

Evaluatiekader Verdere Integratie op de Arbeidsmarkt

Tussenrapport

Projectnummer P0094

Rapport voor het ministerie van Sociale Zaken en Werkgelegenheid

Auteurs: Lennart de Ruig | Gregor Walz | Wim Zwinkels

© 08 November 2018 | De Beleidsonderzoekers | www.beleidsonderzoekers.nl

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt via druk, fotokopie of op welke andere wijze dan ook, zonder voorafgaande schriftelijke toestemming.

Inhoud

1	Inleiding	5
2	Beleidstheorie VIA	7
2.1	Inleiding	7
2.2	Beschrijving programma VIA	7
2.3	Beleidstheorie	10
3	Evaluatiekader programma-evaluatie VIA	19
3.1	Overwegingen ten aanzien van evaluatieniveau	19
3.2	Onderzoeksvragen	23
3.3	Onderzoeksactiviteiten programma-evaluatie	24
4	Evaluatiekader Nudging Werving en Selectie	25
4.1	Beschrijving pilot	25
4.2	Beleidstheorie	26
4.3	Onderzoeksvragen	31
4.4	Evaluatieniveau	33
4.5	Opzet van de evaluatie	35
4.6	Uitkomstmaten	39
4.7	Planning	39
5	Evaluatiekader Culturele Barometer	41
5.1	Beschrijving pilot	41
5.2	Beleidstheorie	42
5.3	Onderzoeksvragen	46
5.4	Opzet van de evaluatie pilotfase	47
6	Tot slot	49

1 Inleiding

Op 30 maart 2018 heeft de Minister van Sociale Zaken en Werkgelegenheid (SZW) de Tweede Kamer geïnformeerd over het programma 'Verdere Integratie op de Arbeidsmarkt' (VIA). Dit programma is bedoeld om de achterblijvende arbeidsmarktpositie van Nederlanders met een migratieachtergrond aan te pakken. Hiervoor zullen onder meer een aantal pilots worden gestart waarin vernieuwde aanpakken worden uitgetest.

Het ministerie van SZW wil bij de vormgeving van de pilots en het programma VIA als geheel vanaf het begin af aan het monitoren en evalueren ervan centraal stellen. Door te monitoren en te evalueren kan goed in beeld worden gebracht welke pilots (onder welke voorwaarden en in hoeverre) bijdragen aan verbetering van de arbeidsmarktpositie van Nederlanders met een migratieachtergrond en wat de werkzame elementen daarbij zijn. Het streven is op termijn de (succesvolle onderdelen van) pilots breder uit te rollen en om waar mogelijk het succes van het programma als geheel in beeld te brengen.

Bij de uitwerking en vormgeving van het programma en de pilots wordt parallel gewerkt aan een monitorings- en evaluatiekader voor het programma. Dit parallelle traject heeft als doel om vanuit een 'evaluatiebril' te adviseren over de vormgeving van de pilots om deze zo evalueerbaar als mogelijk te maken. Het onafhankelijke bureau De Beleidsonderzoekers heeft opdracht gekregen om het evaluatiekader te ontwikkelen.

Dit rapport doet verslag van het eerste tussenproduct van deze opdracht. Het rapport gaat in op de zogeheten beleidstheorie van het programma VIA (H2) en presenteert de hoofdlijnen van het evaluatiekader voor het programma VIA als geheel (H3). Ook bevatten de eerste twee hoofdstukken van het rapport een raamwerk waarmee de evaluatiekaders voor de specifieke pilots kunnen worden ingevuld. Dit raamwerk wordt vervolgens toegepast op de twee pilots die als eerste zijn uitgewerkt. Dit zijn de pilots 'nudging werving en selectie' (H4) en 'culturele barometer' (H5).

Dit tussenrapport wordt in de loop van 2019 verfijnd en aangevuld met het evaluatiekader voor de andere pilots. Het aldus ontstane eindrapport wordt net als het tussenrapport aangeboden aan de Tweede Kamer.

2 Beleidstheorie VIA

2.1 Inleiding

In dit hoofdstuk beschrijven we de beleidstheorie van het programma Verdere Integratie op de Arbeidsmarkt (VIA). We beginnen met een schets van het programma. Vervolgens introduceren we een aantal begrippen en modellen om de beleidstheorie in kaart te brengen. Daarna passen we deze toe op het programma. In het volgende hoofdstuk gaan we in op de aard en opzet van de evaluatie van het programma.

2.2 Beschrijving programma VIA

2.2.1 Aanleiding en probleemdefinitie

Het aantal werkzoekenden in Nederland daalt al vier jaar gestaag. Het aantal banen neemt toe, in sommige sectoren is zelfs sprake van een hogere vraag dan aanbod aan arbeid. Maar niet iedereen is in staat om van deze situatie te profiteren. Met name de arbeidsmarktpositie van Nederlanders met een migratieachtergrond blijft achter. Zo laat de SCP-publicatie 'Integratie in zicht' uit 2016 zien dat Nederlanders met een niet-westerse migratieachtergrond gemiddeld vaker werkloos zijn dan autochtone Nederlanders, een lagere netto-participatie op de arbeidsmarkt vertonen, op een lager beroepsniveau werken en vaker in een baan met een flexibele arbeidsrelatie werkzaam zijn.¹ Deze verschillen houden ook stand wanneer we rekening houden met demografische en sociaaleconomische achtergrondkenmerken.

Wel is volgens het SCP door de jaren heen enige verbetering te zien in de positie van met name de tweede generatie niet-westerse migranten. Die vooruitgang is echter zeer conjunctuurgevoelig: wanneer de economie in zwaar weer verkeert, gaat de positie van Nederlanders met een migratieachtergrond ook weer sneller erop achteruit. Er is kortom nog geen sprake van een evenredige arbeidsmarktpositie. Ook de onderliggende oorzaken voor deze situatie zijn uitvoerig onderzocht. Zowel factoren die verbonden zijn aan de doelgroep zelf (opleidingsniveau, taalvaardigheid, netwerk, sollicitatiestrategieën, netwerk) als mechanismen op de arbeidsmarkt (onbewuste vooroordelen bij werkgevers, discriminatie) belemmeren op aantoonbare wijze de verdere vooruitgang op dit gebied.

De achterblijvende arbeidsmarktpositie heeft nadelige gevolgen voor de groep zelf, voor de Nederlandse economie en voor de samenleving. Zo leidt de achterstand tot een lager

¹ Huijnk, W., Andriessen, I. (2016), *Integratie in zicht?*, Den Haag: SCP

inkomen, minder mogelijkheden tot zelfontplooiing en frustratie onder de groep zelf. Voor bedrijven en werkgevers betekent de situatie dat het aanbod dat de Nederlandse arbeidsmarkt te bieden heeft niet volledig wordt benut. Juist in tijden van hoge conjunctuur en veel vacatures kan dit een rem zijn op de economische groei en ontwikkeling. Voor de maatschappij als geheel leidt deze situatie tot hogere (en onnodige) kosten voor de sociale zekerheid, en bovendien tot een onwenselijke tweedeling in de samenleving.

Het SCP-onderzoek werd uitgevoerd in het kader van de beleidsdoorlichting van het integratiebeleid over de periode 2000-2015. Op andere terreinen, zoals onderwijs, criminaliteit en maatschappelijke participatie wees het SCP juist op positieve ontwikkelingen. De toenmalige minister van SZW noemt in zijn begeleidende brief van januari 2017 de arbeidsmarkt dan ook terecht "het maatschappelijke veld waar zich de meest zorgwekkende ontwikkeling aftekent".² Het is deze ontwikkeling die de huidige minister van SZW met het VIA-programma wil tegengaan. Naast de inhoudelijke focus die in de naam van het programma ("Verdere integratie op de arbeidsmarkt") duidelijk naar voren komt, legt het programma ook de nadruk op een *evidence-based* aanpak: een werkwijze waarin veel aandacht wordt besteed aan de onderbouwing van maatregelen en het onderzoek naar de effectiviteit van de interventies. Net als de inhoudelijke prioritering bouwt deze beleidstechnische insteek voort op de lessen uit eerdere beleidsprocessen. Dit lichten we hieronder verder toe.

In aanvulling op de beleidsinhoudelijke zorgpunten wierp de beleidsdoorlichting van het integratiebeleid uit 2016 ook methodologische vragen op. Zo bleek het meten van de doeltreffendheid en doelmatigheid van het beleid op veel terreinen (en zeker op het terrein van arbeid) voor de onderzoekers niet haalbaar.³ Onder andere het bestaan van generiek beleid in combinatie met het ontbreken van relevante gegevens over specifieke (migranten)groepen maakte de toerekening van veranderingen in de situatie aan de beleidsinzet nagenoeg onmogelijk. Ook hadden de beleidsactoren bij de ontwikkeling van het beleid te weinig concrete doelstellingen en meetbare verwachte effecten afgesproken waardoor de verandertheorie onvoldoende traceerbaar was. Het is dan ook begrijpelijk dat het ministerie in de reactie op de beleidsdoorlichting in 2017 aanstuurt op een verbetering van de beleidsinformatie op integratiegebied en een sterkere focus op de "wijze waarop de voortgang van het beleid wordt gemonitord en geëvalueerd".⁴

Het integratiebeleid past hiermee in een bredere trend in de manier waarop beleid op rijksniveau (verder) wordt ontwikkeld en verbeterd. Waar beleidsevaluatie en beleidsdoorlichting, mede als gevolg van de Regeling Periodiek Evaluatieonderzoek (RPE) en de

² Minister van SZW (2017), Kamerbrief "Beleidsdoorlichting artikel 13, Integratie en Maatschappelijke Samenhang", 30 januari 2017

³ Regioplan (2016), *Synthesestudie beleidsdoorlichting integratiebeleid*, Amsterdam.

⁴ Minister van SZW (2017), Kamerbrief "Beleidsdoorlichting artikel 13, Integratie en Maatschappelijke Samenhang", 30 januari 2017

Comptabiliteitswet, door de decennia heen een vaste plek aan het einde van de beleids-cyclus hebben gekregen, is er de afgelopen jaren ook al in de ontwikkelfase steeds meer aandacht voor de evalueerbaarheid en impactmeting van beleidsprogramma's. Dit is mede het gevolg van de constatering dat beleidsdoorlichtingen in de huidige situatie vaak tekortschieten bij het meten van de doeltreffendheid en doelmatigheid van het beleid.

In deze context is de operatie Inzicht in Kwaliteit ontwikkeld die in het Regeerakkoord 2017 expliciet genoemd wordt. Met deze operatie stuurt het Ministerie van Financiën rijksbreed aan op een betere onderbouwing van beleidsaannames (oftewel een verbeterde focus op de verandertheorie van beleid) en op meer aandacht voor de evalueerbaarheid van programma's en maatregelen. Ook het daadwerkelijke gebruik van de lessen over doeltreffendheid en doelmatigheid moet centraler komen te staan in de verbetering van beleid.

2.2.2 Aard en doelstelling van het programma

Het programma VIA is een beleidsontwikkelingstraject. Door de uitvoering van pilots wil het ministerie van SZW kennis opdoen over effectieve interventies voor het verbeteren van de arbeidsmarktpositie van Nederlanders met een migratieachtergrond. Het is niet aannemelijk dat het programma tot waarneembare macro-effecten op de arbeidsmarkt zal leiden. Dit komt door de relatief kleine omvang van het programma en de grote invloed van conjunctuur op arbeidsmarktpositie.

De kamerbrief over het programma VIA bevat geen expliciete doelstelling. Ten behoeve van de evaluatie van het programma formuleren we op basis van de kamerbrief de volgende expliciete doelstelling:

Het verkrijgen van inzicht in werkzame elementen van interventies die bijdragen aan het verbeteren van de arbeidsmarktpositie van Nederlanders met een migratieachtergrond en het verkrijgen van inzicht in de opschaalbaarheid van deze interventies.

Het programma resulteert dus vooral in kennis ten behoeve van beleid. Meer specifiek in kennis over:

- wat werkt en wat niet werkt om de arbeidsmarktpositie van Nederlanders met een migratieachtergrond te verbeteren;
- hoe interventies werken om de arbeidsmarktpositie van Nederlanders met een migratieachtergrond te verbeteren (werkzame elementen);
- mogelijkheden om interventies op een grotere schaal te implementeren.

2.2.3 Vormgeving en middelen

Voor het programma is een budget van 10 miljoen euro beschikbaar in de jaren 2018-2020. Daarnaast kunnen belanghebbenden geld beschikbaar stellen voor de uitvoering van de pilots binnen het programma. Alle pilots binnen het programma worden gemonitord en geëvalueerd. Daarnaast vindt een overkoepelende evaluatie (programma-evaluatie) plaats van het programma VIA. Bij het vormgeven en uitvoeren van de pilots werkt het programmateam van het ministerie van SZW samen met belanghebbenden. Daarnaast onderhoudt het programmateam contacten met andere programma's, projecten en beleidsinterventies die inhoudelijk gerelateerd zijn aan het programma VIA.

2.3 Beleidstheorie

Wij gebruiken de volgende begrippen en modellen om de beleidstheorie van het programma in kaart te brengen:

- Logic Model. In dit model wordt een relatie gelegd tussen de middelen, de activiteiten en de resultaten van beleid.
- Het begrip 'werkzame elementen'. Om werkzame elementen te beschrijven, gebruiken we het Britse DEFRA 4E-model. Dit model bevat een raamwerk van vier typen beleidsinstrumenten voor duurzame gedragsverandering.
- Discrepantiemodel. Om werkzame elementen te beschrijven, hanteren we ook het discrepantiemodel. Dit is een eenvoudig model van de arbeidsmarkt en beschrijft typen discrepanties die kunnen optreden tussen vraag en aanbod.

Deze begrippen en modellen gebruiken we voor zover mogelijk ook om de beleidstheorie van de afzonderlijke pilots te beschrijven.

2.3.1 Logic model

Als basisstructuur voor het beschrijven van de beleidstheorie hanteren we het Logic Model.⁵ Dit is een instrument dat door onderzoekers en beleidsmakers wordt gebruikt om de effectiviteit van beleid te optimaliseren en evalueren. Er zijn veel vergelijkbare modellen, maar de basisstructuur is meestal die van het Logic Model.

Een belangrijke keuze bij het Logic Model, wanneer we deze toepassen op het programma VIA, is dat deze in feite uit twee niveaus bestaat: het programmaniveau en het pilotniveau. Het programma is daarbij meer dan een optelling van de verschillende pilots.

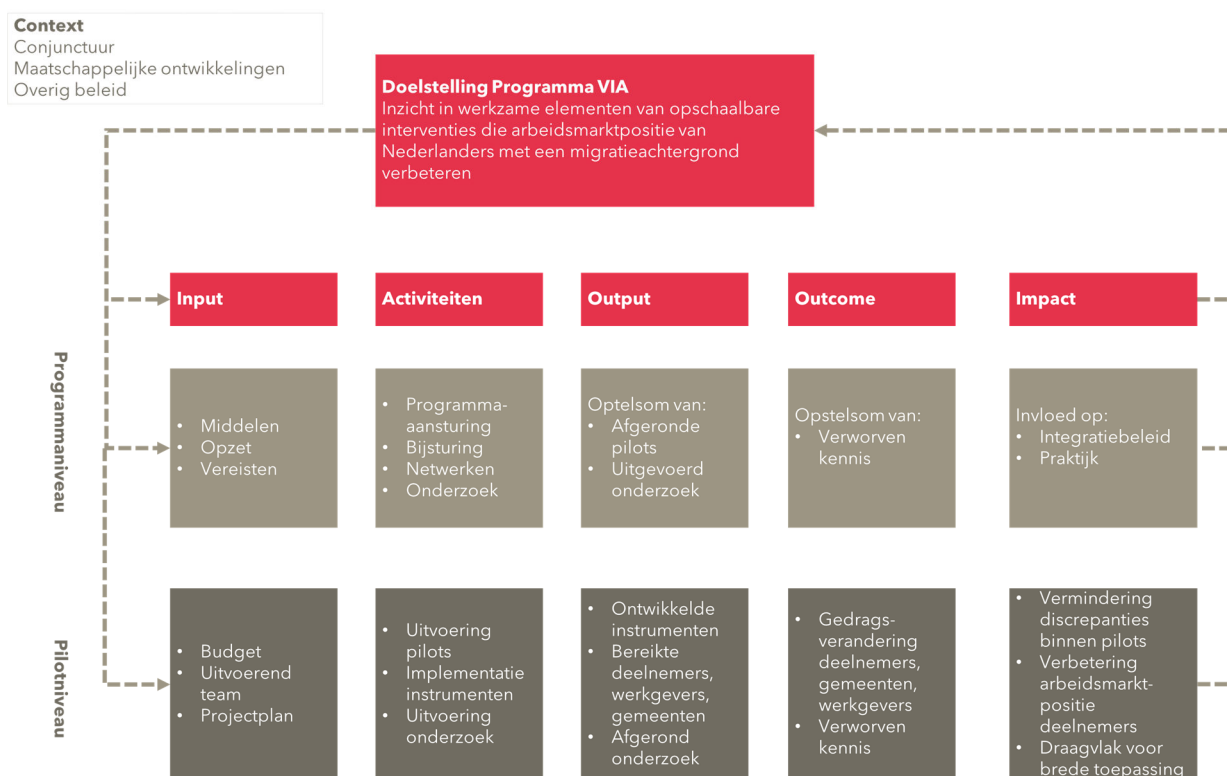
⁵W.K. Kellogg Foundation, Evaluation Handbook, Michigan, 2010. W.K. Kellogg Foundation, Logic Model Development Guide, Michigan, 2004.

Binnen het programma vinden naast de pilots aanvullende activiteiten plaats. En het programma leidt tot eigenstandige effecten.

De volgende onderdelen zijn te onderscheiden (figuur 1.1).

- De **input** heeft betrekking op de beschikbare instrumenten en de materiële en personele middelen (eerste kolom), zowel op het niveau van het programma als de pilots.
- De **throughput** heeft te maken met het uitvoeringsproces, de daadwerkelijke activiteiten binnen het programma en de pilots (kolom 2). De activiteiten hebben tot doel om gedrag te beïnvloeden van partijen of doelgroepen op de arbeidsmarkt.
- De **output** betreft de resultaten van de activiteiten in termen van afgeronde pilots, uitgevoerd onderzoek en bereikte deelnemers (derde kolom).
- De **outcome** betreft de verandering of effecten die de pilots voor de deelnemers teweeg heeft gebracht en de kennis die het programma heeft opgeleverd.
- De **impact** ten slotte betreft de invloed van programma en pilots op de werking van de arbeidsmarkt en op integratiebeleid en -praktijk. Hierbij moet er rekening mee worden gehouden dat de arbeidsmarkteffecten van het programma te klein zullen zijn om op macroniveau te meten. De beoogde impact van het programma en de pilots bestaat enerzijds uit draagvlak bij arbeidsmarktpartijen om de effectieve interventies uit het programma in te zetten. Anderzijds wordt beoogd om het debat over beleid en praktijk van integratie een meer *evidence based* grondslag te geven.
- Er zijn meerdere factoren van invloed op de werking van de arbeidsmarkt dan alleen het programma of de interventie zelf. Rekening moet worden gehouden met de **beleidscontext**, zoals de conjunctuur en andere ontwikkelingen in maatschappij en beleid.

Figuur 1.1 Logic Model



2.3.2 Werkzame elementen van het programma

In de kamerbrief over het programma VIA is te lezen dat ‘de werkzame elementen die opschaalbaar zijn naar landelijk beleid’ worden gedestilleerd uit de monitoring en evaluatie van het programma. Voor een gedegen monitoring en evaluatie is het van belang het begrip ‘werkzame elementen’, ook wel ‘werkzame principes’ of ‘werkzame mechanismen genoemd’, vooraf te definiëren. Wij hanteren de volgende definitie: ‘de elementen van een programma of interventie die er voor zorgen dat die de gewenste uitwerking heeft’.⁶

Werkzame elementen zijn niet alleen achteraf bij een evaluatie te identificeren, maar ook vooraf. Het programma VIA bevat een aantal impliciete en expliciete keuzes en veronderstellingen over de werking. Dit is de essentie van de ‘beleidstheorie’: “het geheel aan veronderstellingen van een actor dat aan een beleid ten grondslag ligt”.⁷ We kunnen hierbij onderscheid maken tussen algemene werkzame elementen en specifieke werkzame elementen. De algemene werkzame elementen gelden altijd, ongeacht doelgroep of de context en beslaan de structuur en methode van het programma en de interventies.

⁶ Wartna, J. et al (2012), “Er is geen enkel werkzaam principe dat altijd werkt”. Een eerste verkenning van het begrip werkzame principes. Wageningen University.

⁷ Hoogerwerf, A. (1989). ‘De beleidstheorie in de beleidspraktijk: een tussenbalans’. In: Beleidswetenschap, nr. 3 p. 320-341.

Om de algemene werkzame elementen te beschrijven, hanteren we een raamwerk dat het Britse departement voor omgeving, voedsel en landbouw in 2008 heeft ontwikkeld voor beleid gericht op duurzame gedragsverandering.⁸ Om te helpen bij het ontwerpen van beleidsprogramma's, biedt het raamwerk een hulpmiddel om een mix van interventies te garanderen. Het raamwerk bestaat uit 4 E's: Enable, Encourage, Exemplify en Engage. Effectief beleid bestaat uit een optimale mix van deze typen instrumenten.

- **Enable** staat voor het mogelijk maken van gewenst gedrag door het verwijderen van barrières en het eenvoudig en vanzelfsprekend maken van bepaalde keuzes. Denk hierbij aan het vastleggen van default-opties, bieden van ondersteuning en voorzieningen, verstrekken van informatie, educatie en training, slim herontwerpen van processen.
- **Encourage** houdt in dat gewenst gedrag op verschillende manieren aangemoedigd wordt. Instrumenten die hierbij horen zijn (financiële) prikkels die gedrag ontmoedigen of aanmoedigen, zoals fiscale voordelen, subsidies, bonussen, malussen of sociale druk.
- **Exemplify** verwijst naar de voorbeeld- en signaalfunctie van de overheid. Het gaat om het delen van goede voorbeelden, consistentie van beleid, 'practise what you preach'.
- **Engage** ten slotte staat voor het betrekken van mensen bij veranderingen. Denk hierbij aan het aansluiten bij of creëren van netwerken, co-creatie, mobilisatie van doelgroepen, werken met intermediairs, bewustmaking.

Wanneer we het 4E-model toepassen op het programma VIA, zien we de volgende impliciete en expliciete keuzes ten aanzien van de werking:

- 1) Door het uitvoeren, monitoren en evalueren van verschillende pilots ontstaat bruikbare kennis over (al dan niet) effectieve interventies. Binnen het raamwerk van het 4E-model kunnen we hierbij spreken over 'Enable': het bieden van informatie om later een juiste keuze te kunnen maken.
- 2) Door bij de uitvoering van het programma samen te werken en verbinding te zoeken met andere programma's, projecten en beleidsinterventies, wordt de kans vergroot dat de inhoudelijk werkzame elementen na afloop van het programma op grotere schaal kunnen worden geïmplementeerd. Er wordt dus aangesloten bij 'energie in de samenleving', waarbij het echter lastig is om de keuzes van andere arbeidsmarktpartijen te sturen omdat er geen sprake is van implementatie van wet- en regelgeving. Binnen het raamwerk van het 4E-model kunnen we spreken over 'Engage': het betrekken van belanghebbenden en het creëren van netwerken.

⁸ Department for Environment, Food and Rural Affairs (2008). A framework for pro-environmental behaviours. Report. January 2008.

- 3) Door tijdens de uitvoering van het programma aandacht te vragen voor kwetsbare migrantengroepen en kennis te delen over effectieve handelswijzen, worden belanghebbenden gestimuleerd aandacht te besteden aan het verbeteren van de arbeidsmarktpositie van Nederlanders met een migratieachtergrond. Ook hierbij kan gesproken worden over 'Engage', maar ook over 'Exemplify': het delen van goede voorbeelden.
- 4) Door de pilots vorm te geven en voor te bereiden met belanghebbenden, zoals werkgevers, gemeenten en de doelgroep, wordt de uitvoerbaarheid van pilots vergroot. Ook hierbij kunnen we spreken over 'Engage'.
- 5) Door de pilots en het onderzoek daarnaar (gedeeltelijk) te financieren als Rijksoverheid, ontstaat een prikkel voor belanghebbenden om deel te nemen en wordt de uitvoerbaarheid van pilots vergroot. In de termen van het 4E-model: encourage.

2.3.3 Werkzame elementen van de pilots in relatie tot het programma

Binnen het programma VIA wordt een aantal pilots uitgevoerd. Deze pilots beogen op een bepaalde manier de werking van de arbeidsmarkt te beïnvloeden. Om de mogelijke werking van de pilots inzichtelijk te maken, maken we gebruik van een vereenvoudigd model van de arbeidsmarkt. In dit model maken we onderscheid tussen de aanbodzijde, de vraagzijde en matching van vraag en aanbod. Tussen vraag en aanbod kunnen knelpunten ontstaan, ook wel discrepanties of mismatches genoemd. Arbeidsmarktbeleid - zoals het programma VIA - is uiteindelijk gericht op het voorkomen en oplossen van knelpunten in de aansluiting tussen vraag en aanbod. Op de arbeidsmarkt zijn globaal gezien drie soorten discrepanties te onderscheiden:

1. Van **kwantitatieve discrepanties** is sprake wanneer er een absoluut tekort is aan vraag of aanbod in bepaalde categorieën. Hiervan is bijvoorbeeld sprake als in een sector, beroepsgroep, of regio te weinig (gediplomeerde) schoolverlaters en andere categorieën werkzoekenden (werkloze werkzoekenden, baanwisselaars uit andere sectoren, en ook medewerkers die binnen de sector van baan willen veranderen) voor de openstaande vacatures zijn dan wel voor deze werkzoekenden te weinig vacatures zijn. In het eerste geval zullen vacatures langer open blijven staan en er naar verhouding weinig werkloze werkzoekenden zijn. Voor werkgevers betekent dat knelpunten in de personeelsvoorziening en voor werkzoekenden goede kansen op een baan. In het tweede geval zullen vacatures sneller vervuld kunnen worden en er relatief veel werkloze werkzoekenden zijn. Dan kennen werkgevers weinig knelpunten in de personeelsvoorziening en werkzoekenden minder goede kansen op een baan.
2. Het voorkomen van **kwalitatieve discrepanties** betekent dat vraag en aanbod er in voldoende mate zijn, maar dat deze niet op elkaar aansluiten. Hiervan is bijvoorbeeld sprake wanneer technisch-instrumentele en/of sociaal-normatieve eisen van de

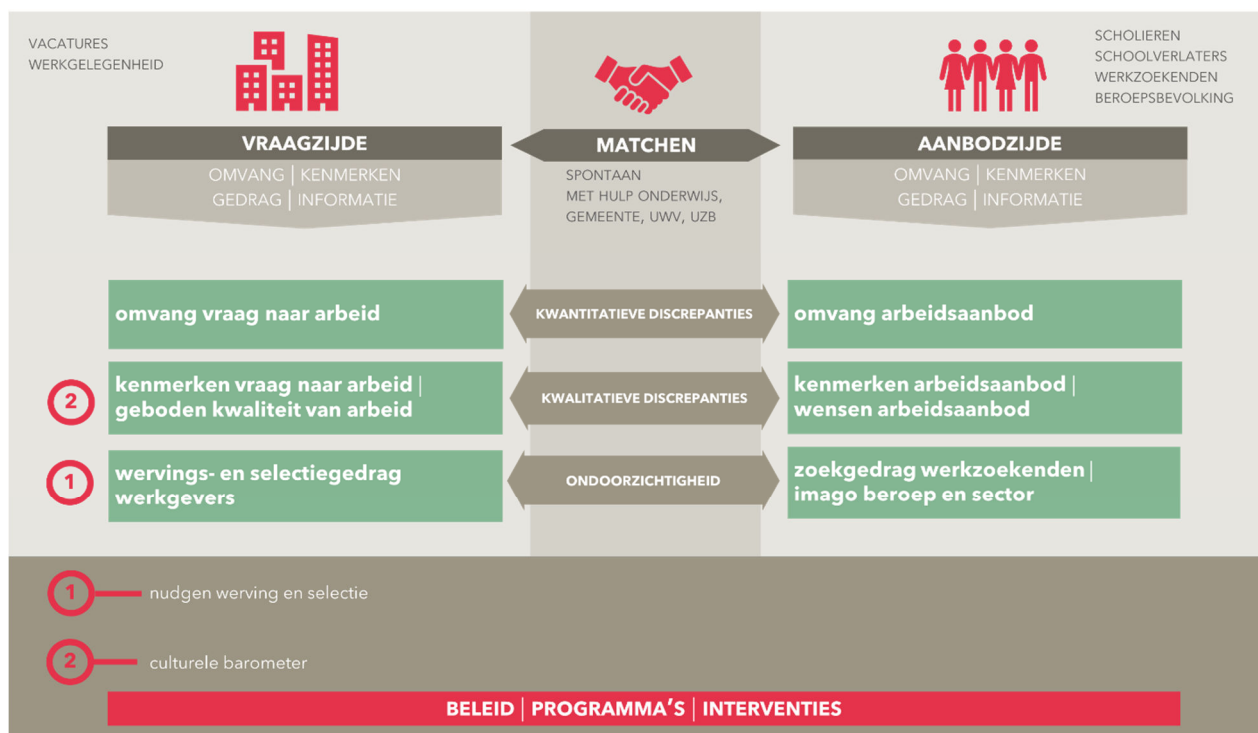
werkgevers hoger zijn dan de kennis, kunde, competenties en/of sociale vaardigheden van (gediplomeerde) schoolverlaters en andere categorieën werkzoekenden. Kwalitatieve discrepanties ontstaan tevens wanneer werkzoekenden hogere eisen stellen aan arbeidsvoorwaarden, -inhoud, -verhoudingen en -omstandigheden dan wat werkgevers binnen de sector willen/ kunnen bieden. Het kan ook zijn dat de kansen op het verwerven van een (passende) baan ongelijk zijn verdeeld en de vervulde vraag geen goede afspiegeling is van de beroepsbevolking.

3. Van **ondoorzichtigheid** (ofwel intransparantie) is sprake wanneer vraag en aanbod er in voldoende mate zijn en ook op elkaar passen, maar elkaar niet kunnen vinden. Dit heeft te maken met wervingsgedrag van werkgevers en zoekgedrag van werkzoekenden. Een beperkte kennis van bijvoorbeeld de beschikbare ondersteunende maatregelen kan zowel bij werkgevers als werkzoekenden hierbij ook een rol spelen.

Discrepanties kunnen ontstaan door zowel feitelijke verschillen tussen vraag en aanbod (in gedrag, aantallen, etc.) als door een gebrek aan juiste informatie over vraag, bemiddeling en/of aanbod (vooroordelen, onbekendheid met de arbeidsmarkt of met werkgeversdienstverlening, etc.).

Op grond van het discrepantiemodel is de (beoogde) werking van de pilots van het programma VIA als geheel te beschrijven. In de hierna volgende paragrafen geven wij bij wijze van invuloefening een eerste aanzet tot de beoogde werking van de pilots. We doen dit voor vraag, aanbod en matching. We karakteriseren de eerste twee uitgewerkte pilots (culturele barometer en nudging werving en selectie) ook kort aan de hand van het 4E-raamwerk. In de loop van 2019 - wanneer er meer pilots zijn uitgewerkt - zal een definitief en volledig ontwikkeld discrepantiemodel worden gepresenteerd. In de volgende figuur staat het voorlopige discrepantiemodel.

Figuur 1.2 Aanzet tot discrepantiemodel voor programma VIA en de eerste twee pilots daarbinnen



De vraagzijde: werkgevers

Betrokkenheid van werkgevers wordt in het programma beschouwd als een basisvoorwaarde voor het verbeteren van de arbeidsmarktpositie van Nederlanders met een migratieachtergrond. De gedachte is dat met name werkgevers in de zogeheten tekortsectoren⁹ vanwege de arbeidsmarktkrapte bereid zijn om werkzoekenden een baan, leerwerkplek of stage te bieden. Ook wordt verwacht dat bedrijven die zich richten op het stimuleren van culturele diversiteit en inclusie open staan voor deelname aan het programma. Deze twee 'typen' werkgevers vormen in beginsel de vraag naar arbeid binnen het programma VIA. We kijken nu naar de verschillende typen discrepanties waarop het programma VIA zich richt.

1: verminderen van ondoorzichtigheid vanuit de vraagzijde

De interventies in het programma zijn op de eerste plaats bedoeld om ondoorzichtigheid van de arbeidsmarkt vanuit de vraagzijde te verminderen. De veronderstelde knelpunten bestaat uit:

- onbekendheid van werkgevers in (o.a.) tekortsectoren met Nederlanders met een migratieachtergrond als werknemer;

⁹ Sectoren waarin relatief veel moeilijk vervulbare vacatures zijn.

- een nadelige invloed van onbewuste vooroordelen bij werving en selectie.

De verwachting is dat meer informatie over het werken met Nederlanders met een migratieachtergrond werkgevers in tekortsectoren kan stimuleren (Exemplify). Daarnaast is de veronderstelling dat het slim vormgeven van HR-processen de invloed van onbewuste vooroordelen bij werving en selectie kan verminderen (Enable). De pilot 'Nudging werving en selectie' is hierop gericht.

2. stimuleren van diversiteit van de vervulde vraag

De relatief slechte arbeidsmarktpositie van Nederlanders met een migratieachtergrond komt onder meer tot uiting in de soms beperkte diversiteit van de vervulde vraag. De hierboven genoemde pilot 'Nudging' moet de baankansen vergroten voor Nederlanders met een migratieachtergrond en leidt zodoende indirect - op lange termijn - tot een grotere diversiteit van de vervulde vraag in de betreffende bedrijven. Ook de pilot 'Culturele barometer' beoogt de diversiteit van de vervulde vraag op een indirecte wijze te stimuleren. Dit gebeurt door een specifieke belemmering weg te nemen: de onmogelijkheid om de diversiteit van de vervulde vraag in beeld te brengen en te vergelijken met de (regionale) beroepsbevolking (Enable).

De aanbodzijde: langdurig werkzoekenden, jongeren en asielstatushouders

De interventies in het programma zijn op de tweede plaats gericht op knelpunten aan de aanbodzijde.

3. verkleinen afstand tot de arbeidsmarkt van het arbeidsaanbod

Verschillende groepen Nederlanders met een migratieachtergrond hebben een relatief grote afstand tot de arbeidsmarkt. Zij ervaren relatief veel en diverse belemmeringen om te participeren, zoals gezondheidsproblemen, schulden, verslaving, laaggeletterdheid/analfabetisme, gebrek aan kennis van de overheidsinstanties, voorzieningen en de Nederlandse arbeidsmarkt.

4. verminderen van ondoorzichtigheid vanuit de aanbodzijde

Een ander knelpunt is de ondoorzichtigheid van de arbeidsmarkt voor scholieren/studenten met een migratieachtergrond en hun ouders. Jongeren kiezen relatief vaak studies met slechte kansen op de arbeidsmarkt. Dit komt doordat zij een beperkt beroepsbeeld hebben. Dit wordt veroorzaakt doordat ouders een belangrijke rol spelen bij de studiekeuze, maar zij weten weinig van het Nederlandse onderwijsstelsel, kampen met een taalachterstand en weten niet hoe zij hun rol moeten vormgeven. Scholen hebben moeite om ouders met een migratieachtergrond te bereiken.

Bemiddeling tussen vraag en aanbod

5. *toeleiden naar leerwerkbanen*

De interventies in het programma VIA zijn ten derde gericht op het knelpunt dat migranten(jongeren) met enige afstand tot de arbeidsmarkt vaak laaggeschoold zijn en enige ondersteuning nodig hebben bij vinden van een baan. Door hen te bemiddelen naar (nieuwe) combinaties van leren werken op het hoogst haalbare niveau, wordt de kans op een match vergroot.

3 Evaluatiekader programma-evaluatie VIA

3.1 Overwegingen ten aanzien van evaluatieniveau

3.1.1 Complexe interventies

Een startpunt bij het bepalen van het evaluatieniveau van het programma VIA en de pilots is de constatering dat het gaat om zogeheten complexe interventies. Complexe interventies bestaan uit meerdere stappen die nodig zijn om de interventie uit te voeren, meerdere werkzame elementen die op elkaar inwerken en meerdere actoren die betrokkenen zijn. Dit in tegenstelling tot eenvoudige interventies zoals de behandeling van patiënten met een pijnstillert, waarbij sprake is van één geïsoleerde interventie die in één stap wordt uitgevoerd (inname medicijn) met één werkzaam mechanisme en een overzichtelijk aantal betrokkenen. Dit heeft consequenties voor de evaluatie en het vinden van werkzame elementen. Pawson et al. (2005) identificeren onder meer de volgende pijlers van complexe interventies¹⁰:

1. Complexe interventies zijn gebaseerd op de hypothese dat wanneer interventies op een bepaalde manier uitgevoerd worden, dit leidt tot bepaalde uitkomsten. De effecten treden dus niet altijd op, maar alleen onder bepaalde omstandigheden en via bepaalde mechanismen. Consequentie voor de evaluatie van het programma VIA en de pilots is dat de effecten altijd gerelateerd moeten worden aan de omstandigheden (context) en werkzame elementen.

2. Complexe interventies hebben een actief karakter, ze werken niet 'als vanzelf', maar behalen hun effecten door middel van actieve input van individuen. Succes en effectiviteit zullen in ieder geval voor een deel verklaard worden door de redeneringen en persoonlijke keuzes van verschillende actoren en deelnemers bij de uitvoering van de interventie. Consequentie voor de evaluatie van het programma VIA en de pilots is dat de interventies niet geïsoleerd bestudeerd moeten worden en dat aandacht nodig is voor de wijze waarop actoren de interventies uitvoeren.¹¹

¹⁰ https://www.researchgate.net/publication/7691447_Realist_review_-_A_new_method_of_systematic_review_designed_for_complex_policy_interventions

¹¹ Interessant is de recente analyse van Hans Bosselaar van de eigenschappen van professionals die bijdragen aan het aan een baan helpen van langdurig werklozen. Zie <https://www.sociale vraagstukken.nl/de-moderne-professionaal-is-vooral-een-dinges/>

3. Complexe interventies zijn langetermijnprojecten. Veel verschillende partijen hebben de opzet en uitwerking in handen. Het succes van een interventie hangt daarom ook af van het cumulatieve succes. Bij de evaluatie van de interventie moet daarom de gehele implementatieketen in ogenschouw worden genomen.

4. Complexe interventies zijn open: andere interventies zullen vaak (succesvolle) onderdelen van de interventie 'lenen' en er zullen leereffecten optreden waardoor de interventies op een andere wijze worden uitgevoerd. Deze vormen van 'besmetting' en het niet modelgetrouw uitvoeren van een interventie zijn binnen een experimentele opzet niet te verkiezen, maar bij complexe interventies onvermijdbaar en zelfs toe te juichen. Consequentie voor de evaluatie van het programma VIA en de pilots is dat er aandacht moet zijn voor de 'lessons learned' bij de implementatie en uitvoering van interventies en de disseminatie van interventies.

3.1.2 Programma-evaluatie

Evaluatie van doeltreffendheid

De evaluatie van het programma VIA is te kenschetsen als een programma-evaluatie. Kenmerkend hiervoor is allereerst dat wordt nagegaan of de doelstellingen van het programma VIA gerealiseerd zijn (doeltreffendheid). Bij het evalueren van het programma komen vragen naar boven zoals: Is het programma volgens planning verlopen? Zijn er wijzigingen opgetreden tijdens de uitvoering van het programma? Zijn de geplande resultaten ook daadwerkelijk gerealiseerd? Hierbij wordt aan de hand van het Logic Model de gehele implementatieketen gevolgd en gerelateerd aan de gestelde doelen. Hiervoor dienen zowel het programma bestudeerd te worden (de ingezette middelen, de uitvoering, etc.) als de evaluaties van de individuele pilots.

Evaluatie van werkzame elementen

Op de tweede plaats brengt de programma-evaluatie de werkzame elementen in kaart, zowel op programmaniveau als op het niveau van interventies (pilots). Hierbij staat de vraag centraal *hoe* en *waardoor* de effecten van de pilots en het programma (al dan niet) gerealiseerd zijn. Ook wordt nagegaan of er bepaalde patronen zijn te ontdekken in de werking van interventies. Interventies zijn bijvoorbeeld te ordenen langs het discrepantie-model, het 4E-model of de uitvoerende actoren (klantmanagers, HR-medewerkers, direct leidinggevenden, etc.).

Tijdens de evaluatie van het programma VIA kan op kwalitatieve wijze worden nagegaan of de beleidstheorie gebaseerd was op de juiste veronderstellingen. Dat wil zeggen: werkte het programma VIA daadwerkelijk via de vooraf geïdentificeerde mix van

instrumenten conform het 4E-model? Is de gekozen mix van instrumenten – in de percepties van betrokkenen – voldoende of ontbreken bepaalde instrumenten? Het antwoord op deze vragen geeft richting aan een advies over de borging van de inzichten uit het programma VIA.

Bij de analyse van werkzame elementen is verder belangrijk om onderscheid te maken tussen de interventies zelf, de context, en de uitvoering van de pilots en het programma. Werkzame elementen kunnen enerzijds betrekking hebben op de inhoudelijke onderdelen van de interventie die ervoor zorgen dat deze de gewenste uitwerking heeft bij de doelgroep¹² en anderzijds op meer praktische aspecten bij de implementatie en uitvoering van de interventie¹³. Ook kunnen werkzame elementen afhankelijk zijn van de specifieke context waarin de interventie wordt uitgevoerd. Denk hierbij aan de regionale werkgelegenheidsstructuur, organisatieomvang, bestuurskracht, et cetera.

Evaluatie van opschaalbaarheid

Op de derde plaats geeft de programma-evaluatie handvatten voor implementatie van de interventies op een grotere schaal. In het programma VIA draait het uiteindelijk om interventies die opschaalbaar zijn. Een pilot zelf is in die zin belangrijk, dat het één manier is om de betreffende interventie te onderzoeken. Als blijkt dat een pilot niet goed is uitgevoerd, zal dat belemmerend werken op de effectiviteit van de interventie. Een niet of minder geslaagde pilot in termen van effectiviteit wil echter nog niet per sé zeggen dat de interventie an sich minder kansrijk is. Andersom kan een pilot zeer goed zijn uitgevoerd maar kan de conclusie zijn dat de betreffende interventie niet kansrijk is. Het is belangrijk dit verschil te onderscheiden, waarbij niet alleen effectiviteit en werkzame elementen een rol spelen, maar ook aspecten als uitvoerbaarheid, draagvlak, legitimiteit, haalbaarheid. Dit betekent dat in de programma-evaluatie ook informatie moet worden opgehaald bij een groter aantal stakeholders en potentiële gebruikers van interventies.

3.1.3 Evaluatieniveau pilots

De programma-evaluatie wordt mede gebaseerd op de evaluaties van de pilots. Op deze plek willen wij enkele algemene aandachtspunten benoemen ten aanzien van het evaluatieniveau van de pilots. Een uitwerking van het evaluatieniveau per pilot komt aan bod in de volgende hoofdstukken.

Een eerste aandachtspunt is dat de pilots in de programma-evaluatie niet eenduidig gerangschikt kunnen worden op een schaal van meer tot minder doeltreffend/effectief. Dat

¹² Denk aan zaken als gebruik van rolmodellen, aansluiten bij de belevingswereld van de doelgroep, eerst plaatsen op werk, dan trainen.

¹³ Denk aan zaken als goede samenwerking tussen betrokkenen, warme overdracht, een goed uitgewerkt stappenplan, voldoende budget.

komt op de eerste plaats omdat de pilots niet allemaal gericht zijn op dezelfde effectindicator. De aard van de effecten van de pilots verschilt sterk. Zo zijn effecten te verwachten in termen van uitstroom uit de uitkering, diversiteit van het personeelsbestand, kansverhoudingen bij werving en selectie, kennis en houding van ouders t.a.v. arbeidsmarkt- en beroepsperspectieven, et cetera. Op de tweede plaats zal het evaluatieniveau van de pilots verschillen. Een hulpmiddel om het evaluatieniveau van interventies te meten, is de Maryland Scale, die hieronder staat weergegeven.

Tabel 2.1 Maryland Scale

Schaalniveau	Omschrijving	Soorten onderzoek
5	Score op de uitkomstvariabele voor en na invoering van de beleidsinterventie, waarbij de beleidsmaatregel 'at random' is toebedeeld aan een experimentele en een controlegroep	<ul style="list-style-type: none"> • Experimenteel onderzoek (RCT)
4	Score op de uitkomstvariabele voor en na invoering van de beleidsinterventie gemeten in een quasi-experimentele setting, met controle voor andere variabelen die de uitkomstvariabele beïnvloeden	<ul style="list-style-type: none"> • Quasi-experimenteel onderzoek (instrumental variables) • Discontinuïteiten benadering
3	Score op de uitkomstvariabele voor en na invoering van de beleidsinterventie, gemeten in een experimentele en vergelijkbare controlegroep	<ul style="list-style-type: none"> • Difference-in-difference • Matching (in combinatie met longitudinaal onderzoek of difference-in-difference) • Longitudinaal onderzoek • Duurmodellen
2	Score op de uitkomstvariabele voor en na invoering van de beleidsinterventie, zonder (vergelijkbare) controlegroep	<ul style="list-style-type: none"> • Matching (cross-sectioneel) • Regressie (cross-sectioneel) • Voor- en nameting
1	Samenhang tussen een beleidsinterventie en uitkomstvariabele na invoering van de beleidsinterventie	<ul style="list-style-type: none"> • Nameting

Bron: <http://www.whatworksgrowth.org/public/files/Scoring-Guide.pdf>

De Maryland Scale wordt binnen het programma VIA gebruikt als hulpmiddel bij de pilotontwikkeling, om bij de vormgeving van de pilots al rekening te houden met eisen die evaluatieonderzoek daaraan stelt, waarbij het doel is om met de via pilots zo hard mogelijk bewijs te verzamelen voor effecten van verschillende instrumenten en aanpakken, maar ze tegelijkertijd praktisch uitvoerbaar te houden. De Maryland Scale is niet te gebruiken om te beoordelen of en hoe pilots opgeschaald kunnen worden. Ook ontbreken in de Maryland Scale kwalitatieve methoden en technieken om werkzame elementen van interventies te analyseren; een belangrijk streven in het programma VIA. In de

evaluatiekaders voor de specifieke pilots besteden we daarom waar nodig en mogelijk aandacht aan aanvullend kwalitatief onderzoek naar opschaling, proces en werkzame elementen.

3.2 Onderzoeksvragen

Voor de programma-evaluatie is de volgende probleemstelling te formuleren:

In hoeverre heeft het programma VIA geleid tot meer kennis van werkzame elementen van interventies, wat zijn deze werkzame elementen en welke leerpunten kunnen worden geformuleerd ten aanzien van opschaling van de interventies?

Hierbij horen de volgende onderzoeksvragen:

1. Input: Welke middelen zijn in de pilots ingezet? In hoeverre zijn de middelen van het programma ingezet?
2. Activiteiten: Hoe is de uitvoering van de pilots verlopen? In hoeverre is het programma conform voornemens uitgevoerd? Waardoor zijn de activiteiten al dan niet conform voornemens uitgevoerd?
3. Output: Wat is binnen de pilots tot stand gebracht in termen van bereikte deelnemers en ontwikkelde instrumenten? Wat heeft het programma tot stand gebracht in termen van afgeronde pilots en uitgevoerd onderzoek?
4. Outcome: Welke gedragsveranderingen zijn gerealiseerd bij deelnemers? Welke kennis over werkzame elementen is verworven bij de uitvoering van de pilots?
5. Impact: Welke langetermijneffecten zijn of worden mogelijk gerealiseerd op het niveau van de pilots en op programmaniveau?
6. Doelrealisatie: In hoeverre zijn de geformuleerde doelen behaald binnen de pilots en binnen het programma? Waardoor zijn doelen wel of niet behaald?
7. Doeltreffendheid: In hoeverre zijn de gevonden effecten bij deelnemers toe te schrijven aan pilots? Zijn de binnen het programma uitgevoerde pilots effectief? Welke overkoepelende conclusies zijn te trekken over werkzame elementen van de pilots en interventies? Wat zijn de werkzame elementen van het programma VIA?
8. Leerpunten: Welke pilots zijn opschaalbaar? Welke leerpunten zijn te formuleren wanneer interventies op grotere schaal geïmplementeerd zouden worden? Wat betekent opschaling voor de vormgeving en uitvoering van de interventies?

3.3 Onderzoeksactiviteiten programma-evaluatie

Voor een programma-evaluatie zijn verschillende bronnen onontbeerlijk, en dienen dus ook meerdere evaluatiemethoden ingezet te worden. Hieronder geven wij een globale opzet van de evaluatie weer. Dit werken we in het eindrapport dat in 2019 volgt nader uit:

Meta-evaluatie. Voor een belangrijk deel is de programma-evaluatie een meta-evaluatie van de evaluaties van de individuele pilots. Hierbij worden alle evaluatierapporten grondig bestudeerd. Door verschillen en overeenkomsten in doelgroepen, actoren, uitvoering, werkzame elementen en effecten te analyseren, kunnen overkoepelende conclusies worden getrokken.

Documentenstudie. In aanvulling op de meta-evaluatie dient documentatie bestudeerd te worden die meer inzicht geeft in de (beoogde) uitvoering van het programma en de pilots. Denk hierbij aan projectplannen, tussentijdse resultaten, financiële verslagen.

Interviews. Interviews op programma-niveau, pilotniveau en op het niveau van stakeholders zijn nodig om de implementatieketen in kaart te brengen, het draagvlak voor bredere implementatie te meten en lessen en leerpunten bij de uitvoering en opschaling van interventies op te tekenen.

Brainstormsessie. Teneinde meer uitgewerkte aanbevelingen te kunnen geven voor de opschaling van interventies, zijn één of meer brainstormsessies met diverse deskundigen raadzaam. Te denken valt aan projectleiders van de pilots, onderzoekers en beleidsmakers op het terrein van integratie.

Enquête(s). Om het draagvlak te meten voor interventies bij potentiële gebruikers, kunnen één of meer enquêtes uitgevoerd worden. Het besluit om enquêtes uit te voeren, dient te worden gebaseerd op een initiële beoordeling van de meest kansrijke interventies. De doelgroepen van de enquête zijn afhankelijk van de interventies en potentiële gebruikers.

4 Evaluatiekader Nudging Werving en Selectie

4.1 Beschrijving pilot

De pilot 'nudging werving en selectie' richt zich op de ontwikkeling van een "set aan kansrijke nudges dat in de praktijk werkt om werkgevers in hun HR werving- en selectieproces de rol van vooroordelen ten opzichte van etnische minderheden te laten verkleinen".

Er is niet één interventie die in de pilot wordt uitgetest, maar een lijst aan instrumenten. De overeenkomst van de instrumenten is dat ze gebaseerd zijn op "nudges", d.w.z. gedragswetenschappelijk onderbouwde ingrepen in het proces die ongemerkt invloed hebben op de manier waarop keuzes worden gemaakt. De instrumenten worden bij een groep van 15-20 werkgevers getest, waarbij werkgevers zelf gaan kiezen welke instrumenten zij in hun bedrijf willen inzetten en laten onderzoeken.

In deze pilot is doorgaans aandacht voor twee centrale aspecten van effectiviteit:

- Netto-effectiviteit van de te ontwikkelen instrumenten op de kansgelijkheid onder sollicitanten met en zonder migratieachtergrond
- Praktische haalbaarheid voor werkgevers voor structurele inlijving in werving & selectieproces en opschaalbaarheid.

De pilot bestaat uit vier onderdelen:

1. Onderzoek naar effectiviteit nudges in wervings- en selectieproces: testen en onderzoek bij bedrijven.
2. Verankering en opschaling: ondersteuning aan werkgevers voor verdere uitrol, verankering en communicatie gericht op andere werkgevers.
3. Vignetstudies (optioneel): aanvullend onderzoek naar werkzame mechanismen.
4. Onderzoek naar algoritmes bij intermediairs (optioneel).

Omdat de conclusies over effectiviteit met name uit het onderzoek op bedrijfsniveau getrokken moeten worden, richt het evaluatiekader zich ook uitsluitend op het eerste onderdeel. Dit betreft de feitelijke implementatie van de instrumenten op bedrijfsniveau en het onderzoek naar de effectiviteit van de instrumenten. De andere onderdelen zijn ondersteunend of aanvullend.

4.2 Beleidstheorie

Om de evaluatie-activiteiten aan te laten sluiten bij de daadwerkelijke opzet van de pilot, moet eerst de beleidstheorie helder in kaart zijn gebracht. Hieronder wordt de beleidstheorie van de interventies op bedrijfsniveau daarom per onderdeel beschreven.

4.2.1 Probleemdefinitie

De pilot 'nudging werving en selectie' vertrekt vanuit de constatering dat onbewuste vooroordelen een rol spelen in het werving en selectiegedrag van werkgevers. Onderzoek laat zien dat deze vooroordelen leiden tot onevenredige kansen van sollicitanten met en zonder migratieachtergrond om uitgenodigd te worden op een sollicitatiegesprek en aangenomen te worden voor een baan. Dit verklaart waarschijnlijk een deel van de achterstandspositie (onverklaard verschil) van de groep werknemers met een migratieachtergrond op de arbeidsmarkt.

Verdere achterliggende aannames (deels gebaseerd op onderzoek) zijn:

- De invloed van onbewuste vooroordelen is het sterkst in de pre-entry fase. D.w.z.: als een persoon eenmaal is aangenomen speelt discriminatie minder een rol.
- Harde doelbewuste uitsluiting komt ook voor maar in mindere mate en wordt aangepakt door middel van handhaving.

4.2.2 Doelgroepen

De doelgroep van de pilot zijn werkgevers in het algemeen, en HR-professionals in het bijzonder. Zij zullen met de te ontwikkelen instrumenten werken.

De profijtgroep van de pilot zijn sollicitanten met een migratieachtergrond. De pilot richt zich met name op de beroepsniveaus hbo en mbo niveau 3/4. Er is geen specifieke keuze gemaakt voor doelsectoren.

4.2.3 Hoofddoelstelling

In het Plan van Aanpak van de pilot zijn de doelstellingen als volgt omschreven:

Op de lange termijn wil het ministerie stimuleren dat alle werkgevers in Nederland met een HR-afdeling relatief eenvoudig en aantoonbaar effectief hun werving- en selectieproces (kunnen) aanpassen, zodanig dat onbewuste vooroordelen minder een rol spelen.

In de pilot gaan we met een voorhoede van 15-20 werkgevers het pad effenen. De doelen van de pilot zijn:

A) In 2021 weten we van een set aan kansrijke nudges wat in de praktijk werkt om werkgevers in hun HR werving en selectieproces de rol van vooroordelen ten opzichte van etnische minderheden te laten verkleinen:

- a. Welke interventies toepasbaar zijn voor werkgevers om binnen hun reguliere HR-processen
- b. Welke interventies goede uitkomsten hebben in termen van het verminderen van vooroordelen in het selectieproces / het aannemen van een etnisch diverser personeelsbestand

B) Tussen 2019 en 2021 is er een groeiende groep werkgevers die positief staat tegenover de interventies en ze begint toe te passen

- a. Pilotdeelnemers houden vast aan wat werkt en blijven dat in hun eigen organisatie toepassen
- b. Pilotdeelnemers nemen een voorbeeldrol en propageren wat werkt bij andere werkgevers en werkgeverskoepels/organisaties in verschillende regio's en sectoren

C) In 2021 zijn de voorwaarden geschapen waardoor de opgedane kennis op langere termijn handzaam beschikbaar blijft voor een brede groep werkgevers

- a. Er is een eigenaar gevonden buiten het ministerie, zodat trainingsbureaus en werkgeversorganisaties, zowel breed landelijk (AWVN of MKB NL), sectoraal (zorg en/of techniek) en regionaal (WSP's), er een beroep op kunnen doen
- b. De kennis over theorie en werkzame mechanismen die we in en buiten de pilots hebben opgedaan is toegankelijk in een toolkit, voor partijen die andere interventies willen ontwikkelen (zoals onderzoeksbureaus en werkgevers).

Doelstelling A) heeft betrekking op de feitelijke interventies die in het kader van de pilot worden uitgetest. Doelstellingen B) en C) hebben betrekking op de flankerende activiteiten en het proces dat vanuit het ministerie wordt aangestuurd en aangejaagd. Omdat de effectiviteitsvraag in de kern onder doelstelling A) valt, is dit evaluatiekader toegespitst op deze doelstelling. Doelstelling B) komt zijdelings aan de orde, omdat deze doelstelling aansluit bij de vraag naar de langere-termijn impact van het programma VIA.

4.2.4 Relatie met VIA

De doelstellingen van de pilot dragen indirect bij aan de doelstelling van het programma VIA. Door het verminderen van de invloed van vooroordelen in het werving en selectieproces worden evenredige kansen bij sollicitatie bevorderd. Hierdoor wordt een drempel bij de toegang tot de arbeidsmarkt van mensen met een migratieachtergrond weggenomen. Vervolgens kan dit leiden tot een betere instroom van deze groep werknemers bij bedrijven. In arbeidsmarktterminologie richt de pilot zich deels op de matching/bemiddeling tussen vraag- en aanbodkant en deels op het bevorderen van transparantie vanuit de vraagkant (werkgevers).

De inzichten die in deze pilot worden opgedaan, zijn ook relevant voor andere pilots. De invloed van onbewuste vooroordelen is van belang voor pilots gericht op de aanbodkant, met name waar het gaat om sollicitatiestrategieën van werkzoekenden met een migratieachtergrond. Dit komt in ieder geval terug in de pilot met betrekking tot combinaties van leren en werken. Verder heeft de pilot duidelijke raakvlakken met activiteiten die in het kader van het Actieplan Arbeidsmarktdiscriminatie worden genomen (directie Arbeidsverhoudingen).

4.2.5 Potentieel werkzame mechanismen, beoogde werking

Alle werkzame mechanismen hebben betrekken op het verminderen van de invloed van vooroordelen op de beslissingen van selecteurs in w&s-procedures. Het gaat hierbij expliciet niet om bewustwording bij selecteurs (bijv. door middel van trainingen), maar om het aanpassen van het proces op een manier die de ruimte beperkt voor een vertekende invloed van onbewuste vooroordelen. De beoogde werking van de instrumenten kan omschreven worden als:

1. Objectiveren van het sollicitatieproces (werving, selectie, gesprek en beslissing) en hierdoor minder ruimte geven aan de invloed van vooroordelen; bijv. instrument 'neutraal maken van wervingsmaterialen' (gericht op werving), 'prominentie van objectieve functie-eisen vergroten' (brieveselectie), '15 minuten voor aanvang voorbereiden' (gesprekken), 'gestructureerde interviews', 'groter gewicht van try-outs en tests' (besluit), 'onafhankelijke derde', 'besluiten in een panel'.
2. Doorbreken van in-group / out-group dynamiek, en hierdoor veronderstelde groepskenmerken minder de kans geven om meegewogen te worden in de selectiebeslissing. Bijvoorbeeld instrument 'video-sollicitaties' (individualiseren van sollicitaties), 'opvragen van referenties' (betrekken bij de in-group). 'Groter gewicht van try-outs en tests' kan hier ook onder vallen.

3. Aansporen of incentives bieden om diverse instroom te bevorderen (expliciete focus op diversiteit). Bijvoorbeeld 'doelen in herinnering brengen'.

Op pilotniveau bestaan er, naast de werkzame inhoud van de maatregelen zelf, een aantal bijkomende mechanismen die onderzocht kunnen worden op doeltreffendheid, met name in de communicatieve sfeer. Dat zijn in ieder geval:

- 1 Creëren van draagvlak onder bedrijven door middel van delen van probleem-eigenaarschap: door werkgevers en HR-professionals vroeg te betrekken en zelf een keuze te laten maken ten aanzien van de te ontwikkelen tools, wordt het draagvlak onder en binnen de bedrijven vergroot.
- 2 Creëren van draagvlak door middel van positieve bedrijfsmatige benadering: door het framen van de pilots in termen van effectievere werving en selectie in plaats van discriminatiebestrijding wordt het draagvlak onder werkgevers versterkt.

4.2.6 Beleidslogica van de pilot

We kunnen nu de beleidslogica definiëren volgens het Logic Model. Om het model werkbaar te houden, wordt hieronder de interventielogica opgesteld voor het onderdeel van de experimenten op bedrijfsniveau. De flankerende activiteiten met betrekking tot communicatie, opschaling, en ondersteunend onderzoek zijn hierin niet meegenomen.

Input

Zowel SZW als de deelnemende bedrijven investeren in de uitvoering van de experimenten:

- Budget SZW voor onderzoek, uitvoering, begeleiding werkgevers
- Budget werkgevers voor uitvoering

Throughput

Vanuit SZW en de bedrijven worden de volgende hoofdactiviteiten uitgevoerd:

- Werving bedrijven
- Aanscherping/standaardisering instrumenten
- Advisering en ondersteuning voor bedrijven
- Implementatie instrumenten bij bedrijven

Output

De output van de experimenten is zowel op pilot- als op bedrijfsniveau te definiëren:

- 10-20 werkgevers betrokken en actief;

- 10 instrumenten uitgewerkt en gestandaardiseerd
- Ondersteuning voor werkgevers gerealiseerd
- Minimaal 100 vacatures ingevuld per instrument
- Minimaal 100 vacatures ingevuld per controlegroep

Outcome

Bij succesvolle implementatie van de instrumenten op bedrijfsniveau wordt de volgende outcome verwacht:

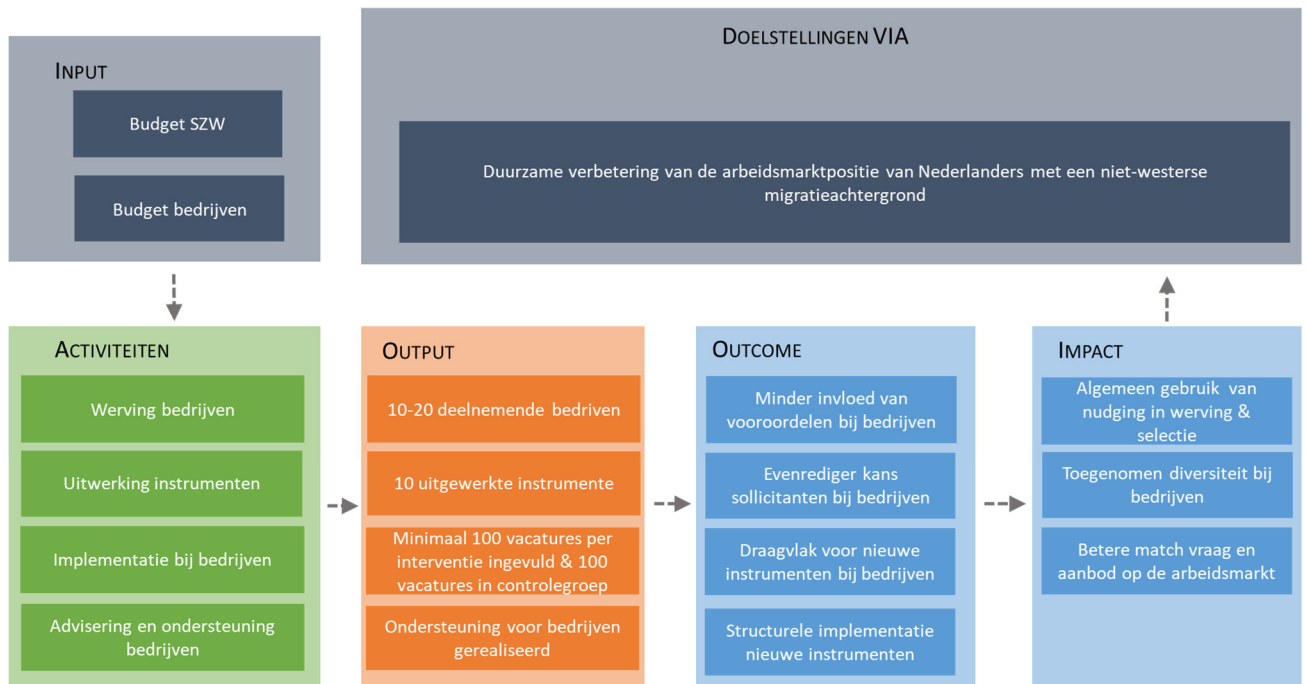
- Verminderde invloed van vooroordelen op de werving en selectie bij deelnemende bedrijven
- Evenrediger kans voor mensen met en zonder migratieachtergrond bij sollicitaties bij deelnemende bedrijven
- Draagvlak bij bedrijven voor nieuwe instrumenten
- Structurele verankering nieuwe instrumenten bij deelnemende bedrijven

Impact

Op de langere termijn heeft deze pilot de volgende impact:

- Toegenomen diversiteit bij deelnemende bedrijven
- Gebruik van nudging in werving en selectie bij Nederlandse bedrijven in het algemeen
- Betere matching vraag en aanbod op de Nederlandse arbeidsmarkt

Hieronder is deze beleidslogica in grafische vorm weergegeven.



4.3 Onderzoeksvragen

Per onderdeel uit de evaluatiematrix kunnen we de onderzoeksvragen specificeren. Voor zover mogelijk zijn deze onderzoeksvragen geformuleerd op het niveau van deelnemende bedrijven:

- 1 Input: Hoe zijn de begrote middelen voor de pilot ingezet?
 - Wat is de feitelijke inzet aan middelen (tijd en geld) voor gebruik van de interventie? Zowel absoluut als relatief ten opzichte van de reguliere sollicitatieprocedure?
 - Wat is de ervaren inzet aan tijd en energie door recruiters bij gebruik van de interventies?
- 2 Throughput: Hoe is de uitvoering van de experimenten verlopen? Waardoor zijn de experimenten al dan niet conform voornemens uitgevoerd?
 - Hoe verliep de werving van deelnemende bedrijven door SZW? In welke mate is daarbij voldaan aan de eis van een aselechte steekproef van deelnemers aan de pilot?
 - Wat was de uitgangspositie van bedrijven? Hoe was het reguliere werving en selectieproces vormgegeven? Wat gaat het bedrijf anders doen? Welke interventies worden ingezet en op welke wijze?

- Hoe verliep de implementatie van de verschillende instrumenten bij bedrijven? Ten aanzien van: uitwerking instrumenten, communicatie intern en extern, aanpassing structuren, daadwerkelijke uitvoering, betrokken actoren. In hoeverre zijn de instrumenten consequent en volgens uitgangspunten ingezet bij de bedrijven?
 - Welke contextfactoren waren van invloed op de invoering van de instrumenten bij bedrijven?
 - Heeft het bedrijf een actief diversiteitsbeleid waarbij gebruik wordt gemaakt van bijvoorbeeld voorkeursbeleid, targets, trainingen voor bewustwording?
 - Is er sprake van een krappe of ruime arbeidsmarkt voor de vacatures? Is dit veranderd gedurende de pilotperiode?
 - Andere ontwikkelingen (voorzien en onvoorzien) binnen of buiten het bedrijf die van invloed zijn op de organisatie.
- 3 Output: Wat is tot stand gekomen in termen van producten, diensten en activiteiten?
- Hoe veel bedrijven hebben deelgenomen aan de experimenten en met welke instrumenten? Hoe veel en welke instrumenten zijn uitgewerkt voor implementatie?
 - Hoe veel vacatures zijn ingevuld in experiment- en controlegroepen? Hoeveel sollicitanten waren betrokken bij de verschillende fases van het sollicitatieproces?
 - Wat was de verdeling tussen sollicitanten met en sollicitanten zonder migratieachtergrond in de verschillende fases van het sollicitatieproces?
 - Van welke ondersteuning hebben bedrijven gebruik gemaakt?
- 4 Outcome: Welke directe effecten zijn gerealiseerd?
- In hoeverre is de kansengelijkheid tussen sollicitanten met en zonder migratieachtergrond toegenomen door invoering van de verschillende instrumenten?
 - Welke definitie van kansengelijkheid is van toepassing op de geïmplementeerde instrumenten? Was er eerder sprake van kansenongelijkheid bij deelnemende bedrijven?
 - In welke mate is de relatieve kans een ronde verder te komen toegenomen voor kandidaten met een niet-westerse achtergrond?
 - In welke mate is de relatieve kans om uiteindelijk aangenomen te worden toegenomen voor kandidaten met een niet-westerse achtergrond?
 - Hoe ervaren de recruiters de meerwaarde van de interventies bij het beoordelen van de potentiële arbeidsproductiviteit van kandidaten? Wat waren de

ervaringen van selecteurs, HR-medewerkers, selectiecommissies? Is de interventie makkelijk of juist moeilijk toepasbaar gebleken?

- Wordt inzet van de interventies na afronding van de pilot voortgezet en opgeschaald naar andere onderdelen? Wat zijn de overwegingen daarbij? Wat zijn de kosten of besparingen bij structurele invoering?

5 Impact: Welke (langetermijn) effecten heeft dit tot gevolg gehad? N.B. Deze vragen zijn niet noodzakelijk voor het directe effectiviteitsonderzoek naar de interventies en zijn pas op de langere termijn te beantwoorden. Daarom worden ze niet verder uitgewerkt in dit evaluatiekader en niet opgenomen in paragraaf 1.5.

- Leidt het structurele gebruik van de instrumenten tot een grotere diversiteit (t.a.v. migratieachtergrond) van het werknemersbestand van de bedrijven?
- Bestaat er belangstelling bij andere bedrijven voor gebruik van de ontwikkelde instrumenten?

4.4 Evaluatieniveau

Voor een ideale evaluatie van de instrumenten (op een zo hoog mogelijk evaluatieniveau) zouden de instrumenten arbitrair (*random*) moeten worden toegepast op een groot aantal bedrijven met uiteenlopende kenmerken. Er zijn in de vormgeving van de pilot echter een aantal keuzes gemaakt die in het licht van de praktische haalbaarheid van de pilot van belang zijn. Hierbij is rekening gehouden met de eis van evalueerbaarheid, en is gepoogd een balans tussen evalueerbaarheid en haalbaarheid in acht te nemen. Dat heeft betrekking op de volgende aspecten:

- Aantal vacatures: het beoogde aantal vacatures per instrument is 100 (zowel voor de controle- als de experimentele groep). Dit aantal is vergelijkbaar met soortgelijke onderzoeken in Nederland en in andere landen. Het aantal sollicitanten dat deelneemt aan de procedures zal beduidend hoger liggen, maar is niet exact te voorspellen. De grootte van de te verwachten effecten is onduidelijk. Dit betekent dat met name bij de interventies en analyses die gericht zijn op de kansen van individuele sollicitanten in de eerste selectiefase de omvang groot genoeg is om ook kleinere statistisch significante effecten vast te stellen. Waar de analyse op het niveau van de vacature plaatsvindt (dus niet op niveau van individuele sollicitant), en bij interventies die gericht zijn op latere fases van het selectieproces (waarbij het aantal sollicitanten dat overblijft steeds beperkter wordt) wordt de kans groter dat effecten van dezelfde omvang bij eenzelfde significantieniveau niet significant zijn en dus niet onderscheiden kunnen worden van toeval. Dit betekent dat voor

de instrumenten in de latere fases van de procedures het aantal vacatures mogelijk niet groot genoeg is om de effectiviteit goed te kunnen beoordelen. Om dit te ondervangen, wordt voor de instrumenten die zich richten op het gesprek en de beslissing aanvullend vignettenonderzoek uitgevoerd. Goede afstemming tussen het evaluatieonderzoek en het vignettenonderzoek is dus van belang.

- Aantal bedrijven: het aantal deelnemende bedrijven is in eerste instantie 10 en uiteindelijk 20. Een hoger aantal is niet haalbaar gebleken voor een spoedig begin van de pilot.
- Kenmerken bedrijven: er wordt gestreefd naar een variatie tussen publieke en private werkgevers. Niet haalbaar gebleken is het vinden van bedrijven die geen actief diversiteitsbeleid voeren. Dit wordt ondervangen door de analyse van de beginsituatie bij bedrijven en de afspraak met bedrijven om geen wijzigingen aan te brengen in het gevoerde beleid gedurende de experimentele fase.
- Keuze van instrumenten: de bedrijven maken zelf een keuze welke instrumenten zij willen implementeren (niet-arbitrair). Ook combinaties van instrumenten worden toegestaan, mits deze geen tegenstrijdige mechanismen in werking zetten en van toepassing zijn op verschillende fases van het selectieproces (d.w.z. maximaal een instrument per fase). Een toedeling van instrumenten 'at random' zou het evaluatieniveau kunnen verhogen, maar is niet haalbaar gezien het belang om bedrijven te werven voor vrijwillige deelname.
- Codering op basis van namen: om het onderzoek haalbaar te maken en de privacy van sollicitanten te beschermen, is gekozen om de codering van sollicitanten vooral te baseren op naam, gender en leeftijd, en niet te corrigeren voor wisselende kwaliteit van CV's die in theorie de uitkomsten in controle- en experimentele groep zouden kunnen beïnvloeden. Om deze correctie mogelijk te maken, zouden bedrijven de CV's van sollicitanten moeten bewaren en zouden de onderzoekers de CV's op inhoud moeten beoordelen. Dit is niet haalbaar vanwege zowel de privacy-gevoeligheid van de gegevens als de grote tijdsinvestering die dit zou vergen van de onderzoekers.
- Gegevensbescherming: met de deelnemende bedrijven worden afspraken gemaakt over de dataverzameling. Deze afspraken zullen voldoen aan de AVG.

Uiteindelijk betekent dit dat bij goede uitvoering en door de quasi-experimentele opzet evaluatieniveau 3 op de Maryland Scale haalbaar is. Een hoger niveau kan alleen bereikt worden door een substantiële uitbreiding van de onderzoekspopulatie (aantal bedrijven) en een gerandomiseerde toewijzing van de instrumenten op de bedrijven.

4.5 Opzet van de evaluatie

De centrale effectiviteitsvraag van deze pilot zal op het niveau van de bedrijfsexperimenten beantwoord worden. Daarnaast moet op pilotniveau overkoepelende monitoring uitgevoerd worden om de bereikte opschaling en mogelijkheden voor verdere uitrol te kwantificeren. Het hieronder uitgewerkte evaluatiekader gaat in op deze twee aspecten, en laat de aanvullende onderzoeksactiviteiten (vignetstudies en algoritmes) buiten beschouwing.

De evaluatie en monitoring van de instrumenten moet leiden tot betrouwbare kennis over de effectiviteit van elk instrument. Daarbij zijn twee aspecten van belang:

- 1 De effectiviteit van de ontwikkelde maatregelen op zichzelf;
- 2 De toepasbaarheid van de maatregelen in de w&s-praktijk van bedrijven.

De ambitie van de evaluatie van VIA is om zo hoog mogelijk op de Maryland Scale te komen. Voor deze pilot is op instrumentniveau een quasi-experimentele opzet mogelijk. Dit is ook praktisch haalbaar. De deelnemende bedrijven committeren zich immers aan de doelstellingen van het onderzoek. Voor het vaststellen van de werkzame mechanismen en het beantwoorden van de vraag naar de praktische uitvoering van de maatregelen is een vorm van procesevaluatie onmisbaar.

Hoewel de instrumenten verschillen, is de basisopzet van het evaluatieonderzoek op bedrijfsniveau dezelfde, ongeacht de ingezette instrumenten. Deze opzet gaat uit van een quasi-experimenteel design met parallel lopende controle- en experimentele groepen zodat een *difference-in-difference* benadering mogelijk is. Dit betekent dat bedrijven in staat moeten zijn om gegevens bij te houden over vacatures en sollicitanten in verschillende fasen van het selectieproces en hun achtergronden (met name wel/geen migratieachtergrond), in zowel de controlegroep als de experimentele groep. Hiervoor bestaan de volgende mogelijkheden:

- De controlegroep bestaat uit sollicitanten/vacatures in een ander, voorafgaand tijdsbestek waarin de maatregel (nog) niet werd toegepast.
- De controlegroep bestaat uit sollicitanten/vacatures in een andere afdeling waarin de maatregel niet wordt toegepast.
- De maatregel wordt willekeurig op een deel van de vacatures/sollicitanten toegepast en op een deel niet.

Een *difference-in-difference* benadering vereist dat controlegroepen op verschillende momenten met elkaar gecombineerd kunnen worden. Dat betekent dat het mogelijk moet zijn om sollicitaties uit de periode voorafgaand aan de testperiode te coderen en in

te delen naar afdeling (of een andere indeling afhankelijk van het bedrijf) en in de testperiode zelf het instrument ook in een deel van de afdelingen toe te passen en in een ander deel niet. De deelnemende bedrijven moeten daarom zowel voorafgaand aan als gedurende de pilotperiode alle relevante gegevens verzamelen en beschikbaar te stellen aan de onderzoekers. Dit vereist extra registratiewerkzaamheden en het openstellen van processen voor externe onderzoekers. Ook moeten sollicitanten geïnformeerd worden over het gebruik van hun gegevens voor extern onderzoek, in verband met hun rechten op gegevensbescherming.

Per bedrijf worden de volgende onderzoeksactiviteiten uitgevoerd:

- **Vooronderzoek bij bedrijven:** voorafgaand aan het experiment, inventarisatie van bestaande HR-procedures, met specifieke aandacht voor reeds bestaand diversiteitsbeleid en mogelijke geplande veranderingen in de komende periode; probleemanalyse t.a.v. diversiteit, vooroordelen, afspiegeling verschillende groepen op diverse functieniveaus; afspraken over controlegroepen en codering van sollicitanten, met het uitgangspunt dat bedrijven minimaal namenlijsten van sollicitanten beschikbaar stellen en idealiter de onderzoekers de CV's laten inzien, voor alle relevante vacatures en de verschillende fases van de procedure.
- **Codering sollicitanten op basis van namen (0-meting):** in controle- en experimentele groep voor een tijdsbestek van een half jaar voorafgaand aan het experiment. De codering wordt door de onderzoekers op basis van voor- en achternaam gedaan - uitgangspunt is de perceptie van achtergrondkenmerken (niet de feitelijke indicatoren zoals geboorteland ouders). Handmatige codering is hierdoor noodzakelijk, gebaseerd op eenduidige instructies. Controle gebeurt door middel van cross-checks door verschillende onderzoekers (steekproefsgewijs bij minimaal 20% van de sollicitanten). De codering gebeurt voor de categorieën niet-Westerse migratieachtergrond en geen niet-Westerse migratieachtergrond en voor de vier subcategorieën Turks, Marokkaans, Surinaams, Antilliaans. Daarnaast vindt codering op gender plaats, op basis van de naam van de sollicitant. Optioneel en waar mogelijk op basis van de beschikbare informatie (afhankelijk van mogelijkheid om inzage in CV te krijgen) kan ook leeftijd meegenomen worden om eventuele neveneffecten in de meting mee te nemen, en eventueel andere herkomstgroepen.
- **Codering sollicitanten op basis van namen (1, 2-meting):** in controle- en experimentele groep gedurende de looptijd van het experiment. Minimaal 2 meetmomenten (T1 en T2). Eenzelfde aanpak als bij 0-meting.
- **Kwantitatieve analyse effecten:** toegepast op verschillende sollicitantengroepen en verschillende fases van het sollicitatieproces. Per sollicitantengroep wordt

de (absolute en relatieve) omvang van de groep per sollicitatiefase berekend, en hiervan afgeleid de kans op succes per ronde, en de verhouding van de kansen van de verschillende geïdentificeerde groepen. In de analyses wordt de individuele kans en de kansverhouding van de sollicitantengroep met migratieachtergrond (en subgroepen) ten opzichte van de sollicitantengroep zonder migratieachtergrond berekend en vergeleken voor T0, T1, T2. Voor de difference-in-difference analyse wordt de verandering van T0, T1, naar T2 tussen de controlegroep en de experimentele groep vergeleken om te corrigeren voor een veranderende context. Er zijn dus verschillende analyses vereist:

- analyse van de verandering in de kans op uitnodiging voor een volgende selectieronde op het niveau van sollicitanten (waarbij rekening gehouden dient te worden met de verschillende contexten van vacatures en bedrijven);
- analyse van de verandering in de relatieve kans op succes van verschillende sollicitantengroepen op het niveau van de vacature;
- statische analyse van de kansverhouding op het niveau van de vacature.

Gezien de eventueel te lage aantallen beschikbare observaties is het ons advies om de hypothesen eenzijdig te toetsen. Er wordt dan alleen getoetst of de interventie de kansen van migrantengroepen in het sollicitatieproces doet verbeteren (niet verslechteren).

- **Interviews met betrokkenen binnen het bedrijf over uitvoering:**

- Interviews met HR-verantwoordelijke en betrokken afdelingsmanagers/directeur - voorafgaand en na afloop van de experimentperiode. Onderwerpen: overwegingen deelname, daadwerkelijke implementatie, bevorderende en belemmerende factoren, perceptie van effectiviteit, lange-termijn perspectief. Richtlijn: minimaal 2 respondenten per bedrijf, persoonlijke interviews.
- Interviews met een selectie van betrokken medewerkers of leden van selectiecommissies, gedurende of na afloop van de experimentperiode, in experimentele groep. Onderwerpen: visie op instrument, ervaring implementatie, perceptie effectiviteit, tevredenheid resultaat, draagvlak. Richtlijn: minimaal 5 respondenten per bedrijf, persoonlijke interviews, optioneel aanvullende telefonische interviews of groepsgesprekken.
- Afhankelijk van de geïmplementeerde instrumenten, moet dit aangevuld worden met observaties die zowel dienen ter controle van de daadwerkelijke uitvoering als om inzicht in praktische dilemma's en mechanismen te verkrijgen. Dit is relevant voor de interventies die zich richten op de latere fases van het sollicitatieproces (gesprek en besluitvorming).

- **Analyse per instrument en overkoepelend:** de kwantitatieve en de kwalitatieve opbrengsten van het onderzoek worden in de analyse bij elkaar gebracht om uitspraken te doen over de relatie tussen effecten en uitvoering, verklaringen voor geconstateerde effecten, en interpretatie van resultaten. Dit vindt zowel op bedrijfs- als op instrumentniveau plaats. Overkoepelend worden de instrumenten met elkaar vergeleken aan de hand van verschillende kenmerken, zoals netto-effectiviteit, toepasbaarheid, intensiviteit, kosten.

Op basis van deze onderzoeksactiviteiten kunnen uitspraken gedaan worden over zowel de effectiviteit als de praktische toepasbaarheid van de instrumenten in de werving en selectiepraktijk van bedrijven. De doelmatigheid van de instrumenten is een onderdeel van het laatstgenoemde evaluatieaspect. Er bestaan geen harde benchmarks voor de acceptabele kosten van dergelijke instrumenten. De doelmatigheid is daarom met name vanuit het perspectief van de deelnemende bedrijven te beoordelen.

Aanvullend op het directe effectiviteitsonderzoek op bedrijfsniveau zullen op pilotniveau interviews gevoerd worden met de verantwoordelijke projectleiders bij SZW, de uitvoerders van de bedrijfsadvisering en de ontwikkelaars van de interventies om het overkoepelende proces in kaart te brengen. Dit wordt aangevuld met een analyse van de beschikbare documentatie (projectplannen, contracten, kostenoverzichten, verantwoordingen) en monitoringgegevens die SZW zelf bijhoudt. Op basis hiervan worden uitspraken gedaan over het aantal bedrijven dat doorgaat met het gebruik van de instrumenten.

Eisen aan het onderzoeksresultaat

Bovenstaande onderzoeksactiviteiten zijn een richtlijn voor de uitvoering van het evaluatieonderzoek. Van doorslaggevend belang is dat het evaluatieonderzoek betrouwbare informatie oplevert die inzicht levert in:

- Daadwerkelijke implementatie van de instrumenten op bedrijfsniveau in relatie tot beleidslogica (kwantitatief en kwalitatief).
- Effectiviteit van de verschillende instrumenten met oog op de primaire uitkomstmaat, de kanselijkheid bij werving & selectie (kwantitatief).
- Ervaringen met implementatie en perceptie van effectiviteit bij betrokken personen (kwalitatief).
- Conclusies over praktische toepasbaarheid van instrumenten in reguliere HR-procesen bij bedrijven (kwalitatief).

4.6 Uitkomstmaten

Om te kunnen bepalen of de pilot en de instrumenten succesvol zijn, zijn voorafgaand uitkomstmaten vastgesteld die bepalend zijn voor de afweging of een instrument wel of niet breder uitgerold zou moeten worden.

De uitkomstmaten zijn afhankelijk van het instrument dat ingezet wordt. In zijn algemeenheid is vast te stellen dat de belangrijkste uitkomstmaat de (veranderende) kansverhouding van verschillende groepen sollicitanten in het werving- en selectieproces is. De pilot als geheel zet in op het bevorderen van een zo gelijk mogelijke kansverhouding (bij gelijke geschiktheid). De focus zou daarom niet (alleen) moeten liggen op het bevorderen van de kansen van een specifieke doelgroep maar op het resultaat dat de verhouding van kansen van verschillende doelgroepen evenrediger wordt.

We stellen de volgende uitkomstmaten voor:

Uitkomstmaat	Gegevensbron
Aantal betrokken bedrijven	Projectmonitoring
Aandeel bedrijven dat doorgaat met gebruik instrument	Projectmonitoring
Verbetering kansengelijkheid in praktijksetting	Quasi-experimenteel onderzoek, per instrument

4.7 Planning

De planning van het evaluatieonderzoek sluit aan op de planning van de pilot als geheel. Dat betekent dat er wordt uitgegaan van een start van de activiteiten in het vierde kwartaal van 2018, twee 'batches' van 10 werkgevers, en een looptijd van het onderzoek van 2018 t/m 2021.

Q4 2018-Q1 2019	Vorbereiding: onderzoeksplan, afspraken met bedrijven over onderzoek en controle/experiment groepen, vooronderzoek bedrijven, 0-meting bedrijven
Q2-Q4 2019	Experimentele fase batch 1, 1-meting augustus 2019, 2-meting december 2019. Voor elke tussenmeting wordt een publicabel tussenrapport opgeleverd.
Q3 2019	Vooronderzoek batch 2, afspraken en 0-meting
Q3 2019 - Q3 2020	Experimentele fase batch 2, 1 meting april 2020, 2-meting september 2020

5 Evaluatiekader Culturele Barometer

5.1 Beschrijving pilot

In de Pilot barometer culturele diversiteit wordt een instrument ontwikkeld dat primair voor werkgevers en secundair voor andere (intermediaire en beleids)organisaties de culturele diversiteit van het personeelbestand en het aanbod van werkzoekenden in kaart brengt.

De barometer bestaat uit drie onderdelen.

1. Per bedrijf dat meedoet aan de pilot (acht werkgevers zijn voorzien) worden op basis van CBS-gegevens drie aandelen werknemers gepubliceerd: a) werknemers met een westerse migratieachtergrond, b) werknemers met een niet-westerse migratieachtergrond en c) werknemers met Nederlandse achtergrond. Indeling geschiedt op basis van land van herkomst. Het gaat om een eenmalig standcijfer met peildatum december 2016. Daarnaast is er de optie voor werkgevers om hun eigen werknemersbestand met BSN te uploaden bij CBS. CBS koppelt deze gegevens vervolgens op persoonsniveau. Er kunnen dan recentere cijfers gebruikt worden, mogelijk op vestigingsniveau met uitbreidingen en/of uitsplitsingen naar instroom/uitstroom, functiecategorie, divisie, aantal dienstjaren of leeftijd.
2. Het tweede onderdeel is een benchmark zodat bedrijven zich kunnen spiegelen. De peildatum is eveneens december 2016. De benchmark geeft de totale verdeling werknemers in Nederland voor bovengenoemde drie groepen voor:
 - bedrijfstak en uurloon; een organisatie kan in de CBS-tabel de verdeling in de voor hem relevante bedrijfstak(ken) en uurloonklasse(s) opzoeken en zich hiermee vergelijken (2017).
 - bedrijfstak x regio (2016).
 - Beroepsgroep x regio. Beroep is gebaseerd op de Enquête Beroepsbevolking (EBB, een steekproefonderzoek). In de pilot wordt nagegaan of de celvulling voldoende is, ook omdat respons onder de personen met een migratieachtergrond vaak lager ligt. Om de steekproefomvang en daarmee het haalbare detailniveau te verhogen worden drie jaren van de EBB samengenomen (2015-2017).

3. Aanbod van werkzoekenden. Doel hiervan is inzicht te geven in hoeverre (toekomstige) werknemers van verschillende migratieachtergronden beschikbaar zijn op de arbeidsmarkt. Dit wordt gepresenteerd in absolute cijfers. Dit omvat:
 - Migratieachtergrond van studenten naar studieniveau, studierichting en regio (gegevensbestanden DUO).
 - Werkzoekenden naar beroepsgroep en regio (UWV).

Na de pilot zal er een go/no-go beslissing plaatsvinden. Bij een go kunnen ook nog (kleine) wijzigingen aangebracht worden. De go/no-go hangt af van de technische en juridische haalbaarheid en praktische bruikbaarheid van de culturele barometer. Ook hangt de inhoud (wel/niet uitgebreide versie) af van de CBS-kosten en het (verwacht) potentieel gebruik. Afhankelijk van de resultaten van de pilot vindt uitrol op landelijk niveau plaats (Q1 2019). Er is bij het ministerie van SZW reeds een 'wachlijst' met werkgevers die hieraan deel willen nemen.

Werkgevers uit de pilotfase gaan mogelijk door met een mogelijk tussentijds gewijzigde versie van de barometer. De effecten van het gebruik van de definitieve versie van de barometer zal geëvalueerd worden onder de acht werkgevers uit de pilotgroep en de nieuwe deelnemers.

In dit tussenrapport wordt alleen het evaluatiekader uitgewerkt voor de pilotfase, dus voor de go/no-go beslissing. Als op basis van de pilot duidelijk is hoe de definitieve barometer eruit komt te zien en hoe de verdere uitrol vormgegeven wordt, volgt een evaluatiekader voor deze uitrolfase.

5.2 **Beleidstheorie**

5.2.1 **Probleemdefinitie**

Er is thans gebrekkig inzicht in de culturele diversiteit van (potentiële) werknemers. De culturele barometer is een instrument dat bedoeld is om individuele werkgevers te ondersteunen door het geven van inzicht in de arbeidsmarktsituatie van personen met een (niet-westerse) migratieachtergrond.

5.2.2 **Hoofddoelstelling**

De hoofddoelstelling van de barometer is om te voorzien in de behoefte van werkgevers om inzicht te krijgen in de culturele diversiteit van hun personeelsbestand en het arbeidspotentieel van potentiële werknemers met een migratieachtergrond. Met dit inzicht

kunnen acties ondernomen/beleid gevoerd worden om de arbeidsmarktsituatie van personen met een (niet-westerse) migratieachtergrond te verbeteren.

De hoofddoelstelling van de pilot is te verkennen of een barometer culturele diversiteit haalbaar is.

De hoofddoelstelling van het VIA-programma is volgens de brief van de minister de duurzame verbetering van de arbeidsmarktpositie van Nederlanders met een niet-westerse migratieachtergrond.¹⁴ De barometer is primair bedoeld voor werkgevers en is daarmee vraaggericht en gericht op het vergroten van het inzicht in de diversiteit van de vervulde vraag en de diversiteit van het arbeidsaanbod.

De pilot staat los van andere pilots in het VIA-programma. Mogelijk zou het type CBS-informatie wel bruikbaar kunnen zijn voor andere projecten binnen het VIA-programma. Het gaat dan niet zozeer om de barometer zelf maar om de onderliggende bestanden.

5.2.3 Potentieel werkzame mechanismen, beoogde werking

In het 4E-model¹⁵ (Enable, Encourage, Exemplify en Engage) gaat het in deze pilot voornamelijk om Enable. Partijen krijgen inzicht zodat zij mogelijk acties gaan ondernemen die bijdragen aan het hoofddoel: het verbeteren van de arbeidsmarktsituatie van personen met een niet-westerse migratieachtergrond. Het instrument lokt niet direct actie uit (Encourage), maar indirect door informatie te ontsluiten. Ook, maar in mindere mate dan Enable, kan Engage een rol spelen. Er ontstaat via de pilotgroep een netwerk van werkgevers die kunnen helpen bij veranderingen en het mobiliseren van andere actoren. Het Engage-aspect wordt vergroot op het moment dat meer werkgevers zich gaan aansluiten na de pilotfase.

De verwachte werking van het instrument is gebaseerd op drie aannames. Allereerst is er behoefte aan het instrument onder werkgevers. Jaarlijks melden zich enkele bedrijven bij CBS met het verzoek om gegevens over de culturele diversiteit van hun personeelsbestand. Ook zouden zij zich willen benchmarken met hun sector en zouden zij het aanbod van relevant cultureel-divers arbeidspotentieel willen kennen voor hun personeelsplanning.

Ten tweede zijn er geen alternatieve informatiebronnen. De huidige data over culturele diversiteit onder werkzoekenden volstaan niet voor effectieve sturing op culturele

¹⁴ Minister van SZW (2018), *Verdere Integratie op de Arbeidsmarkt: de economie heeft iedereen nodig!*, Kamerbrief, 30 maart 2018

¹⁵ Department for Environment, Food and Rural Affairs (2008). A framework for pro-environmental behaviours. Report. January 2008.

diversiteit. Uit gesprekken met werkgevers blijkt dat de data van UWV en arbeidsregio's niet altijd voldoende concreet zijn, om werkgevers te helpen bij hun arbeidsmarktstrategie.

Ten derde biedt vrijwillige registratie onder werknemers onvoldoende bruikbare feitenbasis over culturele diversiteit. Uit de verkenning naar vrijwillige registratie blijkt dat dit proces zeer omslachtig is voor werkgevers en bovendien per definitie onvolledig blijft, omdat niet alle medewerkers hun gegevens aanleveren. Daarnaast zijn er uit het verleden negatieve ervaringen met de (in 2004 afgeschafte) Wet Samen, waarin werkgevers registratie van personen met migratieachtergrond dienden bij te houden.

De verwachte werking bouwt verder voort op het adagium meten is weten. Door informatie te ontsluiten kunnen gebruikers van de informatie van de barometer activiteiten ondernemen die de culturele diversiteit kunnen verbeteren. De werking is daarbij indirect: de barometer is een ondersteunend middel. Er zal door gebruik van de barometer niet direct een verbetering van de positie van personen met migratieachtergrond optreden. Bedrijven wordt niets dwingend opgelegd. Wel is de barometer een hulpmiddel om de situatie in kaart te brengen. Bedrijven kunnen met de informatie via acties en beleid de arbeidsmarktsituatie van personen met migratieachtergrond proberen te verbeteren. Bijvoorbeeld door specifieke aandacht in het HR-beleid of het vrijwillig instellen van een quotum/streefcijfer. Er is bijvoorbeeld een werkgever in de pilotgroep die aangegeven heeft een target te willen formuleren voor het aandeel personen met migratieachtergrond maar niet weet waarop die target gebaseerd moet worden.

5.2.4 Beleidslogica van de pilot

We kunnen nu de programmalogica definiëren volgens het Logic Model.

Input

Het ministerie van SZW stelt middelen ter beschikking om CBS de pilotversie van de barometer te laten ontwikkelen. Daarnaast zijn projectmedewerkers VIA actief. Deelnemende werkgevers stellen capaciteit gratis beschikbaar.

Throughput

Het project kent de volgende fasen die deels parallel lopen:

1. Praktische en juridische (privacy, onthullingrisico) voorverkenning
2. Opdrachtverlening CBS
3. Werving pilotdeelnemers
4. Ontwikkeling en oplevering barometer door CBS, publicatie dient openbaar te zijn

Output

De output wordt gevormd door de opgeleverde barometer culturele diversiteit. Acht werkgevers doen mee in de pilot. Een opbrengst is ook dat netwerkvorming optreedt tussen deelnemers aan de pilot. Dit netwerk kan later zijn diensten bewijzen wanneer de barometer uitgerold wordt.

Outcome

De barometer geeft inzicht in culturele diversiteit op bedrijfsniveau. Het gebruik van deze resultaten kan bedrijven helpen bij HR-beleid rond culturele diversiteit. Bedrijven kunnen bijvoorbeeld cijfers gebruiken om streefcijfers te formuleren voor aandelen personen met een (niet-westerse) migratieachtergrond in het personeelsbestand en op basis van informatie maatregelen nemen om instroom te bevorderen en uitstroom te voorkomen.

De werkgevers die graag diversiteit willen bevorderen ervaren geen knelpunten in de zin dat relevante informatie ontbreekt.

Impact

De barometer is een instrument waarmee werkgevers een feitenbasis gegeven wordt om meer culturele diversiteit mogelijk te maken en daarmee de arbeidsmarktpositie van de groep met een migratieachtergrond te verbeteren. De wijze waarop werkgevers dat doen wordt niet ingegeven door de barometer.

Daarnaast kunnen de totaalgegevens van de benchmark (onderdeel 2) en de gegevens van het arbeidsaanbod gebruikt (3) worden om een beter gebruik te maken van het arbeidspotentieel van arbeidsmigranten.



5.3 Onderzoeksvragen

Per onderdeel uit de evaluatiematrix kunnen we de onderzoeksvragen specificeren:

1. Input: Hoe zijn de begrote middelen voor de pilot ingezet?
 - Welk budget is gereserveerd en hoe is dit uitgegeven?
2. Throughput: Hoe is de uitvoering van de maatregelen verlopen? Waardoor zijn de maatregelen al dan niet conform voornemens uitgevoerd?
 - Is de barometer technisch haalbaar gebleken?
 - Is de inhoud van de barometer conform plan verlopen?
3. Output: Wat is tot stand gekomen in termen van producten, diensten en activiteiten?
 - Is de barometer gebruiksvriendelijk?
 - Welk organisaties hebben gebruik gemaakt van de barometer?
 - In hoeverre is het beoogde bereik (aantal deelnemende organisaties) gerealiseerd?
4. Outcome: Welke effecten (kunnen) worden gerealiseerd?
 - Welke inzichten geeft de barometer voor deelnemende organisaties?
 - Is dit inzicht waardevol?
 - Welke verwachte gevolgen heeft dit binnen de organisatie voor het HR-beleid?
5. Impact: Welke (langetermijn) effecten heeft dit tot gevolg gehad?
 - Leiden de genoemde outcome indicatoren tot meer culturele diversiteit binnen de arbeidsorganisatie?

In de pilotfase ligt de nadruk meer op de vaststelling of de ontwikkelde barometer bruikbaar en waardevol is. De impact doet zich voor op langere termijn en kan daarmee niet vastgesteld worden in de pilotfase, en wordt dus ook niet verder uitgewerkt in de volgende paragraaf.

5.4 Opzet van de evaluatie pilotfase

In deze notitie wordt de evaluatie beperkt tot de pilotfase. Deze valt uiteen in twee delen

- I. Technisch-juridische evaluatie.
 - II. Evaluatie gebruikers met inschatting gebruik, verwachte outcome, verwachte impact.
-
- I. Technische en juridische haalbaarheid. De vragen waarop deze evaluatie een antwoord geeft, zijn:
 - o Wat zijn de technische (on)mogelijkheden gebleken van de barometer? Op welke punten heeft CBS af moeten wijken van de oorspronkelijke doelstellingen?
 - o Hebben alle desaggregaties op macro- en regionaal niveau (SBI-niveaus, indeling beroepsgroepen) de gewenste informatie opgeleverd en zo ja via welke aanpak/bestanden?
 - o Wat is het resultaat en de reikwijdte (aantal lege cellen door onthullingsrisico) van de culturele barometer?
 - o Welke overige knelpunten of aandachtspunten voor het vervolg is CBS tegengekomen bij de ontwikkeling?
 - o Welke bevindingen zijn er met het oog op de verdere uitrol? Wat is nodig voor gebruiksvriendelijke en juridisch correcte ontsluiting voor meer werkgevers, welke delen lenen zich voor uitrol, welke aanpassingen zijn nodig en wat betekent dit voor eventuele kosten en doorlooptijd?
 - o Kan CBS aanvragen monitoren voor gebruik barometer? Welke afspraken kunnen gemaakt worden over monitoren daarvan communicatie daarover.

Aanpak:

Analyse schriftelijke evaluatie CBS.

Gesprek met CBS, projectleider/projectorganisatie Pilot barometer.

Planning:

Na oplevering barometer (december 2018)

- II. Evaluatie van gebruikers met een inschatting van het gebruik, de verwachte outcome en de verwachte impact.
- Throughput
 - Hoe verliep het proces en de samenwerking met CBS (met name van belang voor organisaties die eigen gegevens geupload hebben)?
 - Output
 - Wat vinden organisaties van de diverse onderdelen van de barometer (eigen werknemers en benchmarkgegevens)? Zijn de gegevens bruikbaar? Zijn de gegevens actueel genoeg?
 - Is het werken met de barometer gebruikersvriendelijk?
 - Welke verbeter suggesties zijn er? Welke onderdelen kunnen worden aangepast/uitgebreid/vervallen?
 - Verwacht gebruik:
 - Verwachten werkgevers de barometer te gaan gebruiken? Welke onderdelen? Waarvoor? En in welke mate (periodiciteit)?
 - Wat kan er aanvullend gedaan worden om het gebruik of de gebruikersvriendelijkheid te vergroten (bijvoorbeeld toelichting/gebruiksaanwijzing, voor wie HR-managers/beleid)?
 - Hoe is de werkgever van plan gebruik binnen de organisatie te monitoren, indien de barometer na de pilot fase verder uitgerold wordt?
 - Zijn werkgevers bereid 'afhaken' te melden aan SZW met inbegrip van redenen, indien de barometer na de pilot fase verder uitgerold wordt? Hoe?
 - Verwachte outcome:
 - Geeft de culturele barometer bedrijven naar verwachting goed inzicht in de culturele diversiteit binnen het bedrijf?
 - Zal de barometer leiden tot aanpassing van HR-beleid? Welke aanpassingen? Waarom wel/niet?
 - Zal de barometer tot ander gedrag leiden bij HR-medewerkers/beleid? Welke veranderingen verwacht men?

Aanpak:

- Survey onder alle deelnemers na oplevering van de barometer.
- Groepsdiscussie twee maanden na oplevering.

6 Tot slot

In dit tussenrapport hebben we een eerste aanzet gegeven voor het evaluatiekader voor het programma VIA. Het rapport geeft de stand van zaken weer per november 2018, legt een basis onder de evaluatiekaders voor de verschillende pilots, geeft een eerste aanzet voor het evaluatiekader van de programmaevaluatie en presenteert voor twee specifieke pilots uitgewerkte evaluatiekaders. In de loop van 2019 zullen meer pilots zijn uitgewerkt en volgt een uitgebreid en verfijnd eindrapport.

De Beleidsonderzoekers

Vestwal 2-4
2312 NP Leiden

071 566 59 47
info@beleidsonderzoekers.nl