in twee groepen. De ene groep krijgt het medicijn waarvan de effectiviteit bepaald moet worden (interventiegroep) en de andere groep krijgt een placebo (controlegroep). De patiënt (en het liefst ook de arts) weet niet wie een placebo heeft gekregen en wie het medicijn. Verschillen in de ontwikkeling van het genezingsproces zijn dan geheel toe te rekenen aan het medicijn. Steeds vaker wordt de RCT ook ingezet als methodiek om de effectiviteit van een (beleids)interventie vast te stellen.

Methodie: causale verbanden bepalen

Met een RCT kun je causale effecten (oorzaak-gevolgrelaties) bepalen van een interventie op gedrag. Er wordt een steekproef getrokken en die mensen worden aselekt verdeeld over de controle- en interventiegroep.

Soms wordt gedrag getest in een laboratoriumsituatie, een gecontroleerde omgeving, en soms in een praktijksituatie, een natuurlijke omgeving. Die laatste situatie heet een veldexperiment. De voorkeur ligt bij veldexperimenten, omdat de validiteit van laboratoriumexperimenten voor veel gedragingen vaak beperkter is. De volledige context is niet na te bouwen in het lab. Soms kun je labexperimenten gebruiken als eerste inschatting om te bepalen in hoeverre een interventie kan werken om deze daarna in de praktijk te testen. In het veld is een RCT soms lastig uit te voeren. De interventie moet correct, nauwkeurig en consequent verlopen. Een grote uitvoeringsdiscipline is daarbij vereist: alle betrokkenen moeten de goede dingen doen. Dat maakt dat de doorlooptijd vaak langer is dan technisch noodzakelijk.

Toepassing: voor de evaluatie van een interventie, vooraf of achteraf

RCT's zijn bij uitstek geschikt om een voorgenomen interventie of varianten van interventies te testen op effectiviteit. Je kunt door middel van een pilot bepalen wat het gedragseffect is van ontworpen interventies. Het is dan een evaluatie, maar vóór de daadwerkelijke inzet van de interventie (ex-ante-evaluatie). Als de RCT positieve gedragseffecten laat zien, kan de interventie breed toegepast worden.

Type onderzoek: kwantitatief

Een RCT is een kwantitatieve methode. Met een goede steekproef en voldoende grote groepen levert het onderzoek representatieve resultaten op.

Sterke en minder sterke punten van de RCT

Sterke punten:

- De methodiek is zeer geschikt om oorzakelijke verbanden aan te tonen tussen het effect en de interventie. Stevig bewijs van de effectiviteit is soms wenselijk. Wil je beleid op kennis baseren, in plaats van op waarnemingen op het oog of op inschattingen, dan is het belangrijk andere oorzaken van veranderingen in het gedrag te kunnen uitsluiten. Afhankelijk van de opzet (de grootte van de onderzoeksgroep) zijn met een RCT ook kleine effecten meetbaar.
- Ook bewijst de RCT zijn nut in complexe situaties: soms blijken de effecten van een interventie op gedrag lastig te voorspellen. Neem een boete voor ouders die hun kinderen te laat ophalen van de crèche. Vooraf werd gedacht dat ouders door zo'n interventie wel op tijd bij de crèche zouden zijn. Een RCT wees echter uit dat een boete niet hielp. Uit nader onderzoek bleek dat ouders hun kind juist later ophaalden 'want je betaalde er toch ook voor'.

Minder sterke punten:

- De RCT vraagt gedegen methodologische kennis en ervaring. Die is nodig om bij de opzet van het experiment al zo veel mogelijk alternatieve verklaringen en potentieel versturende factoren uit te kunnen sluiten.
- Je kunt met een RCT slechts beperkt gedragsmechanismen achterhalen: hóe bereikt de interventie de gemeten effecten? Als

je vooraf een gedegen hypothese hebt over achterliggende gedragsmechanismen, is die te toetsen in een opzet met meerdere interventiegroepen. Als je die niet hebt, ligt het meer voor de hand een verkennende methode te kiezen zoals diepte-interviews. Ook een combinatie met een enquête geeft een completer beeld van de achterliggende mechanismen. Zie methodieken 1 en 6.

Praktijkvoorbeelden: effect op energiebesparing en duurzame consumptie

- **RVO-mails:** in een veldexperiment is een alternatieve e-mail getest van de Rijksdienst voor Ondernemend Nederland (RVO) aan bedrijven die veel energie verbruiken. Hierin was gebruik gemaakt van gedragsinzichten. Wat bleek? De bedrijven die de alternatieve, korte mail kregen, vroegen drie keer vaker informatie over hun energieverbruik op dan met de oorspronkelijke mail. Met behulp van een enquête kon bepaald worden dat dit ook leidde tot evenredig vaker vertonen van energiebesparingsgedrag. RVO heeft de alternatieve e-mail nu standaard ingevoerd.
- **Energieverbruiksmanager:** enkele energiebedrijven willen vaststellen welke effecten hun energieverbruiksmanager (EVM) heeft op energiebesparing. Dat doen ze met een RCT. Zo krijgen ze beter zicht op de daadwerkelijke effecten van hun product, in plaats dat de bedrijven uit moeten gaan van schattingen op basis van buitenlandse experimenten. Effecten in het buitenland kunnen afwijken van die in Nederland omdat de context verschilt. Daarnaast krijgt het ministerie van Economische Zaken en Klimaat zicht op de effecten van de inspanningen van deze energiebedrijven. De experimenten lopen.
- **Vleesconsumptie:** met een veldexperiment, in combinatie met een enquête, heeft het Planbureau voor de Leefomgeving het

draagvlak onderzocht voor verduurzaming van consumptie onder burgers. Deelnemers kregen de keuze tussen biologisch vlees of gangbaar vlees onder verschillende situaties: wat koop je? Consumenten bleken 'conditioneel coöperatief' hun aankopen te doen. Zij willen wel duurzamer kopen, maar niet één van de weinigen zijn ('gekke Henkie'). Motiverende interventies (zoals subsidies) hebben de voorkeur boven meer dwingende maatregelen.

Meer informatie

Planbureau voor de Leefomgeving (PBL):
Kees Vringer, kees.vringer@pbl.nl.

“Een kans om aan te tonen dat ons werk verschil maakt”

Ruth Gonschorrek

medewerker toezicht bij ACM, over Randomized Controlled Trials

“Mijn afdeling van de Autoriteit Consument en Markt stuurt vaak brieven naar een klein aantal bedrijven. Dat is lastig als je effect wilt meten. Terwijl je je wel wilt verantwoorden over projecten. Toen er een brief uit moest naar een **grotere groep**, alle postbedrijven en telecomaanbieders, dacht ik gelijk: dit is onze kans! We konden een RCT doen.

Alle 312 bedrijven in deze branche moeten jaarlijks een formulier insturen met hun omzetgegevens. Dat is om de toezichtskosten te kunnen berekenen. 60% had al gereageerd. Vraag aan mij als gedragsbeïnvloeder was: hoe halen we die overige 40% over met zo min mogelijk moeite? Ik heb een variant gemaakt op onze standaard-**rappelbrief** waarin ik gedragsinzichten heb toegepast. De brief werd korter, leesbaarder en in doelgroepgerichte taal geschreven. Verder gebruikte ik de sociale norm: de meeste bedrijven hebben de gegevens al ingediend. En ik benadrukte: als je niet reageert, beschouwen wij dat als een bewuste keuze. Uit ander onderzoek was bekend dat dit uitstelgedrag kan voorkomen.

Deze drie aanpassingen zaten tegelijk in de brief. Idealiter zou je ze **apart testen**, dat is zuiverder, maar de groep was te klein om op te splitsen. Het mogelijke effect stond bij ons voorop. De ene helft kreeg de standaardbrief, de andere de nieuwe brief. Bij de telecomaanbieders bleek er geen verschil. De herinnering aan sich was blijkbaar al effectief. Bij de postbedrijven was er wel een verschil: degenen die de aangepaste brief hadden gekregen, reageerden **significant** vaker. Misschien logisch want dat zijn Nederlandse bedrijven die elkaar goed kennen. Het kan zijn dat de sociale norm daar eerder werkt. Misschien spreken die internationale telecomaanbieders wel geen Nederlands? Dat was wel een les: we weten maar weinig over onze doelgroep.

Het was best nog wat werk om de nieuwe brief te verwerken in de processen van ACM. De groepen moesten handmatig **random** verdeeld worden, bestanden moesten gekoppeld. Dat is ook te automatiseren als je zulke experimenten vaker doet. We zijn al met de volgende bezig.”

! Tip van Ruth voor collega's

Ken je doelgroep. Anders meet je niet wat je wilt meten. Voor het nieuwe experiment stuurden we een brief eerst naar 40 deelnemers die ik vervolgens nabelde met de vraag wat ze van de brief vonden. Heel inzichtelijk! Met zo'n voortraject kun je daarna een betere RCT houden.

Onderzoeksmethodiek 14:

Praktijkexperimenten

Een praktijkexperiment is een experiment waarbij de onderzoeksgroepen niet in een laboratoriumsetting getest worden maar in de werkelijke wereld. Er zijn veel situaties waarin je zo'n experiment kunt houden, bijvoorbeeld op scholen, in winkels, ziekenhuizen, in kantoorgebouwen of op straat. Met een praktijkexperiment zijn causale verbanden te onderzoeken, bijvoorbeeld wanneer je benieuwd bent naar het effect van één of meer interventies in de praktijk.

Methodie: interventie- en controlegroep in het veld

In een praktijkexperiment wordt de onderzoeksgroep bij voorkeur willekeurig ingedeeld in verschillende interventiegroepen en een controlegroep (die geen interventie krijgt). Omdat het experiment in het dagelijks leven wordt gehouden, kun je niet altijd de opzet van een Randomized Controlled Trial aanhouden (zie methodiek 13). Dan zijn mensen niet willekeurig toe te wijzen, maar probeer je die willekeur te benaderen. Bijvoorbeeld door wel/geen interventie te verdelen over natuurlijke groepen, zoals vergelijkbare klassen op school. Of door één groep varianten van interventies te geven die met elkaar vergeleken worden, zoals wanneer bedrijven verschillende sollicitatiebrieven beoordelen.

Zo kun je bijvoorbeeld etnische discriminatie op de arbeidsmarkt onderzoeken. Twee (of meer) sollicitanten die alleen van elkaar verschillen in etnische achtergrond, reageren op dezelfde vacature. Vervolgens meet je het gedrag van de werkgever: worden mensen uit de ene groep bijvoorbeeld minder vaak uitgenodigd voor een gesprek dan uit de andere groep? Ook is het mogelijk om het effect van etnische achtergrond te onderscheiden van het effect van productieve kenmerken zoals werkervaring en een goed cv.

Toepassing: causaal verband achterhalen ter verkenning of evaluatie van interventie

Praktijkexperimenten kunnen worden ingezet wanneer er vragen zijn over een (mogelijk) causaal verband. Zoals tijdens een probleemanalyse: heeft etnische achtergrond een effect op de kans op een baan? Of om mogelijke interventies te evalueren: is jeugdcriminaliteit te voorkomen door een lesprogramma dat de politie verzorgt op scholen? Leiden andere standaardopties bij uitvoeringsorganisatie DUO tot lagere studieschulden? Door de

methodologische opzet zijn dergelijke vragen met de hoogst haalbare zekerheid te beantwoorden.

Type onderzoek: kwantitatief onderzoek

Een praktijkexperiment is een vorm van kwantitatief onderzoek. Om te kunnen vaststellen of de verschillen tussen de groepen geen toeval zijn en iets zeggen over het onderzochte verband, zijn er groepen nodig die groot genoeg zijn voor een statistische toets. Het experiment is uiteraard uit te breiden met kwalitatief onderzoek, zoals diepte-interviews (zie methodiek 1).

Sterke en minder sterke punten van praktijkexperimenten

Sterke punten:

- Een praktijkexperiment is een betrouwbare manier om een oorzaak-gevolgrelatie vast te kunnen stellen tussen een interventie en gedrag.
- Doordat het experiment wordt gehouden in de praktijk waarin het gedrag natuurlijk voorkomt, zegt het gemeten effect vaak meer over de werkelijkheid dan in labonderzoek.

Minder sterke punten:

- De opzet en uitvoering van een praktijkexperiment kosten relatief veel tijd. Behalve de onderzochte interventie(s) moeten alle andere factoren constant blijven.
- Je weet na afloop of iets effectief is of niet, maar niet waarom dat zo is. Dit kun je mogelijk ondervangen door te werken met variaties in de opzet, waardoor je meer zicht krijgt op mechanismen. Ook kun je kwalitatief onderzoek toevoegen om informatie te verzamelen over de beweegredenen van mensen.

- De methode vraagt om verschillende behandeling van verschillende groepen. Dit is niet altijd haalbaar of wenselijk.

Praktijkvoorbeeld: op afkomst afgewezen

Het Sociaal en Cultureel Planbureau (SCP) heeft in 2014 een praktijkexperiment uitgevoerd naar etnische discriminatie op de arbeidsmarkt. Bekeken werd in hoeverre er sprake is van discriminatie van Marokkaanse en Hindoestaanse Nederlanders die werk zoeken op laag en midden functieniveau in de regio Haaglanden. In het onderzoek reageerden gelijkwaardige (fictieve) kandidaten op dezelfde (echte) vacature. Ze verschilden alleen in etnische achtergrond. Zo kon een verschil in uitkomst worden toegeschreven aan de etnische achtergrond van de sollicitanten. Er zijn 504 sollicitaties verstuurd naar 176 verschillende vacatures in de regio Haaglanden. De reacties van werkgevers werden geturfd: ófwel de sollicitant werd uitgenodigd voor een sollicitatiegesprek (of de werkgever gaf op een andere manier aan interesse in de sollicitant te hebben); ófwel de sollicitant kreeg een afwijzing of helemaal geen reactie.

Uit het experiment blijkt dat er wordt gediscrimineerd op de arbeidsmarkt van Haaglanden. De autochtone sollicitanten hadden 34% kans op succes. Bij Hindoestaans-Nederlandse sollicitanten was dat 23% en bij Marokkaans-Nederlandse sollicitanten 19%, terwijl ze een gelijkwaardig cv hadden. Met andere woorden: autochtone kandidaten hadden zeker 1,5 keer zoveel kans op een succesvolle sollicitatie. Het SCP publiceerde hierover 'Op afkomst afgewezen' (2017). Eerder vergelijkbaar onderzoek is bijvoorbeeld 'Liever Mark dan Mohammed?' (2010).

Meer informatie

Sociaal en Cultureel Planbureau (SCP):
Iris Andriessen, i.andriessen@scp.nl
Marcel Coenders, m.coenders@scp.nl.

Onderzoeksmethodiek 15:

Propensity Score Matching

Propensity Score Matching (PSM) is bruikbaar om effecten te meten van een interventie als je geen experiment kunt opzetten zoals een RCT (zie methodiek 13). In een experiment wijs je personen of bedrijven vooraf willekeurig toe aan een controle- of interventiegroep. Bij een PSM maak je de groepsindeling achteraf, als de interventie al is ingezet. Dat maakt het een quasi-experiment. Uit de beschikbare data wordt een controlegroep geconstrueerd die zo goed mogelijk lijkt op de interventiegroep. Die matchende groepen vergelijk je vervolgens met elkaar, zoals bij een echt experiment.

Methodie: een geschikte controlegroep selecteren

Met PSM is uit een grote groep ondernemingen of personen die geen gebruik maken van een instrument, een geschikte controlegroep samen te stellen. Namelijk zó dat die op zoveel mogelijk relevante kenmerken vergelijkbaar is met de interventiegroep. Alleen het gebruik van het instrument verschilt. Neem als voorbeeld een ondersteuningsregeling voor ondernemingen die in financieel zwaar weer verkeren. De vraag is: voorkomt deze regeling dat bedrijven kopje onder gaan? Een deel van de ondernemingen die het instrument zouden mógen gebruiken, doet dat echter niet. Wanneer je deze groep lukraak met de interventiegroep zou vergelijken, kan dit leiden tot een onderschatting van het effect van het beleidsinstrument. Mogelijk gaat het hier om gezondere bedrijven die het zonder het instrument ook wel redden. Met PSM wordt per bedrijf in de interventiegroep een bedrijf gevonden dat op zoveel mogelijk relevante en meetbare kenmerken vergelijkbaar is.

Dat matchen gebeurt aan de hand van *propensity scores*: scores voor de kans dat een bedrijf gebruik maakt van het instrument, of ze dit nu daadwerkelijk doen of niet. Vóór matching zitten er in de controlegroep vaak veel bedrijven waarvan het onwaarschijnlijk is dat ze gebruik maken van het instrument. Na matching is de verdeling in de controlegroep vergelijkbaar met die in de interventiegroep. Zo kunnen de gevonden verschillen met meer zekerheid verklaard worden door de interventie.

Toepassing: voor effectonderzoek

PSM wordt gebruikt voor evaluaties, waarin je het effect van een beleidsinstrument (interventie) vaststelt. Andere methodieken voor

quasi-experimenteel onderzoek zijn: Regression Discontinuity (methodiek 16) en Difference-in-Difference (methodiek 17).

Soort onderzoek: kwantitatieve analyse

PSM is een (analyse)methodiek die wordt toegepast binnen kwantitatief, quasi-experimenteel, onderzoek. Vaak wordt aanvullend onderzoek ingezet om te achterhalen waaróm een bepaald effect optreedt.

Sterke en minder sterke punten van Propensity Score Matching

Sterke punten:

- Met PSM is achteraf informatie over een causaal verband te destilleren uit de data, zonder dat vooraf een zuivere RCT is opgezet.
- Vaak verschillen personen of bedrijven op veel kenmerken van elkaar. Normaal gesproken moet dan worden 'gewogen' op al die verschillende relevante variabelen om ze te kunnen vergelijken. Met PSM is dat niet meer nodig, omdat deze methodiek die variabelen terugbrengt tot slechts één variabele: de *propensity score*.
- Je kunt met PSM voorwaarden stellen aan de manier van matchen. Je kunt 'hard' matchen op bepaalde kenmerken. Er kan bijvoorbeeld voor gezorgd worden dat bedrijven die actief zijn in de industrie, alleen aan andere industriebedrijven worden gekoppeld. Ook is een maximum te stellen aan de mate waarin koppeltjes personen of bedrijven van elkaar mogen verschillen. Dit beperkt het risico van een slechte match.

Minder sterke punten:

- De methodiek stelt hoge eisen aan de data. Zo moet de groep personen of bedrijven waaruit een controlegroep wordt samen-

gesteld, voldoende groot zijn. Verder moeten er data beschikbaar zijn over de relevante kenmerken waarop gematcht wordt. En er moet gematcht worden op variabelen die niet beïnvloed zijn door gebruik van het instrument. Vaak betekent dit dat je ook data nodig hebt van het jaar voorafgaand aan het instrumentgebruik.

- Er is niet één manier van matchen die de beste resultaten oplevert. Zo is het bijvoorbeeld mogelijk om per waarneming in de interventiegroep één of juist meerdere personen of bedrijven te koppelen. Het kan tijdrovend zijn om het juiste matchingsalgoritme te vinden, en dit moet voor ieder onderzoek opnieuw gebeuren.
- Matching werkt alleen voor die waarnemingen waarvan de *propensity score* gelijk is. Als de kans op instrumentgebruik in de interventiegroep altijd erg hoog is en in een groep waaruit een controlegroep samengesteld wordt altijd erg laag, wordt het moeilijk om paren van waarnemingen te vinden. De bedrijven of personen verschillen dan te veel.

De nadelen van PSM zijn deels te ondervangen door een gestructureerde dataverzameling. Zo werkt het Centraal Bureau voor de Statistiek (CBS) voor dit soort analyses bij voorkeur met integrale registerdata. Verder zorgt overleg tussen inhoudelijk en technisch experts voor goed onderbouwde keuzes in het matchingsproces.

Praktijkvoorbeeld: effectonderzoek SZW-regeling

Het CBS heeft PSM onder meer gebruikt in een tussentijdse evaluatie van de Overbruggingsregeling transitievergoeding voor het ministerie van Sociale Zaken en Werkgelegenheid. Hierin is het effect onderzocht van de regeling op de levensvatbaarheid en de bedrijfsprestaties van kleine werkgevers. Gematcht is op achtergrondkenmerken van bedrijven en op variabelen die van invloed

zijn op hun prestaties, zoals: de omzet, het aantal werknemers, het aantal ontslagen in het voorgaande kwartaal, het eigen vermogen, de leeftijd en sector. Ook de beschikbaarheid van data bepaalde de keuze. De analyse is uitgevoerd op het deel van de gebruikers waarvoor het CBS een vergelijkbare niet-gebruiker heeft gevonden. In de interventiegroep blijken inderdaad minder faillissementen voor te komen dan in de controlegroep. De regeling doet waar die voor bedoeld is: werkgevers in staat stellen om personeel te ontslaan om faillissementen (en daarmee meer ontslagen) te voorkomen. De werkgevers die de regeling gebruiken, staan er vervolgens echter niet veel beter voor dan werkgevers die dat niet doen.

Meer informatie

Centraal Bureau voor de Statistiek (CBS):
Tommy Span, t.span@cbs.nl.

Onderzoeksmethodiek 16:

Regression Discontinuity

Regression Discontinuity (RD) gebruik je in quasi-experimenteel onderzoek om achteraf vast te stellen of een interventie effectief is geweest. Het is toepasbaar als er een strikte grens is gesteld voor deelname aan de interventie, bijvoorbeeld een regeling voor mensen van 45 jaar of ouder. Door de uitkomsten voor 44-jarigen (die niet in aanmerking komen) te vergelijken met die van 45-jarigen (die wel in aanmerking komen), is het effect van de regeling te meten.

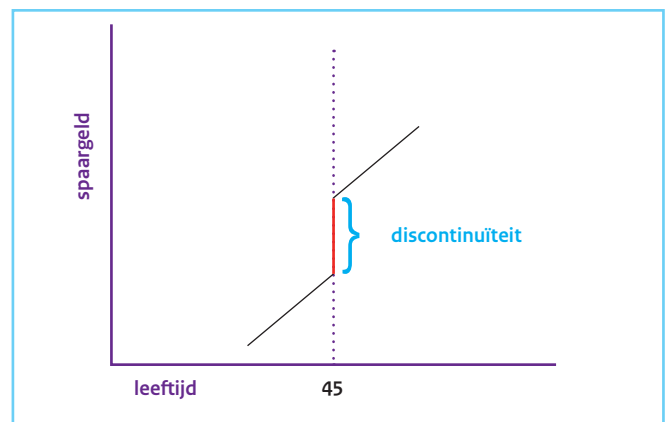
Methode: twee specifieke groepen vergelijken voor effectmeting

Niet altijd kun je in de praktijk uittesten of een interventie werkt, bijvoorbeeld omdat het niet wenselijk is om te loten wie de interventie krijgt en wie niet (zoals in experimenteel onderzoek, met een Randomized Controlled Trial, zie methodiek 13). Dan kun je kiezen voor een quasi-experimentele onderzoeksopzet. Hierbij maak je gebruik van 'toevallige randomisatie' in het beleid zelf, waardoor er verschillen tussen groepen ontstaan die in alle andere opzichten erg op elkaar lijken. Regression Discontinuity is zo'n opzet, naast bijvoorbeeld Difference-in-Difference (zie methodiek 17).

In het voorbeeld van een spaarregeling voor 45-plussers: als er een duidelijk verschil in uitkomst is tussen de 44- en 45-jarigen, zal dat komen door de regeling. En niet door hun verschil in leeftijd. Dat verschil is verwaarloosbaar klein, is de veronderstelling. En hoe kleiner de range van leeftijden gemaakt wordt (van 44,9-45,1 jaar), hoe aannemelijker die veronderstelling is. Zie de figuur: de relatie tussen leeftijd en spaargeld maakt een sprongetje, oftewel is discontinu, rond de leeftijdsgrens van 45 jaar. Waar de doorgaande lijn onderbroken wordt, zit het effect van de regeling. Deze methodiek werkt niet alleen met een leeftijdscriterium, maar ook met een geografisch criterium, positie op een ranglijst, et cetera.

Toepassing: in fase van evaluatie

Het is een methode voor ex-postevaluatieonderzoek. Je kunt RD gebruiken om aannemelijk te maken óf een beleidsinterventie een causaal effect heeft gehad en om inzicht te geven in hoe groot dat effect is.



Type onderzoek: kwantitatieve analyse

Regression Discontinuity is een kwantitatieve methode voor analyse van al verzamelde data. Als je wilt weten waarom een bepaald effect optreedt, wordt vaak gekozen voor aanvullend (kwalitatief) onderzoek.

Sterke en minder sterke punten van Regression Discontinuity

Sterke punten:

- Het is een relatief goede benadering van een RCT voor de groep rond de discontinuïteit: het is aannemelijk dat controlegroep en interventiegroep weinig van elkaar verschillen. Daardoor kun je vrij betrouwbare conclusies trekken over de effectiviteit van de beleidsinterventie.
- De onderzoeker heeft in principe genoeg aan één waarneming per individu. Een longitudinale dataset – ofwel meerdere

waarnemingen voor hetzelfde individu – is niet noodzakelijk, wat vaak veel tijd en geld kan schelen.

Minder sterke punten:

- Externe validiteit: in hoeverre zijn de resultaten te vertalen naar andere (deel)groepen of zelfs naar de hele populatie? Een quasi-experiment op basis van een strikte grens levert als resultaat: het effect voor mensen die zich dicht bij die grens bevinden. In het eerdere voorbeeld is dat het effect van de regeling op mensen rond de 45 jaar. Welk effect die regeling heeft op 25-jarigen of juist op 50-plussers is op basis van deze ene leeftijdsgrens niet zonder meer te bepalen. De externe validiteit is makkelijker te onderbouwen als er sprake is van meerdere grenscriteria, bijvoorbeeld bij een staffel (zoals drie extra vrije dagen voor 50-plussers, vijf voor 55-plussers en tien voor 60-plussers).
- RD stelt eisen aan de beschikbare data: door te focussen op een beperkte subpopulatie, zoals alleen de 44- en 45-jarigen, zijn er veel observaties nodig voor voldoende power in de analyses. Aan de andere kant: observaties verder weg van die grens zijn niet nodig.

Praktijkvoorbeelden: effecten op onderwijsdeelname

- **Het effect van de leerplichtwijziging van 1971 op onderwijsdeelname, inkomen en de kans op werk:** een klassiek voorbeeld van een effectmeting via RD gaat over de leerplicht. In 1971 is de leerplichtige leeftijd opgetrokken van 14 naar 15 jaar. Wat zijn de effecten geweest? De RD-analyse laat zien dat de wetwijziging heeft geleid tot een daling van het aantal voortijdig schoolverlaters, maar dat leerlingen niet meer zijn gaan verdienen of betere kansen op de arbeidsmarkt hebben gekregen.

- **Het effect van scholingsbeurzen voor leraren op hoger onderwijsdeelname en –afroding:** het Centraal Planbureau (CPB) heeft in deze studie de effecten geschat met een RD-design. Er was een regressiediscontinuïteit ontstaan in de kans om een beurs toegewezen te krijgen. Dat kwam doordat het aantal aanvragen in de eerste ronde veel hoger was dan het budget toeliet en de lerarenbeurzen vervolgens zijn toegewezen op volgorde van binnenkomst van de aanvragen. Het geschatte effect op de kans voor leraren om in het onderwijs blijven, was positief.

Meer informatie

Centraal Planbureau (CPB):

<https://www.cpb.nl/contactformulier>.

Onderzoeksmethodiek 17:

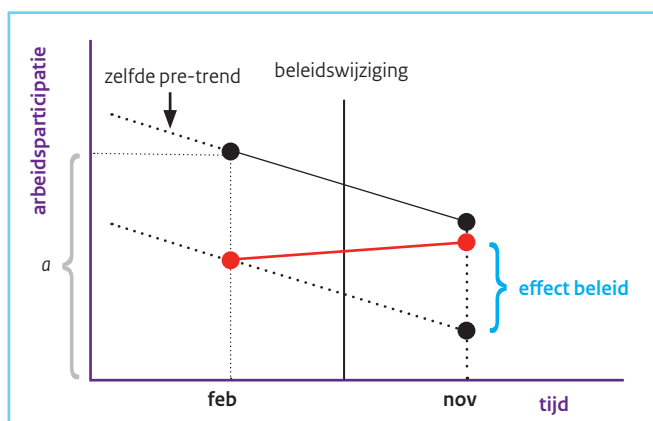
Difference-in-Difference

Difference-in-Difference (DiD) is een methode om achteraf vast te stellen of een interventie effectief is geweest. Het is toepasbaar als twee groepen voorafgaand aan de interventie een zelfde trend in gedrag vertonen en de interventie zich op één van beide groepen richt. Het ‘verschil in hun verschil’, de verandering in de trend van de interventiegroep, wijst dan op het effect van de interventie. DiD valt daarmee onder het quasi-experimenteel onderzoek.

Methode: twee parallelle groepen vergelijken voor effectmeting

Niet altijd kun je in de praktijk met een Randomized Controlled Trial (zie methodiek 13) uittesten of een interventie werkt, bijvoorbeeld omdat het niet wenselijk is om te loten wie de interventie krijgt en wie niet. Dan kun je kiezen voor een quasi-experimentele onderzoeksopzet. Difference-in-Difference is zo'n opzet, naast Propensity Score Matching en Regression Discontinuity (zie methodieken 15 en 16).

Een Difference-in-Difference (DiD)-design vergelijkt de uitkomsten van twee groepen die gelijk opgaan vóór de interventie en waarvan de ene groep wel de interventie heeft ondergaan en de andere niet. De figuur hieronder geeft aan hoe je uit de vergelijking tussen deze twee groepen het beleidseffect kunt identificeren.



Denk bijvoorbeeld aan een werkbonus om de arbeidsparticipatie van een bepaalde groep te stimuleren. De beleidsinterventie heeft geen effect op de controlegroep (bovenste zwarte stippen), maar wel op de groep voor wie die bedoeld is, de interventiegroep (rode stippen). De controlegroep bootst zo de situatie na in het geval dat het beleid hetzelfde was gebleven. Een juiste controlegroep lijkt zoveel mogelijk op de interventiegroep. De onderzoeker kan corrigeren voor verschillen in waargenomen kenmerken, bijvoorbeeld leeftijd.

Toepassing: in fase van evaluatie

DiD is bruikbaar voor ex-postevaluatieonderzoek. Je kunt aannemelijk maken óf een beleidsinterventie een causaal effect heeft en hoe groot dat effect is.

Type onderzoek: kwantitatieve analyse

Het betreft een kwantitatieve methode. Om te achterhalen waarom een effect optreedt, zal vaak aanvullend (kwalitatief) onderzoek gewenst zijn.

Sterke en minder sterke punten van Difference-in-Difference

Sterke punten:

- Een goed uitgevoerde DiD-analyse kan achteraf informatie over een causaal verband destilleren uit de data, zonder dat vooraf een zuivere RCT is opgezet.
- Je vindt een effect dat geldig is voor de gehele interventiegroep en geen lokaal effect (zoals bij Regression Discontinuity waar de effecten gelden voor een deel van de groep).

Minder sterke punten:

- Je bent afhankelijk van de relatief sterke aanname dat er sprake is van een zelfde trend in gedrag tussen de interventiegroep en de controlegroep en dat zij niet verschillen op relevante niet-waargenomen kenmerken. Dit is niet evident en ook niet direct te testen. Een zelfde trend is aannemelijker als aangetoond kan worden dat tot aan het moment van de onderzochte beleidsinterventie de ontwikkeling inderdaad parallel was. Als groepen hetzelfde reageren op eerdere beleidsveranderingen, dan weegt dit aandachtspunt minder zwaar. Dit vergroot wel het belang van een langere tijdreeks voor de analyse.
- Voor dit type onderzoek zijn (anders dan bij Regression Discontinuity) data nodig over een voldoende lange periode, bij voorkeur vanaf een aantal jaren voor de wijziging.

Praktijkvoorbeelden: effect van kinderopvangsubsidies

- **Kinderopvangsubsidies en arbeidsparticipatie; empirisch bewijs van een grote Nederlandse hervorming:** in de periode 2005-2009 heeft de overheid de subsidies voor kinderopvang flink verhoogd, waardoor de gemiddelde ouderbijdrage halveerde. Het Centraal Planbureau (CPB) schatte met een DiD-design het effect hiervan op het arbeidsaanbod: gaan moeders meer werken? Controlegroep waren ouders met een ouder kind. Het effect bleek niet groot. Het aantal gewerkte uren van moeders steeg met 1,1 uur per week. Analyse laat zien dat dit vooral het gevolg is van een toename in het aantal uren per werkende, en in mindere mate van niet-werkenden die gaan werken.
- **Quasi-experimenteel bewijs voor de relatie tussen subsidies voor kinderopvang en de kwaliteit van kinderopvang:** ook hier gebruikte het CPB een DiD-design. Daaruit bleek dat een lagere subsidie voor kinderdagopvang gemiddeld leidde tot een

significante vermindering van de kwaliteit van de aangeboden opvang ten opzichte van een controlegroep van peuterspeelzalen waar geen wijzingen in de financiering waren.

Meer informatie

Centraal Planbureau (CPB):

<https://www.cpb.nl/contactformulier>.

“Hoe een RCT een quasi-experiment werd”

Pim den Hertog

partner van innovatiebureau Dialogic, over quasi-experimenten

“Wij waren een Randomized Controlled Trial van plan, maar het werd een quasi-experiment. Wat was er aan de hand?”

Het ministerie van Economische Zaken wilde het effect weten van de zogeheten *service design voucher*. De **maakindustrie** gaat steeds meer product-dienstcombinaties verkopen in plaats van enkel producten. Ze bieden klanten zagezegd geen kopieerapparaat, maar kopieën. De kleine maakbedrijven vinden het lastig om die switch te maken, om primair te denken in oplossingen voor de klant. Daarom is de service design voucher bedacht – een soort kennistegoedbon. De mkb’er kan die inwisselen voor kennis bij bijvoorbeeld een adviesbureau in de creatieve industrie.

RVO had geld om 80 vouchers toe te kennen. We wilden **mkb’ers loten** zodat we een gerandomiseerd experiment konden houden. Maar er kwamen ondanks uitgebreide voorlichting slechts 110 aanvragen van mkb’ers binnen. Hiervan bleken er maar 60 aan de voorwaarden te voldoen. Deze kregen daarom allemaal een voucher. Er was een officiële regeling van gemaakt, dus RVO kon het niet maken het bedrag niet uit te keren. Uiteindelijk hebben 44 bedrijven die tegoedbon ook echt verzilverd.

We hadden dus wel een interventiegroep, maar geen gerandomiseerde controlegroep. Toen hebben we er een quasi-experiment van gemaakt en een **controlegroep samengesteld**. Daarin zaten die 16 bedrijven die de bon niet verzilverd hadden, de 50 bedrijven waarvan de aanvraag afgewezen was en de overige bedrijven die ook aan de o-meting hadden meegedaan. Van interventiegroep en controlegroep konden we de verschillen zien in hun dienstenontwikkeling tussen de o- en de 1-meting, dus vooraf en achteraf.

Daarop deden we een Difference-in-Difference-analyse: nemen de verschillen in de ene groep nu harder toe dan in de andere groep? Je hebt te maken met een *moving target*: bedrijven met en zonder voucher ontwikkelen zich allebei in de interventieperiode. En je wilt het netto-effect weten van de interventie, dus je kijkt naar het **verschil in het verschil**. Eerst is er binnen de groepen gekeken of de gemeten kenmerken zijn toe- of afgenomen tussen de o- en 1-meting. Daarna worden deze toe- en afnamen vergeleken tussen de groepen.

Zo bleek dat de **voucher** effect had gehad, een bescheiden effect. Maakbedrijven met de service design voucher zijn duidelijk intensiever bezig met diensteninnovatie. Ze bieden ook vaker productgerelateerde diensten aan.”

! Tip van Pim voor collega's

Houd er rekening mee dat er niet altijd draagvlak is in het veld voor dit soort experimenten. De brancheorganisaties vonden het bijvoorbeeld lastig om hun leden te verleiden tot iets, terwijl maar de helft die *treatment* krijgt. Voor de beleidsmaker die snel wil leren van het experiment, is een praktijksetting voor het onderzoek misschien te tijdrovend. De doorlooptijd van ons experiment was anderhalf jaar; ook omdat het een officiële regeling werd, inclusief publicatie in de Staatscourant et cetera.

Onderzoeksmethodiek 18:

Voor- en nameting

Leidt een gedragsinterventie tot het gewenste effect? Dat kun je onderzoeken met een voor- en nameting. Een interventie is daarbij breed te interpreteren. Het kan gaan om een programma om jeugdige delinquenten op het rechte pad te krijgen, een campagne om donorregistraties te stimuleren, een subsidie op de aanschaf van elektrische auto's of om een wet die telefoongebruik in het verkeer verbiedt. De opzet van dergelijk effectonderzoek kan verschillen, maar de essentie blijft zo zuiver mogelijk het effect van de interventie vast te stellen.

Methode: minimaal één voor- en nameting

In effectonderzoek wil je bepalen in hoeverre een gedragsverandering plaatsvindt in de beoogde richting ten opzichte van de situatie voorafgaand aan de inzet van de interventie. Naast gedragseffecten kan het ook gaan om andere doelstellingen, zoals kennis. Hiervoor is minimaal een voormeting en een nameting nodig. Je bekijkt dan de verschillen in resultaten van de voor- en de nameting. Die verandering kan echter ook opgetreden zijn door iets anders dan je interventie, bijvoorbeeld een ontwikkeling in de maatschappij. Om die andere factoren zo veel mogelijk uit te sluiten, meet je idealiter ook vooraf en achteraf bij een controlegroep, een groep die niet aan de interventie wordt blootgesteld (zoals in een RCT). Maar dat is in de praktijk niet altijd mogelijk, bijvoorbeeld bij tv-campagnes. Dan is één voor- en nameting, aangevuld met kwalitatief onderzoek of extra analyses, een goed alternatief.

Toepassing: in fase van evaluatie

Een voor- en nameting zet je in om te evalueren of een interventie het effect heeft gehad dat vooraf werd beoogd.

Soort onderzoek: kwantitatief

De voor- en nameting maakt gebruik van kwantitatieve onderzoeksmethoden. Data kunnen verzameld worden met vragenlijsten, maar er zijn ook registratiedata te gebruiken. Voorbeelden van geregistreerde data zijn het aantal mensen dat op tijd belastingaangifte doet of het aantal automobilisten dat te hard rijdt.

Als aanvulling op de kwantitatieve dataverzameling kun je ook kwalitatief onderzoek doen. Dit kan nodig zijn om resultaten te kunnen duiden. Soms is het wenselijk om ook een kwalitatieve procesevaluatie te houden. Daarin wordt onderzocht of de

interventie wel wordt uitgevoerd zoals bedoeld. Als je bijvoorbeeld een lesprogramma wilt evalueren, moet de docent dat lesprogramma wel op een juiste manier begeleiden. Als de uitvoering afwijkt, kan dat het effect van de interventie verstoren.

Sterke en minder sterke punten van de voor- en nameting

Sterke punten:

- Effectonderzoek via voor- en nameting geeft inzicht in de verandering die je tijdens de looptijd van de interventie bewerkstelligt en in de specifieke doelgroepen waarbij een effect te zien is.
- Je krijgt ook inzicht in de factoren die hebben bijgedragen aan deze effecten.
- Een voor- en nameting is makkelijker te organiseren dan een RCT of quasi-experimenteel onderzoek (zie methodieken 13, 15, 16 en 17).

Minder sterke punten:

- Het causale verband tussen de interventie en het effect is alleen te beredeneren als de invloed van andere factoren zo veel mogelijk is uitgesloten. Lukt dat niet, dan kan het zijn dat je de verkeerde conclusies trekt. Er moet dus voldoende controle zijn op de situatie en het onderzoek. Dit is een kwetsbaarheid van het effectonderzoek. Een paneldesign komt hieraan tegemoet. Meestal wordt bij een nameting een nieuwe groep respondenten ondervraagd. Bij het paneldesign worden de mensen uit de voormeting na de interventie nogmaals ondervraagd. Zo is te analyseren of de groep die is blootgesteld aan de interventie, grotere effecten laat zien dan de groep die de interventie niet heeft gekregen. Dan is het waarschijnlijk dat de effecten

veroorzaakt zijn door de interventie, bijvoorbeeld de publieks-campagne, en niet door externe factoren, zoals publiciteit in de nieuwsmedia.

Praktijkvoorbeelden: effectonderzoeken van DPC en WODC

- **Evaluatie van voorlichtingscampagnes:** Dienst Publiek en Communicatie doet effectonderzoek via een voor- en nameting voor alle voorlichtingscampagnes van de Rijksoverheid met een groot mediabudget (boven de 150.000 euro). Dit gebeurt jaarlijks voor circa vijftien televisiecampagnes en een wisselend aantal campagnes zonder televisiespot. Veel van deze campagnes hebben gedragsdoelstellingen. Het onderzoek laat zien wat ermee is bereikt. Is er iets veranderd in wat mensen doen, bespreken, weten, vinden en voelen? De resultaten worden gebundeld in een jaarevaluatie voor verantwoording aan de Tweede Kamer. Daarnaast worden ze gebruikt om steeds betere campagnes te kunnen ontwikkelen.
- **Effecten van JenV-beleid:** het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) voert regelmatig onderzoek uit naar de effecten van beleidsinterventies van het ministerie van Justitie en Veiligheid. In 2019 is bijvoorbeeld de invloed onderzocht van een speciale afdeling in het Pieter Baan Centrum voor verdachten die weigeren mee te werken aan gedragskundig onderzoek. Zij zijn daarvoor aangemeld omdat vermoed werd dat ze een psychische stoornis hadden toen zij een delict pleegden. Resultaten van het verblijf in de speciale afdeling (nameting/experimentele groep) zijn vergeleken met die van het verblijf in het centrum voor de invoering van de speciale afdeling (voormeting/controlegroep). Na de invoering kon significant meer inzicht in de verdachten worden verkregen, zo bleek. Het centrum kon

vaker zeggen bij wie stoornissen zijn uitgesloten en bij wie de kans op herhaling hoog wordt ingeschat.

Meer informatie

Dienst Publiek en Communicatie (DPC):
Simon Zebregs, s.zebregs@minaz.nl.

“We konden Bob steeds bijsturen”

Koos Tamis

voormalig campagnemanager bij lenW, over effectmetingen

“Zitten we op de goede weg met onze **verkeersveiligheidscampagnes**? Dat was én is een belangrijke vraag bij lenW. Daarom doen we effectonderzoek met voor- en nametingen. Het campagne-effectonderzoek houdt bij hoeveel mensen een tv-spot of andere uiting gezien hebben. Maar meer nog dan om output, gaat het om outcome: treedt er een verandering op in de sociale norm, in de houding, in de gedragsintentie van de weggebruikers die met de campagne in aanraking zijn gekomen?”

Het blijft met enquêtes natuurlijk **zelfgerapporteerd gedrag**: wat mensen zeggen te doen of te gaan doen. Daarnaast worden ook feitelijke waarnemingen gedaan, zoals bij gordelgebruik, fietsverlichting en alcohol in het verkeer. Dat gebeurt in steekproeven langs de weg: dragen mensen een gordel, wat komt er uit blaastests et cetera?

De uitkomsten worden gebruikt voor verantwoording aan de minister en aan de Kamer. En natuurlijk om de volgende campagneflight te kunnen optimaliseren. De Bob-campagne loopt bijvoorbeeld al achttien jaar, en we volgen de effecten dus over een langere termijn. Een voorbeeld van hoe je daarmee de campagne kunt bijsturen? We zagen in het onderzoek een sterke stijging van het aantal mensen die zeiden de Bob te zijn, bij het stappen of naar de sportclub. Dat is natuurlijk mooi. Maar in kwalitatief onderzoek ontdekten we dat mensen die de Bob-rol hadden, toch nog één of twee glazen dronken. Die **pseudo-Bobs** vonden wij een gevaarlijke situatie, want alcohol heeft een go-effect. Na één drankje krijg je trek in meer. Er was daardoor een behoorlijk risico dat mensen toch met te veel alcohol op gingen rijden. In het effectonderzoek hebben we daarover toen vragen toegevoegd en die bevestigden het beeld uit het kwalitatief onderzoek. Niet elke Bob bleek nuchter achter het stuur te zitten.

Vervolgens hebben we in de campagne de inmiddels bekende kreet geïntroduceerd ‘**100% Bob, 0% op**’. Leidde die boodschap tot een afname van het aantal pseudo-Bobs en tot een groter aandeel bestuurders dat aangaf zonder alcohol te rijden? Dat effect werd in de volgende metingen inderdaad zichtbaar. Zo biedt effectonderzoek de mogelijkheid om campagnes steeds te blijven bijsturen.”

! Tip van Koos voor collega's

Als je echt inzicht wilt krijgen in wat je doelgroep beweegt, breid je effectonderzoek dan uit. Doe niet alleen voor- en nametingen, maar vul het aan met kwalitatief onderzoek en het liefst ook met feitelijke waarnemingen als dat lukt. Die combinatie is echt sterk, als je achter gedrag en gedragsfactoren wilt komen.

Meer weten over BIN NL?

Deze methodiekenbundel is een uitgave van alle ministeries gezamenlijk, verenigd in het Behavioural Insights Netwerk Nederland (BIN NL). BIN NL geeft deze uitgave op eigen initiatief uit.

BIN NL is een samenwerkingsverband van alle ministeries voor de toepassing van gedragsinzichten in beleid, uitvoering, toezicht en communicatie. Het netwerk wisselt kennis en ervaring uit door bijvoorbeeld publicaties uit te brengen en jaarlijks de Dag van het gedrag te organiseren. Zie voor meer informatie: binnl.nl.

Werkgroep onderzoeksmethodieken BIN NL:

- Joost Loef (AZ)
- Aletta Boele (EZK)
- Noor Berkman (IenW)
- Kim Herbel (Belastingdienst)
- Hadewych van Kempen (Inspectieraad)
- Maureen Turina-Tumewu (JenV).

Met dank aan iedereen die een methodiek heeft beschreven of op een andere manier bij deze publicatie betrokken is geweest:

- **CBS:** Tommy Span en Magchiel van Meeteren (inmiddels BZK)
- **PBL:** Kees Vringer
- **SCP:** Annemarie Wennekers, Anne Roeters, Pepijn van Houwelingen, Iris Andriessen, Marcel Coenders, Martin Olsthoorn en Crétien van Campen
- **CPB:** Rik Dillingh, Nicole Bosch (inmiddels Belastingdienst)
- **ADR:** Ellen Couwenberg, Marlies Ypma, Roos Kalker en Martine de Zeeuw
- **DPC:** Lizet van Tilburg, Marit Groeneveld, Martijn Warmoeskerken, Simon Zebregs, Margot Lagendijk en Joost Loef
- **Geïnterviewden:** Astrid Paulussen (RWS), Koos Tamis (IenW), Jo van Kruchten (Belastingdienst), Joost Huurman (CBS), Ruth Gonschorrek (ACM) en Pim den Hertog (Dialogic)
- **Oud-werkgroepleden:** Thomas Dirkmaat (EZK), Charlotte Janssen (voormalig Fin), Odette van de Riet (inmiddels Fin), Jan Bade (BZ).

Gedrag veranderen is complex. Onderzoek kan helpen deze kluwen te ontwarren. Waarom gedragen mensen zich zoals ze zich gedragen? Welke interventies zijn kansrijk? En hoe meet je of de gekozen interventie effectief is? Er bestaan veel onderzoeksmethodieken voor. In deze publicatie staan er 18 toegelicht, inclusief toepassingen uit de praktijk van de Rijksoverheid. Van focusgroep tot big data-analyse, van keuze-experiment tot RCT. Een behulpzaam overzicht – ter inspiratie en als wegwijzer – voor wie gedragsonderzoek wil inzetten.

Deze publicatie is een uitgave van

Ministerie van Economische Zaken en Klimaat
Postbus 20401 | 2500 EK | Den Haag

november 2019

Publicatienr. #####