Vertrouwensexperiment Oss
Zelf aan het stuur en op maat. Werkt het?

Ruud Muffels (projectleider), Kirsten Blom-Stam (onderzoeker)
& Stefan van Wanrooij (student assistent)

Tilburg-Oss, 13 januari 2020

© Tilburg University/Reflect

Tranzo/ReflecT - Research Institute for Flexicurity, Labour Market Dynamics and Social Cohesion, Tilburg University, PO Box 90153, 5000LE Tilburg, Tel.: +31 13 466 2181, reflect@tilburguniversity.edu

Inhoudsopgave

1. **Inleiding** (doel concept rapportage; inhoud rapportage, databronnen) 2
1.1 Aanleiding voor het onderzoek; beleidsvragen gemeente 3
1.2 Theoretisch kader 5
1.3 Onderzoeksvragen 13

2. Opzet onderzoek 14
2.1 Experimentgroepen 14
2.2 Procedure van random selectie en toewijzing 15
2.3 Onderzoekspopulatie 16
2.4 Product- en procesevaluatie 19
2.5 Dataverzameling product- en procesevaluatie 20

3 Resultaten **22**
3.1 Productevaluatie op basis van uitkomstmaten 22
3.1.1 Resultaten 24
3.1.2 De invloed van duur nader onderzocht: matching modellen 30
3.1.3 Conclusies 34
3.1.4 Secundaire uitkomstmaten 34

3.2 Procesevaluatie: vragenlijst 53
3.2.1 Gehanteerde uitkomstmaten 53
3.2.2 Analyse methode 53
3.2.3 Resultaten analyse procesevaluatievragenlijst 54
3.2.4 Conclusies 72

3.3 Focusgroep gesprekken met consulenten 73
3.3.1 Topics 74
3.3.2 Analysemethode 74
3.3.3 Resultaten 74
3.3.4 Samenvatting 83

4 Samenvattende bevindingen en voorlopige conclusies 85
4.1 Opzet van het experiment en werving 85

Literatuur 91
Bijlagen

1 Inleiding

Dit concept eindverslag van het Vertrouwensexperiment Oss bespreekt de uitkomsten van het onderzoek naar de effecten van een andere manier van bejegening en ondersteuning van mensen in de bijstand, wat bij aanvang een regelluwe bijstand werd genoemd, vergeleken met de reguliere aanpak. Het onderzoek bestaat uit een product -en een procesevaluatie. De productevaluatie onderzoekt de uitkomsten ofwel effecten van de andere aanpak vergeleken met de standaard aanpak niet alleen op uitstroom naar werk maar ook op welbevinden, gezondheid, keuzevrijheid, sociale participatie en zelfredzaamheid; de procesevaluatie onderzoekt de uitvoering van het experiment door de consulenten bij de sociale dienst van de gemeente Oss. Met andere woorden, hoe verloopt het proces van begeleiding en ondersteuning en wat gebeurt er in de spreekkamers waarin consulenten en bijstandsgerechtigden het experiment vorm geven. Daarbij gaat het onder andere om de vraag wat de andere manier van bejegening in vergelijking met de standaardaanpak doet met de consulenten en de deelnemers met het oog op de ondersteuningstaken.

Dit eindverslag vormt het derde verslag voor de gemeente - eerder zijn twee tussenrapportages uitgebracht in juni 2018 en april 2019 – maar het is wel het eerste verslag waarin de uitkomsten van de twee treatment -of interventiegroepen eigen regie en intensieve begeleiding in vergelijking met de standaard treatment worden gerapporteerd. In de tussentijdse verslagen zijn geen effecten van de afzonderlijke treatments gerapporteerd aangezien dat het experiment zou kunnen beïnvloeden waardoor een zuivere meting van de effecten niet meer mogelijk is. Wel zijn in de tussentijdse verslagen, en met name het tweede tussentijdse verslag, de stand van zaken, voortgang en de wijze waarop het uitvoeringsproces van het experiment verloopt weergegeven. Hierdoor is inzicht ontstaan in de ontwikkeling van het proces van experimenteren met de andere aanpak.

Landelijke experiment en LOEP

Het onderzoek staat niet op zichzelf maar is onderdeel van een landelijk experiment in de bijstand waarvoor zes gemeenten zich hebben gekwalificeerd. In totaal nemen meer dan 5000 mensen in de bijstand deel aan dit landelijke experiment. Het is daarmede een van de grootste experimenten ooit, niet alleen in Nederland maar wereldwijd. Het landelijke experiment Participatiewet is ook opgenomen in de RCT registry van de American Economic Association (AEA). De onderzoekers van de zes zogenaamde AMvB-gemeenten in Nederland maar ook de vijf niet-AMvB gemeenten hebben zich verenigd in het Landelijk Overleg Experimenten Participatiewet, LOEP genaamd. In dat overleg heeft tijdens de tweejarige periode voortdurende afstemming plaatsgevonden over de opzet van het onderzoek, de verwerking van de data, de analyses, de rapportages en de uitkomsten. De analyses in dit rapport zijn deels dan ook gebaseerd op een gezamenlijk analyseplan dat door de onderzoekers verenigd in LOEP is opgesteld. Daarnaast zijn aanvullende lokale analyses uitgevoerd onder andere vanwege de specifieke lokale aanpak. In Oss is bijvoorbeeld een procesevaluatie uitgevoerd welke niet in deze vorm is uitgevoerd in de AMvB gemeenten met uitzondering van het onderzoek van de experimenten in Tilburg en Wageningen waarvoor Tilburg University ook verantwoordelijk is.

1.1 Aanleiding voor het onderzoek: achtergrond

In de tussentijdse verslagen is de geschiedenis en achtergrond van het experiment reeds besproken. Per 1 oktober 2017 is Oss van start gegaan met wat destijds een regelluwe bijstand werd genoemd en daarna het vertrouwensexperiment (Muffels & van der Klink, 2017). Het experiment in Oss betreft evenals die in Apeldoorn en Epe een experiment buiten de AMvB[1] om, dus binnen de kaders van de bestaande Participatiewet. De gemeente wilde een nieuwe aanpak waarin wederzijds vertrouwen tussen cliënt en gemeente alsook eigen regie en maatwerk belangrijke uitgangspunten zijn van de begeleiding. Daarom wordt het experiment ook wel "Zelf aan het stuur en op maat" genoemd. Onderdeel van het experiment is om positieve prikkels in te zetten zoals het motiveren en belonen van het eigen initiatief van de deelnemer. Bij de opzet van het experiment is overigens aangesloten bij de opzet van de experimenten in de AMvB gemeenten. Dat zijn er zes: Tilburg, Wageningen, Groningen, Deventer, Nijmegen en Utrecht. Dat betekent dat ook in Oss en de andere niet-AMvB gemeenten is gekozen voor een 'random control trial' (RCT) design waarbij deelnemers aselect worden toegewezen aan de 'treatments' of behandelingsgroepen (zie onderzoeksvoorstel: Muffels & van der Klink, 2017). De treatment groepen zijn wel verschillend doordat niet-AMvB gemeenten niet mogen experimenteren met een ruimere vrijlating maar dus wel met versoepeling van de verplichtingen (eigen regie) en een andere bejegening en begeleiding van de cliënt (maatwerk en intensieve begeleiding).

Voortrekkersrol

Oss was samen met Apeldoorn en Epe een van de eerste gemeenten en daarmee voortrekker in Nederland om te starten met een experiment met regelluwe bijstand in de Participatiewet buiten de kaders van de AMvB om. Niet-AMvB gemeenten hebben meer vrijheidsgraden bij het ontwerp van het experiment dan de AMvB gemeenten doordat zij niet gebonden zijn aan de nogal als beperkend ervaren voorwaarden van artikel 83. Anderzijds hebben zij niet de mogelijkheid om de vrijlatingsbepalingen te verruimen. Nochtans zijn ook de niet-AMvB gemeenten gebonden aan de kaders van de huidige Participatiewet. De toetsing van de rechtmatigheid van de uitkering blijft bestaan en categorale ontheffing van verplichtingen voor specifieke treatment groepen is niet toegestaan.

Wetgeving: de invoering van de Participatiewet per 1 januari 2015

Sinds 1 januari 2015 vallen bijstandsuitkeringen onder de regelingen van de Participatiewet (PW), die de oude Wet werk en bijstand (Wwb) heeft vervangen. Doel van de Participatiewet is om zoveel mogelijk mensen naar betaald werk te helpen of, als dat niet kan, op een andere manier te laten participeren. De wet regelt de uitgifte van bijstandsuitkeringen en geeft een kader voor de regels en verplichtingen die bijstandsgerechtigden moeten naleven. Hoewel de wet de nationale kaders vastlegt, ligt de verantwoordelijkheid voor de uitvoering van de wet bij de gemeenten. Het doel van de wet was om meer mensen sneller te laten uitstromen naar betaald werk (zie SCP rapport, 2019). Na invoering van de wet op 1 januari 2015 is bij meerdere gemeenten de behoefte ontstaan om de doeltreffendheid van de nieuwe wet te onderzoeken, met andere woorden of de Participatiewet de gestelde doelen bereikt; werkt het

[1] AMvB is een acroniem voor "Algemene Maatregel van Bestuur", welke nodig was om toepassing van het experimenteerartikel 83 in de Participatiewet door gemeenten mogelijk te maken.

goed en zo niet, wat en hoe kan het dan beter werken. Doeltreffendheid richtte zich daarbij zowel op de toeleiding naar betaald werk als ook op de bevordering van de participatie in de samenleving. Bij het tweede doel kan gedacht worden aan vrijwilligerswerk, psycho-sociale of andere hulpverlening zodat mensen beter om kunnen gaan met hun persoonlijke en/of sociale problemen en daardoor beter kunnen participeren. In algemene zin vroegen de gemeenten zich af of de P-wet voldoende is toegerust om ondersteuning te kunnen bieden voor verbetering van de sociale participatie door middel van het bevorderen van de zelfredzaamheid, gezondheid en welbevinden van de bijstandsgerechtigde welke hiervoor bepalende factoren zijn. Bovendien vonden een aantal gemeenten de regelingen en verplichtingen in de Participatiewet niet altijd goed aansluiten bij de huidige praktijk. Het stelsel van regels en verplichtingen in de Participatiewet werd door gemeenten als complex en bureaucratisch ervaren, met weinig ruimte voor maatwerk en vormen van begeleiding waarbij de bijstandsgerechtigde zélf meer regie krijgt. Uiteindelijk zijn 11 gemeenten aan de slag gegaan om te experimenteren met een andere manier van begeleiding waaronder zes gemeenten die na een aanvraagprocedure zijn toegelaten als experiment in het kader van het experimenteer artikel 83 van de Participatiewet.

Wat houdt het onderzoek in?

Het onderzoek bestaat zoals eerder opgemerkt uit twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de cliënt; de nieuwe regelluwe aanpak of de reguliere aanpak die gekenmerkt wordt door een strenge toetsing op de bestaande verplichtingen en de toepassing van maatregelen en kortingen wanneer onverhoopt de bijstandsgerechtigde hieraan niet voldoet. Over een periode van twee aar zijn deelnemers aan het experiment gevolgd. In het onderzoek wordt er niet alleen gekeken naar of mensen eerder betaald werk vinden maar ook wat het met mensen doet, met hun gezondheid, hun welbevinden, hun zelfvertrouwen en hun zelfredzaamheid. Daartoe vullen deelnemers aan het begin van het experiment, na een jaar en na twee jaar een online vragenlijst in. De eerste meting noemen we de nulmeting en de twee anderen, de eerste en tweede vervolgmeting.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en de consulenten als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meer -of minderwaarde de consulenten ervaren van de andere bejegening van de cliënt voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe zijn om de acht maanden, in totaal vier, focusgroep gesprekken gehouden met de consulenten en de projectleider(s). Ten tweede vult de coach na elke 6-9 maanden op basis van de gesprekken in deze periode voor elk van hen een korte procesevaluatievragenlijst in over de afspraken en bevindingen van de afgelopen periode. In totaal zijn door de coaches drie vragenlijsten per deelnemer ingevuld. Tot slot is aan de consulenten gevraagd een vierde, meer algemene vragenlijst over de bevindingen gedurende de totale looptijd van het experiment voor alle door hen begeleide deelnemers in te vullen. De eerste door hen ingevulde procesevaluatievragenlijst vormt tezamen met het eerste focusgroep gesprek met de consulenten en projectleider(s) de nulmeting van de procesevaluatie. De vier focusgroep gesprekken daarna en de vier ingevulde vragenlijsten zijn onderdeel van de vervolgmetingen van de procesevaluatie. Vanaf 1 oktober 2018 zijn de consulenten vervolgens gestart met de invulling van de procesevaluatie vragenlijsten voor

elke deelnemer. De focusgroep gesprekken met de consulenten en de projectleiders van Oss hebben vervolgens op 6 december 2017, 3 oktober 2018, 12 juni 2019 en 30 oktober 2019 plaatsgevonden.

Inhoud concept eindverslag

Dit eindverslag is als volgt opgezet. In deel 1 wordt de opzet van het experiment en het onderzoek nader toegelicht. Dit houdt in dat de beleidsuitgangspunten en filosofie van het experiment, het theoretisch kader, zowel het wetenschappelijke als het beleidstheoretische kader, de werving van de deelnemers, de keuze voor de treatmentgroepen, de wijze van steekproeftrekking, de gestelde onderzoeksvragen en de manier van dataverzameling wordt toegelicht. Vervolgens worden in hoofdstuk 3 de resultaten besproken van de beschrijvende analyses van de product -en procesevaluatie op basis van de vragenlijsten. Hoofdstuk 4 bespreekt in dit concept summier de belangrijkste voorlopige conclusies van het onderzoek.

Gebruik van CBS microdata en planning rapportage

Dit eindverslag is een concept eindverslag. Het rapporteert over de effecten van het experiment op basis van de informatie die op lokaal niveau bij de gemeenten is verzameld (vragenlijsten deelnemers en procesevaluatie vragenlijsten consulenten) dan wel beschikbaar is middels de administratie van de gemeente Oss (BUS ofwel bijstandsuitkeringsstatistiek). In overleg met de projectleider(s) van de gemeente Oss is afgesproken dat een definitieve rapportage volgt nadat de effecten op volledige en gedeeltelijke uitstroom naar betaald werk zijn geanalyseerd gebruik makend van de microdata van het CBS. Deze analyse zal in het voorjaar van 2020 plaatsvinden in gezamenlijk overleg met de onderzoekers verenigd in het landelijk overleg experimenten Participatiewet (LOEP). De rapportage daarvan staat gepland op 1 april 2020. De zes AMvB-gemeenten maken eveneens gebruik van de microdatabestanden van het CBS om de effecten op de zogenaamde primaire uitkomstmaat te analyseren. Deze analyse gebeurt in overleg met het Centraal Planbureau dat door het Ministerie van Sociale Zaken en Werkgelegenheid is verzocht om de evaluatie van deze uitkomstmaat voor haar rekening te nemen. Door gebruik te maken van de door het CBS bewerkte microdata en een iets andere analysemethode kunnen de uitkomsten van de beide analyses verschillen. De onderzoekers zullen deze verschillen in kaart brengen en van toelichting voorzien opdat eenduidige beleidsconclusies mogelijk zijn. In het technische deel van de rapportage dat later volgt wordt een meer gedetailleerde en technische toelichting te geven op de manier waarop het onderzoek is uitgevoerd. Dit gedeelte is bedoeld voor lezers die meer achtergrondinformatie willen hebben van het onderzoek. Dit deel zal onderdeel uitmaken van het definitieve verslag op 1 april 2020.

1.2 Theoretisch kader

Tegen de achtergrond van de vragen die de gemeenten zich stelden over de doeltreffendheid van de nieuwe Participatiewet die per 1 januari 2015 werd ingevoerd en de wens ook in de gemeente Oss om te onderzoeken of een andere aanpak 'beter werkt' zijn door de onderzoekers drie treatments uitgewerkt welke zijn gebaseerd op verschillende theoretische inzichten uit de literatuur (zie Muffels & van der Klink, 2017). Twee daarvan, namelijk eigen regie en intensieve begeleiding zijn ook in de opzet van het experiment in Oss opgenomen. Voor de derde treatment van 'extra vrijlating van verdiensten' bestond er geen wettelijke basis, daarvoor moest gebruik worden gemaakt van artikel 83 welke alleen door de AMvB

gemeenten gebruikt mag worden. De theoretische begrippen die aan de basis van de ontwikkelde treatments staan zijn: wederkerigheid, vertrouwen, intrinsieke motivatie, keuzevrijheid, belonen (ook niet-financieel door positieve bejegening), zelfredzaamheid en individueel maatwerk (afgestemd op de wensen en mogelijkheden van de bijstandsgerechtigde). Deze begrippen zijn afgeleid uit de gedragseconomie (belonen, vertrouwen, intrinsieke motivatie), psychologische cognitie -en motivatietheorieën (wederkerigheid, zelfredzaamheid en intrinsieke motivatie) en de capability theorie van Amartya Sen (keuzevrijheid en maatwerk) (zie ook Groot, Muffels en Verlaat, 2019). Deze zullen hieronder nader worden toegelicht. Eerst gaan we in op de achterliggende beleidstheorie die aan de basis heeft gestaan van de vragen die de gemeente zichzelf stelde bij aanvang van het experiment.

Beleidstheorie

In artikel 83 van de Participatiewet wordt als 'primaire uitkomstmaat' van het experiment de uitstroom naar voltijd betaald werk genoemd. Dat betekent dat wanneer mensen minimaal 27 uur gaan werken tegen het minimumloon zij volledig onafhankelijk worden van de uitkering en dus uitstromen uit de bijstand die immers voor een alleenstaande 70% van het minimumloon bedraagt. Het primaire doel van de Participatiewet is zoals opgemerkt om meer mensen sneller naar voltijds betaald werk toe te leiden en daardoor participatie in de samenleving te bevorderen. Lukt dit niet dan zijn tussenstapjes op weg naar dit doel (activering) evenwel ook mogelijk (deeltijdwerk, vrijwilligerswerk, cursussen, training en opleiding, beschut werk, proefplaatsingen, aanpassingen werkplek en loonkostensubsidie voor(gedeeltelijk) arbeidsongeschikten). Loonkostensubsidie en beschut werk zijn nieuw in de P-wet. In de praktijk bleek dit doel zelfs in de herstelperiode van de arbeidsmarkt vanaf 2015 dus na de economische crisis moeilijk haalbaar. De arbeidsmarkt herstelde zich langzaam in de eerste jaren na de ernstigste naoorlogse economische recessie die duurde van 2008 tot einde 2014. Vanaf 2017 ging het beter en veranderde de arbeidsmarkt hetgeen in de jaren 2018 en 2019 leidde tot een erg krappe arbeidsmarkt. Het aantal tijdelijke banen bleef hoog en zelfs stijgen maar tegelijkertijd nam het aantal vaste banen toe. Dit gebeurde in de tijd dat het experiment van start ging. Het experiment vond dus plaats in een gunstig economisch tij waardoor met name jongere bijstandsgerechtigden een plekje konden vinden op de arbeidsmarkt. Voor velen in de bijstand (mensen met slechte gezondheid,ouderen, geen startkwalificatie) was dat ook in deze periode niet mogelijk.

Tussenstapjes op weg naar dit doel zijn evenwel ook mogelijk (deeltijdwerk, vrijwilligerswerk, beschut werk, proefplaatsingen, loonkostensubsidie voor (gedeeltelijk) arbeidsongeschikten). In de praktijk bleek dit doel zelfs in de herstelperiode van de arbeidsmarkt vanaf 2015 dus na de economische crisis moeilijk haalbaar. De arbeidsmarkt herstelde zich langzaam in de eerste jaren na de ernstigste naoorlogse economische recessie die duurde van 2008 tot einde 2014. Vanaf 2017 ging het beter en veranderde de arbeidsmarkt hetgeen in de jaren 2018 en 2019 leidde tot een erg krappe arbeidsmarkt. Het aantal tijdelijke banen bleef hoog en zelfs stijgen maar tegelijkertijd nam het aantal vaste banen toe. Dit gebeurde in de tijd dat het experiment van start ging. Het experiment vond dus plaats in een gunstig economisch tij waardoor met name jongere bijstandsgerechtigden een plekje konden vinden op de arbeidsmarkt. Voor velen in de bijstand (mensen met slechte gezondheid, ouderen, geen startkwalificatie) was dat ook in deze periode niet mogelijk. De meerderheid van de mensen in de bijstand betreft immers een zeer kwetsbare groep waarvan slechts een klein deel kan worden toegeleid naar betaald werk (7,5-

10% per jaar in Oss). Bovendien is de kans om in een tijdelijke baan terecht te komen groot (80%) waardoor de kans op terugkeer in de bijstand op korte termijn (binnen 2 jaar) ook groot is. In Oss is deze kans tussen 2017 en 2019 ongeveer 15%. Bijna driekwart van de bijstandsgerechtigde in Oss heeft geen startkwalificatie (MBO-niveau 2) en is daardoor voor de meeste werkgevers onvoldoende geschikt voor een baan. Daarnaast vinden velen een tijdelijke baan maar met een betrekkelijk gering aantal arbeidsuren per week waardoor ze deels afhankelijk blijven van de bijstand. Voor zover mensen gedeeltelijk of volledig uitstromen naar betaald werk lukt dat alleen met (intensieve) persoonlijke begeleiding en maatwerk (zie o.a. Muffels e.a. 2016).

De stap naar betaald werk is voor een groot deel van de bijstandsgerechtigden dus te groot; voor velen zijn tussenstapjes nodig voordat een terugkeer in betaald werk mogelijk is, voor anderen in het zogenaamde "granieten bestand" is uitstroom naar betaald werk een onhaalbare kaart zelfs op de wat langere termijn. Zij hebben speciale ondersteuning nodig met een specifieke maatgerichte aanpak in verband met hun persoonlijke en maatschappelijke problemen en problemen met hun fysieke of mentale gezondheid. Op basis van de informatie in de bijstandsuitkeringsstatistiek (BUS) bleek dat ruim 60% van de mensen in de bijstand in Oss met gezondheidsproblemen kampten.

In de eerste tussenrapportage is met behulp van latente klasseanalyse[2] een analyse gemaakt van de deelnemers aan het experiment op basis van informatie in de eerste vragenlijst. Het zoek -en sollicitatiegedrag van de deelnemer is geanalyseerd plus de belemmeringen die zij ervaren onder andere vanwege hun gezondheidsproblemen en een laag niveau van arbeidsgeschiktheid (inclusief hun kwalificaties). Destijds was voor 189 mensen informatie beschikbaar. Later is deze analyse herhaald op alle mensen die de nulmeting hebben ingevuld en waarvoor de informatie beschikbaar was (n=248). Het beeld bleef hetzelfde. Zes kenmerken van de deelnemers in Oss zijn gebruikt om de profielen te maken. Deze hebben allen te maken met het zoeken naar werk en de perceptie van de mate van arbeidsgeschiktheid inclusief de fysieke en mentale beperkingen die mensen ervaren om te gaan werken. De beste oplossing uit de analyse was een indeling in drie groepen. De eerste groep noemen we de 'actief op zoek, arbeidsgeschikte' groep en bestaat uit mensen die aan het werk willen en actief op zoek zijn. Dat is het geval voor een op de twee deelnemers. Zij zijn actief op zoek naar betaald werk (54%) en zien daar ook nog mogelijkheden voor. Dan zijn er twee groepen die ofwel berustend en ontmoedigd zijn (16%) ofwel vanwege gezondheidsproblemen passief zijn en nog weinig kansen zien voor het vinden van betaald werk (30%). Deze tweede, passief, arbeidsongeschikte groep kenmerkt zich doordat ze weinig solliciteren en ook niet actief meer zoeken maar bovenal aangeven een slechte fysieke gezondheid (26%) en vooral een slechte mentale gezondheid (69%) te hebben. De eerste berustende, ontmoedigde groep zijn 'willers'

[2]Latente klasse analyse probeert aan de hand van een aantal waargenomen variabelen (bevraagd in de nulmeting) een achterliggende) niet waargenomen variabele te vinden (zie https://nl.wikipedia.org/wiki/Latenteklassemodel). Minimaal bestaat deze variabele uit twee categorieën verwijzend naar twee latente klassen. Een latente klasse wordt in de analyse weergegeven door voorwaardelijke kansen die voor iedere mogelijke score op de waargenomen variabelen wordt berekend. Deze kans geeft aan hoe waarschijnlijk een deelnemer die tot een van de klassen behoort een bepaalde score heeft op de waargenomen variabelen. De oplossing is technisch gezien uitstekend: er is een sterk onderscheid tussen de klassen en het aantal classificatiefouten is gering (6%). De analyses zijn uitgevoerd met het programma Latent Gold 5.1.

die ook wel 'kunnen', ze zijn arbeidsgeschikt, maar zoeken niet meer omdat ze ontmoedigd zijn vanwege het zoeken naar werk maar geen werk kunnen vinden.

Vanuit dit besef hebben de gemeenten bij aanvang van het experiment de wens geuit om naast de primaire uitkomstmaat van volledige uitstroom naar betaald werk ook een aantal secundaire uitkomstmaten te gebruiken om te bepalen of het experiment al dan niet werkt voor de deelnemers. Juist voor mensen waarvoor de stap naar betaald werk te groot is kunnen deze secundaire uitkomsten van groot belang zijn om later ook ander stappen te kunnen zetten. De uitkomstmaten die de gemeente Oss heeft gekozen zijn: vertrouwen, keuzevrijheid, bevordering van zelfredzaamheid, subjectief welbevinden, gezondheid (fysiek en mentaal) en sociale participatie. Daarnaast werd belang gehecht aan een betere relatie tussen burger en overheid door een andere bejegening van de burger uitgaande van vertrouwen, maatwerk en aandacht in plaats van 'loslaten', 'weinig of geen contact' en wantrouwen welke begrippen in de ogen van de gemeenten aan de basis staan van de reguliere aanpak. Veel gemeenten die vanaf de invoering van de nieuwe P-wet werden geconfronteerd met de volledige verantwoordelijkheid voor de financiering van de bijstandsuitkeringen middels de BUIG-middelen zochten een weg om de wet doeltreffender uit te voeren. Een stringente toetsing op verplichtingen en voorwaarden was een van de manieren om dat te doen. Sommige gemeenten zijn die weg op gegaan, anderen zochten juist naar mogelijkheden om de effectiviteit van de re-integratie te verbeteren middels een minder strikte toepassing van de arbeids- en re-integratieplicht en meer ruimte voor eigen regie van de deelnemer. Die weg is Oss opgegaan in de vorm van het inrichten van het vertrouwensexperiment waarin naast eigen regie en het minder stringent toepassen van de verplichtingen een intensievere en meer op maat gesneden begeleiding en ondersteuning wordt aangeboden vooral ook aan mensen in het zogenaamde "granieten" bestand. Daarom wordt het experiment ook wel aangeduid met 'zelf aan het stuur' en 'begeleiding op maat'.

Wetenschappelijk kader

Het experiment is theoretisch ontleend aan een viertal inzichten, die ontleend zijn aan een brede onderzoeksliteratuur in de gedragseconomie, de empirische sociologie en de cognitie -en motivatiepsychologie (zie Groot, Muffels en Verlaat, 2018):

- Het eerste inzicht gaat over recente onderzoeksresultaten over de invloed van armoede op de 'mindset' ofwel geestesgesteldheid van mensen. Onderzoek in dit relatief nieuwe onderzoeksgebied laat zien dat (financiële) schaarste en stress door armoede de cognitieve hulpbronnen van mensen verminderen (Manietal, 2013; Mullainathan en Shafir, 2013). Men heeft alle energie, geestkracht en inventiviteit nodig om de eindjes voor henzelf, eventuele partner en kinderen aan elkaar te knopen door het ontwikkelen van zogenaamde huishoudbezuinigingsstrategieën om in de eigen behoeften en die van het huishouden te kunnen voorzien. Indien financiële schaarste en het voldoen aan bijstandsverplichtingen een groot deel van de cognitieve hulpbronnen van mensen innemen, is er weinig ruimte voor belangrijke en cognitief uitdagende taken, zoals het bijscholen voor een andere baan, het onderhouden van het sociaal netwerk of het actief zoeken naar betaald werk (Groot, Muffels en Verlaat, 2019).

- Het tweede inzicht is afkomstig uit de gedragseconomie en gaat over de invloed van instituties op de arbeidsmarkt en in de sociale zekerheid alsmede de impliciete waarden die zij vertegenwoordigen op het gedrag van mensen. De impliciete waarden die ten grondslag liggen aan deze instituties en die tot uiting komen in uitkeringsvoorwaarden, regels en verplichtingen beïnvloeden het gedrag van de gebruikers in de vorm van bijvoorbeeld het zoekgedrag op de arbeidsmarkt, maar ook de keuze tussen werken en vrije tijd in de vorm van bijvoorbeeld het aantal uren dat men wil werken. Daarnaast beïnvloeden zij andere economische beslissingen in de sfeer van consumptie en bestedingen, huisvesting, onderwijs en gezondheid niet alleen van henzelf maar ook van anderen in het huishouden waaronder ook eventuele kinderen. Onderliggende waarden zijn bijvoorbeeld reciprociteit (wederkerigheid, 'tit for tat') en vertrouwen. Reciprociteit betekent dat individuen een goede behandeling of het ontvangen van vertrouwen (een investering in sociale relaties), belonen door bijvoorbeeld zich extra in te spannen (positieve reciprociteit), terwijl ze het omgekeerde doen als ze slecht worden behandeld of op basis van wantrouwen worden bejegend (negatieve reciprociteit) (Fehr en Schmidt, 2003). Negatieve prikkels, zoals opgelegde financiële sancties en kortingen zijn niet noodzakelijkerwijze de beste manier om mensen te prikkelen mee te werken en zich aan de regels te houden. Economisch en sociologisch onderzoek laat zien dat er weliswaar bescheiden korte termijn effecten optreden (men accepteert sneller werk ook al is de 'fit' minder) maar dat de effecten op de langere termijn onzeker zijn bijvoorbeeld doordat een slechte 'fit' tussen mens en werk op de iets langere termijn welbevinden, gezondheid en de loopbaan negatief beïnvloeden (zie SCP, 2019; Blonk, 2018). Bevindingen uit de experimentele economie laten ook zien dat mensen in ruil voor het vertrouwen dat zij ontvangen extra gemotiveerd zijn en meer hun best doen voor hun taak en daarmee dus ook degene die hen vertrouwt beloont (Groot, Muffels en Verlaat, 2019). Op deze manier leidt vertrouwen dus tot gevoelens van positieve reciprociteit en daardoor ook tot blijvende inzet en toegenomen productiviteit (Bohnet et al., 2001).
- Het derde inzicht is afkomstig uit de psychologische motivatie theorie en leert ons dat extrinsieke prikkels de intrinsieke motivatie kunnen wegnemen (Frey en Jegen, 2001). De zelfbeschikkingstheorie (Deci en Ryan, 1985) stelt dat intrinsiek gemotiveerde mensen een activiteit ondernemen omdat ze het plezierig en interessant vinden, waardoor ze meer effectiviteit en volharding laten zien in hun gedrag en een verbeterd welbevinden (Ryan et al., 1997). Deze theorie stelt ook dat het geven van vertrouwen aan mensen een gevoel van zelfsturing oplevert, die weer effect heeft op het werkzoekgedrag en duurzame werkgelegenheid. Voorgaand onderzoek wijst ook uit dat de intrinsieke motivatie kan worden versterkt door een activiteit als keuze aan te bieden in plaats van als controlemiddel (Groot, Muffels en Verlaat, 2019).
- Het laatste inzicht gaat over 'vrije keuze' en komt voort uit Sen's 'capability theorie'. Binnen deze theorie zijn 'capabilities' de keuze opties die mensen worden aangeboden om de dingen te doen die zij belangrijk vinden in hun eigen leven. Op deze manier krijgen mensen kansen aangeboden waardoor ze persoonlijke autonomie en zelfvertrouwen ontwikkelen en daarmee hun welbevinden kunnen vergroten (Sen, 2004; 2009). Op basis van de inzichten van Sen wordt in recente theorievorming rond duurzame inzetbaarheid het waarde-aspect van werk benadrukt

alsook het belang van eigen regie (Van der Klink et al., 2016). De focus op keuzemogelijkheden en eigen regie vergroot dus ook de intrinsieke motivatie om te gaan werken.

Deze vier inzichten vormen de uitgangspunten voor de manier waarop de eerder genoemde alternatieve interventie-benaderingen (of treatments) zijn vormgegeven binnen het vertrouwensexperiment. Tevens zijn ze inspiratiebron geweest voor het opstellen van vragen in de vragenlijsten aan de deelnemers en de coaches. De inzichten zijn niet alleen relevant voor de productevaluatie waarin het effect wordt onderzocht op arbeidsinschakeling en op welbevinden, gezondheid en sociale participatie maar ook voor de procesevaluatie omdat de procesevaluatie onderzoekt in hoeverre sturing op zelfredzaamheid, intrinsieke motivatie en keuzevrijheid ook beter werkt in de uitvoering. We hebben aan het begin devolgende hypothesen geformuleerd.

Hypothesen

- Op grond van gedragseconomische inzichten verwachten we dat positieve prikkels in de vorm van bijvoorbeeld een andere, positievere en meer ondersteunende bejegening met minder verplichtingen en gebaseerd op vertrouwen leidt tot een grotere motivatie en tot meer coöperatief gedrag bij de cliënten hetgeen zich uit in een grotere bereidheid tot deelname alsook inzet om de re-integratie tot een succes te maken. We verwachten positieve effecten op motivatie en daarmede op participatie en arbeidsinschakeling in deze groepen in vergelijking met de controlegroep.
- Omgekeerd verwachten we op grond van het wetenschappelijk empirische onderzoek dat een systeem van strenge handhaving, monitoring en sanctionering zoals bij de controle –en referentiegroep positieve effecten kan hebben op inschakeling in werk maar dat de vraag is of deze effecten korte termijn effecten zijn en dat duurzame uitstroom mogelijk minder zal zijn bij deze groep dan bij de treatment groepen
- Een streng systeem van controle en handhaving kan negatieve effecten hebben op de motivatie(motivatietheorieën) en de mentale belasting en stress die het systeem van controle en straffen met zich meebrengt met negatieve gevolgen voor welbevinden en de (mentale) gezondheid.
- In overeenstemming met de uitkomsten van het Fit4Work programma en de uitkomsten van review studies naar de effecten van IPS (individual placement support) en ES ('employment support') interventies verwachten we sterke effecten van intensieve begeleiding op maat op de arbeidsinschakeling van vooral de meest kwetsbare groepen in de bijstandspopulatie (het granieten bestand).
- De focus op eigen regie en zelfredzaamheid geeft mensen handelings- en keuzevrijheid terwijl intensieve begeleiding en maatwerk hen extra keuzemogelijkheden en kansen biedt. Hierdoor neemt hun zelfvertrouwen en welbevinden toe waardoor het zoekgedrag toeneemt alsook de intrinsieke motivatie om te gaan werken hetgeen uiteindelijk ook de arbeidsinschakeling bevordert. Dit verwijst naar Sen's 'capability' theorie en Ryan en Deci's zelfdeterminatie theorie.

Procesevaluatie

De procesevaluatie beoogt inzicht te krijgen in hoe het proces van experimenteren verloopt in de spreekkamer van de professional waar de ontmoeting met de burger plaatsvindt. Hoe vertaal je de treatments naar de uitvoeringspraktijk; wat betekent eigen regie en hoe intensief is intensief in de treatment intensieve begeleiding? Hoe worden de treatments uitgevoerd, op welke manier wordt een vertrouwensband met de deelnemer opgebouwd, welke condities belemmeren een goede uitvoering van de treatments etc.? De beschikbare wetenschappelijke literatuur over 'wat en waarom werkt' in het rentegratie beleid is beperkt. In de oratie van Blonk (2018) wordt gesteld dat de meeste studies gericht zijn op beleid en beleidsevaluatie meer dan op het begrijpen van de individuele gedragsverandering en het bieden van handvatten aan de professional. Evaluatiestudies van re-integratie beleid laten zien dat de effecten van dat beleid (bemiddeling, werkcreatie, loonkostensubsidies) niet erg groot zijn, en veelal korte termijn effecten laten zien (zie de overzichtsstudies van Card, Kluve en Weber, 2010; 2015). De effecten van het activerend arbeidsmarktbeleid bedragen gemiddeld over alle studies voor de middellange termijn (12 tot 24maanden) 11,5%. Voor experimenteel onderzoek bedraagt het gemiddeld percentage 5,6%. Dit betreft het gemiddelde van begeleiding bij het zoeken naar werk, werk creatie in publieke sector, controle/sancties, subsidies voor private werkgelegenheid en training. Dit betreft dus zowel mensen in de werkloosheidsverzekering als in de bijstand. De studies tonen wel aan dat training en ontwikkeling dus investeren in menselijk kapitaal effectiever is dan de andere instrumenten voor de wat langere termijn. Een recente studie van een soortgelijk 2-jarig experiment met intensieve begeleiding en maatwerk in Lausanne in Zwitserland tussen 2015-2017 laat zien dat het effect op uitstroom naar werk gemidddeld zo'n 9% bedraagt over de gehele periode (Bigotta, Bonoli, Fossati, Lalive, and Oesch 2018). Internationaal onderzoek toont ook aan dat Intensieve begeleiding van mensen met een psychische beperking werkt en effectief om deze mensen naar werk toe te leiden. De effecten hiervan zijn groter dan de hierboven genoemde effecten en variëren gemiddeld tussen 10 tot 25% (gemiddeld 15%). De meest sprekende voorbeelden van succesvolle begeleiding naar werk van mensen met (mentale en fysieke) gezondheidsbeperkingen komen dan ook uit de psychologische literatuur rond "Individual Placement Support" en "Employment Support" (zie Bond et al. 2010; Marshall et al. 2014). Daarin wordt specifiek en gericht individueel maatwerk geboden (zowel aan de aanbod als de vraag -of werkgeverskant) welke leidt tot de genoemde additionele uitstroompercentages.

Blonk (2018) stelt in zijn oratie daarom het volgende: "Je moet dus investeren in mensen en dat vooral volhouden. Dat is het beste dat je kunt doen. Maar we weten ook dat niet elke training helpt en ook niet voor iedereen in gelijke mate. Om daar meer grip op te krijgen is inzicht nodig in de werkzame bestanddelen van en de condities rond een training. Niet alleen wat werkt, maar waarom, zodat de professional onderbouwde keuzes kan maken wat wordt ingezet. Die kennis is nog maar beperkt beschikbaar. Er is wat dat betreft nog steeds een hoog 'black box' gehalte. Om dat op te lossen zijn meer studies nodig waarin naar gedragsaspecten wordt gekeken. Aspecten zoals kwaliteit van zoekgedrag, naar de intensiteit daarvan en de focus, naar vertrouwen in eigen in dat zoekgedrag en naar vaardigheden die zelfregie ondersteunen en hoe op deze aspecten te interveniëren. Wat dat betreft zijn we nog maar net begonnen." (citaat oratie Blonk, 2018). Vervolgens verwijst Blonk naar studies van Liu, Wang en Huang uit 2014 waarin wordt onderzocht wat werkt bij een aantal specifieke solicitatietrainingen die hebben geleid tot een driemaal grotere kans om werk te vinden voor de deelnemers. Het is volgens Blonk de combinatie

van motivatie en gedrag die bepalend zijn voor het succes van deze trainingen. Deze overwegingen hebben aan de basis gestaan van de procesevaluatie en de invulling van de 'skills' training die aan consulenten is gegeven om deelnemers te leren zelfredzaam te zijn of te worden. De basis van de training wordt gevormd door de theorie van 'planned behavior' ofwel het integratieve gedragsmodel van Fishbein & Ajzèn (2010). In het model (Fig. 1.1) staat zoekgedrag naar werk centraal. Het gaat om intenties ofwel motivatie en vaardigheden. De. motivatie is afhankelijk van het belang van het te bereiken doel, de ervaren druk van de omgeving en joe je daarmee omgaat en de verwachting dat je het doel kunt halen. Bij vaardigheden gaat het om vaardigheden die je ten dienste staan bij het eigen regie nemen en houden over je leven, over zelfredzaamheid, volhouden en niet bij de pakken te gaan neerzitten ondanks je belemmeringen. Deze vaardigheden zijn te leren. De laatste component in het model zijn de mentale en fysieke beperkingen maar ook kenmerken zoals leeftijd, nationaliteit of culturele achtergrond. Het omgaan met deze beperkingen en kenmerken is onderdeel van zelfredzamheid en eigen regie.De training is een aantal keer in Oss gegeven. In de focusgroepgesprekken met de consulenten is de betekenis en rol van de training in het uitvoeringsproces voortdurend besproken. In de verslaglegging van deze geprekken wordt hier nader op ingegaan.

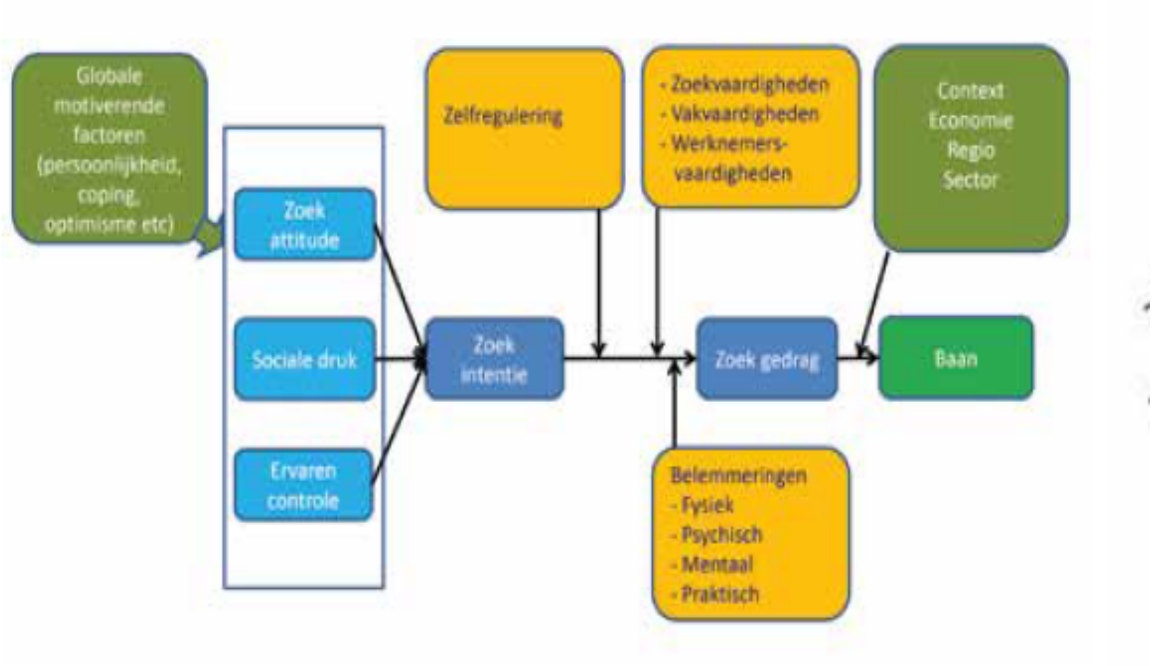


Fig. 1.1. Het integratieve gedragsmodel van Fishbein & Ajzèn (2010).
Bron: Oratie Blonk (2018):

Vanuit wetenschappelijk oogpunt is nu vooral van belang om te toetsen of als de uitvoering niet tot het gewenste resultaat leidt dit het gevolg is een 'theory failure', de uitgangspunten zijn niet goed gefomuleerd of een 'uitvoeringsfailure', de treatments zijn niet conform de onderliggende uitgangspunten uitgevoerd. De vragenlijsten bij de coaches en de focusgroepgesprekken hebben tot doel

om hier meer inzicht in te krijgen. Daarbij analyseren we in samenhang de invloed van contextuele factoren (politiek, gemeentelijke uitvoering), het gedrag van de coaches en deelnemers, de bedoelde en onbedoelde effecten van de treatments, en de uitkomsten en effecten.

1.2 Onderzoeksvragen en hypothesen

Het doel van het onderzoek is om te achterhalen wat het effect is van de twee alternatieve interventies of aanpakken op de gekozen uitkomstmaten in vergelijking met de standaard aanpak. Het idee is dat door een andere meer positieve bejegening (inhoudende eerder belonen dan straffen) en een andere manier van het 'gesprek' aangaan, meer vanuit intrinsieke motivatie en vertrouwen dan vanuit extrinsieke prikkels en wantrouwen, het gedrag positief wordt beïnvloed in de richting van meer eigen initiatief en zelfredzaamheid. Uiteindelijk wordt daarmede ook het hoofddoel van de P-wet om meer mensen sneller toe te leiden naar betaalde arbeid beter bereikt.

Wetenschappelijk gezien zijn de experimenten interessant omdat er nog weinig bekend is over de effectiviteit van een alternatief beleid van motiveren en belonen (positieve prikkels) in plaats van een beleid van straffen en sanctioneren (negatieve prikkels). De doeltreffendheid van het beleid is voor dit onderzoek breed gedefinieerd. Dat houdt in dat er wordt gekeken naar effecten op vijf domeinen: (1) arbeidsparticipatie, (2) sociale participatie en vertrouwen, (3) gezondheid en welbevinden en (5) de financiële situatie van de deelnemers. De doeltreffendheid wordt vastgesteld door de twee alternatieve interventies te vergelijken met de standaardaanpak. Hieruit resulteren een aantal beleids-en onderzoeksvragen:

Beleids -en onderzoeksvragen

- In het experiment worden de verplichtingen, die gekoppeld zijn aan de inspanningen die van cliënten worden verwacht om te participeren en re-integreren, bijvoorbeeld de sollicitatie -en reintegratieplicht versoepeld (met name ook in de eigen regiegroep). De vraag voor beleid is dan wat is het effect van deze ontheffingen op de gekozen uitkomstmaten?
- Een tweede beleidsvraag betreft het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name intensieve begeleidingsgroep).
- Een derde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooien, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, intensieve begeleiding).
- Een belangrijke vraag voor de gemeente was wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en gezondheid (eigen regie, intensieve begeleiding).
- In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de consulenten. De consulenten krijgen ook meer tijd (lagere caseload) om de deelnemer te begeleiden. De gemeente wil graag weten hoe dit proces

verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Hoe verloopt de interactie met de deelnemer? Wat werkt in de uitvoering, in de spreekkamers beter en wat werkt niet ook vanuit het perspectief van de deelnemer? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).

2 Opzet van het onderzoek

2.1 Experimentgroepen

Oss kent evenals de andere niet-AMvB gemeenten waarvoor Tilburg University de evaluatie verricht drie treatment of interventiegroepen: 1. Eigen regie; 2. Maatwerk/intensieve begeleiding en 3. De reguliere ofwel standaard aanpak.

1. *Eigen regie of 'zelf aan het stuur'*. Dit is de groep waarbij de focus ligt op het minder stringent toepassen van de verplichtingen en mensen wordt geleerd zelfredzaam te zijn en zelf de regie te nemen om de problemen aan te pakken. Van deze groep wordt verwacht dat zij zelfredzaam zijn of worden om terugkeer naar werk te realiseren, door de deelnemers te leren hoe ze proactief en met zelfvertrouwen (en dus zelfsturend) moeten handelen. Bij deze groep zijn de verplichtingen van de sollicitatie en re-integratieplicht minder strikt toegepast, om zo meer ruimte te laten voor eigen regie van de deelnemer. Sancties worden in deze groep slechts in uitzonderlijke gevallen toegepast. Voor het overige blijft alles zoals in het normale regime, bijvoorbeeld bijverdiensten worden hetzelfde behandeld als voor bijstandsgerechtigden die niet deelnemen aan het experiment.
2. Op maat: dit is de groep die individueel maatwerk ontvangt en extra persoonlijke begeleiding krijgt. Bij de intensiveringsgroep blijven alle regels en regelgeving van het bestaande bijstandsregime gelden, waaronder de toepassing van de re-integratieplicht. Wel ontvangt deze groep extra hulp en/of meer contact door middel van op maat gemaakte begeleiding en intensieve bemiddeling van de gemeente. Gemiddeld hebben deze deelnemers twee keer zoveel contactmomenten met de gemeente dan de deelnemers in de controlegroep.
3. Regulier: dit is de groep die de standaard aanpak krijgt. Voor hen gelden de bestaande strenge aanvraag- en re-integratieverplichtingen.

'Hawthorne' effecten

De reguliere/standaardaanpak kan in de loop van de tijd veranderen omdat er in de gemeentelijke organisatie uitstralingseffecten van het experiment zijn op de begeleiding van de bijstandscliënten in de reguliere aanpak (zogenaamde 'Hawthorne' effecten). Dit heeft uiteraard in dat geval invloed op het onderzoek naar de effecten van het experiment. Om die reden is ook nog een referentiegroep onderscheiden waarin alle tot de populatie behorende bijstandsgerechtigden die niet aan het experiment deelnemen zijn vertegenwoordigd. Deze worden administratief gevolgd. De onderzoekers zullen uiteraard in het onderzoek met deze verandering in de standaardtreatment sinds de start van het experiment per 1 oktober 2017 rekening houden. Van belang is dat ook de overige 'treatments' zoveel als mogelijk is in stand worden gehouden en niet veranderen maar ook dat is niet vanzelfsprekend aangezien beleid

voortdurend verandert en het experiment invloed heeft op de begeleiding door de consulenten in de treatmentgroepen. In de praktijk blijkt dit een lastige opgave omdat er experiment -en leereffecten optreden bij de consulent waardoor de treatment zich wijzigt. De vergelijking met de referentiegroep is daardoor erg belangrijk vooropgesteld dat de treatmentgroepen en de referentiegroepen wat samenstelling betreft vergelijkbaar zijn. Deze vergelijking met de referentiegroep is overigens alleen mogelijk bij de uitkomstmaat uitstroom naar werk, maar niet bij de overige uitkomstmaten van welbevinden, gezondheid, zelfredzaamheid, keuzemogelijkheden en sociale participatie. Daarvoor ontbreekt een referentiegroep omdat over deze uitkomstmaten geen administratieve data beschikbaar zijn.

2.2 Procedure en werving

Voor het onderzoek is er gebruik gemaakt van een RCT design, dat wil zeggen een gerandomiseerd experiment met een controlegroep[3], waarbij deelnemers aselect worden toegewezen aan de 'treatments' of behandelingsgroepen. De doelgroep omvatte alle bijstandsgerechtigden in Oss met uitzondering van enkele specifieke groepen die zijn uitgesloten van deelname aan het experiment omdat ze niet voldeden aan de uitsluitingscriteria. Uitgesloten zijn mensen die geen recht (meer) hebben op een uitkering Participatiewet, mensen die in een inrichting verblijven, dak -en thuislozen, mensen die gedurende het experiment met pensioen gaan (>65 jaar), mensen die de Nederlandse taal niet machtig zijn, mensen die volledig arbeidsongeschikt zijn en mensen die andere regelingen zitten (IOAW, IOAZ, BBZ) bij aanvang van het experiment.

Steekproef en werving

In Oss is besloten om de (gestratificeerde) steekproef voor de helft te trekken uit het zittende bestand en voor de helft uit de "nieuwe instroom". Daarbij werden alle mensen met een lopende uitkering op 1 juni 2016 beschouwd als mensen uit het zogenaamde zittende bestand en alle mensen die op de peildatum van 1 juni 2017 na 1 juni 2016 zijn ingestroomd, of nog instromen na 1 juni 2017, als nieuwe instroom. Op deze manier zou volgens berekening de helft van de populatie bestaan uit het zittende bestand en de helft uit nieuwe instroom. De verwachting was dat bij een streefgetal van 300 deelnemers en dus 150 nieuwe instromers er 6-9 maanden van werving van deelnemers nodig zijn om dit streefgetal te bereiken. Dit is ook gelukt. De laatste instromer dateert van 25 juni 2018. Deze deelnemer kan dus slechts 15 maanden meedoen aan het experiment terwijl de deelnemers uit het zittende bestand die op 1 oktober 2017 zijn ingestroomd de volle twee jaar deelnemen. Op 1 juni 2017 bestond de totale doelpopulatie, dus voor toepassing van de uitsluitingscriteria, uit 1726 mensen met een lopende uitkering in de Participatiewet. Door toepassing van de uitsluitingscriteria werden 255 bijstandsgerechtigden (15%) niet toegelaten tot het experiment.

Het resterende bestand van lopende bijstandsuitkeringen aan het begin van het experiment kwam daarmede in totaal per 1 juni 2017 op 1471 bijstandsgerechtigden. Deze zijn vervolgens benaderd om mee te doen aan het experiment. Vanaf 1 oktober 2017, de startdatum van het experiment, zijn ook

[3] Engels: Randomised Controlled Trial (RCT).

nieuwe instromers gevraagd mee te doen aan het experiment. De deelname van mensen uit het zittende bestand was belangrijk omdat bekend is op basis van de literatuur dat het grootste effect van de andere bejegening juist verwacht mocht worden bij deze groep (zie Muffels & Klink, 2017). Uit de bijstandspraktijk is immers bekend dat na verloop van tijd het aantal contacten met de cliënten met mensen uit het 'granieten bestand' erg gering is. Juist deze cliënten blijken baat te hebben bij 'extra aandacht'. Om het gewenste aantal deelnemers te bepalen is in maart 2017 een poweranalyse uitgevoerd[4]. Daarbij is uitgegaan van een power van 95%. De uitkomst daarvan was dat een aantal van 100 deelnemers per treatment groep gewenst was, dus bij drie groepen 300 in totaal. In totaal hebben zich tot 25 juni 2018, de dag waarop de laatste deelnemer is geworven, 380 mensen aangemeld waarvan 345 voldeden aan de toelatingscriteria maar waarvan 44 zich weer hebben afgemeld. Deze 44 afmelders deden zich uitsluitend voor bij het zittende bestand. Dat zijn alle mensen met een uitkering die is gestart voor 1 oktober 2017 waardoor er netto 301 deelnemers zijn waarvan 44% afkomstig is uit het zittende bestand (131) en 56% (170) uit de nieuwe instroom.

Afmeldingen

De deelnemers worden aselect toegewezen aan de drie groepen, maar deelnemers mogen na indeling alsnog weigeren deel te nemen aan het experiment. Ze kunnen dan echter niet worden geplaatst in een andere groep. Uiteindelijk zijn er 79 afmeldingen geregistreerd in Oss hetgeen 21% van het aantal aanmeldingen is (N=380). De meeste afmeldingen (drie van de vijf) komen uit het zittende bestand. Steeds is door de gemeente de reden van afmelding geregistreerd. Nadere analyse van de 79 afmeldingen leerde dat 35 aangemelde deelnemers bij aanmelding niet voldeden aan de toelatingscriteria zodat het aantal afmeldingen uitkomt op 44 bij een aantal van 345 aanmeldingen die voldoen aan de criteria, dat is 13%. Slechts 16 deelnemers hebben mogelijk afgezien van deelname vanwege de indeling in de groep of omdat men het nut van deelname niet in zag. Dat is slechts 4,2% van het totaal van 380 aanmeldingen. De andere afmeldingen zijn ontstaan doordat men uiteindelijk toch geen uitkering heeft gekregen (n=14 of 3,7%) of dat men niet tot de doelpopulatie behoorde doordat men taalproblemen of ernstige fysieke en mentale gezondheidsproblemen had (n=23 ofwel 6%). Daarover was op voorhand geen of onvoldoende informatie beschikbaar en kon slechts worden vastgesteld tijdens het eerste gesprek met de cliënt. Het resterende deel is uitgestroomd vanwege uitstroom naar betaald werk voordat de toestemmingsverklaring was getekend (n=14), studie (n=4), of vanwege detentie, verhuizing (n=2) of stopzetting van de uitkering door toepassing van de rechtmatigheidstoets (n=4). Soms belden mensen zelf om die reden af.

2.3 Onderzoekpopulatie: deelnemers in de verschillende treatments

In Tabel 2.3.1 wordt weergegeven in hoeverre de afmeldingen al dan niet gelijkelijk zijn verdeeld over de treatments. Tevens geeft de tabel de bereikte aantallen weer voor elk van de drie groepen inclusief en exclusief de 44 afmelders. De verwachting was dat het meer moeite kost om mensen in de

[4] De power geeft eenvoudig gezegd de kans weer dat een werkelijk verschil tussen de interventiegroepen ook daadwerkelijk wordt gevonden. Een poweranalyse is dus bedoeld om na te gaan welke aantallen deelnemers nodig zijn in een experiment om met een bepaalde kans een werkelijk verschil te vinden. Deze kans hebben we op 95% gesteld. Dat betekent dat we met deze aantallen met 95% betrouwbaarheid of zekerheid een verschil vinden als dat verschil in werkelijkheid ook bestaat.

standaardgroep te krijgen en te behouden aangezien zij in beginsel weinig baat hebben bij deelname. Dit bleek uiteindelijk erg mee te vallen. De netto aantallen, dus na aftrek van de afmeldingen, zijn goed gespreid over de drie interventiegroepen.

Tabel 2.3.1: Deelnemers Oss per 'treatment groep', afmeldingen en type uitstroom per 1 oktober 2019

	Deelnemers	afmeldingen	netto aantal	Uitstroom werk	Uitstroom overig
Zelf aan het stuur (eigen regie)	111	13	98	38 (32,2%)	24 (21,6%)
Op maat/intensieve begeleiding	115	14	101	33 (28,7%)	39 (33,9%)
Standaardgroep	119	17	102	35 (29,4%)	30 (25,2%)
Totaal	345	44	301 (100%)	106 (30,7%)	93 (27.0%)
in %	100.0%	12.7%	87.0%	31.0%	27.0%

Bron: Gegevens gemeente Oss, Tilburg University controlebestand en BUS-gegevens, 1 oktober 2019

Het succes van de start van het experiment en de werving kan voor een belangrijk deel verklaard worden door de persoonlijke motivatie, enthousiasme en inzet van de medewerkers van de gemeente Oss (projectleiders, procesanalist, consulenten, afdelingshoofd, programmamanager, politieke steun vanuit college i.c. de wethouder). Wat goed heeft gewerkt is een intensieve persoonlijke benadering waarin de cliënt persoonlijk werd geënthousiasmeerd voor deelname hetgeen vanuit het gezichtspunt van het aanwezige 'wantrouwen' van de cliënt ten opzichte van de gemeente geen sinecure is.

Representativiteit

Doordat de deelnemers aselect zijn toegewezen aan de drie interventiegroepen verwachten we geen grote verschillen tussen de deelnemers en de doelpopulatie waaruit de steekproef is getrokken. Mensen moeten zich evenwel aanmelden waardoor er zelfselectie zou kunnen ontstaan. Daarnaast kan zelfselectie ontstaan doordat mensen met bepaalde kenmerken zich na de aanmelding alsnog afmelden dan andere. Het zou kunnen zijn dat lager opgeleiden of ouderen boven 55 jaar doordat ze meer moeite hebben met invulling van online vragenlijsten ondervertegenwoordigd zijn. Voorts verwachten we dat mensen met weinig kansen op de arbeidsmarkt minder genegen zijn zich aan te melden, veelal betreft dat mensen die reeds lang in de bijstand zitten en tot het zogenaamde 'granieten bestand' behoren. In de volgende drie tabellen worden de deelnemers vergeleken met de referentiegroep (dat is de populatie van bijstandsgerechtigden die voldoet aan de toelatingscriteria minus degenen die zich hebben aangemeld als deelnemer). Daarbij hebben we gebruik gemaakt van de BUS gegevens op individueel niveau bij de start van het experiment. Voor het zittende bestand zijn dat gegevens per 30 september 2017. Voor de nieuwe instromers zijn dat de gegevens ultimo van de maand die zo dicht mogelijk ligt bij de start van het experiment veelal de maand ervoor. De informatie beschikbaar in de BUS betreft gegevens zoals geslacht, leeftijd, hoogst afgeronde opleidingsniveau, duur uitkering, nationaliteit, leefvorm en huisvestingssituatie. De referentiegroep omvat 1503 mensen waartoe mensen behoren met een lopende uitkering per 1 oktober 2017 of mensen die tussen 1 oktober en 1 juli 2018 zijn ingestroomd in de bijstand maar hebben besloten niet deel te nemen aan het experiment. De drie treatment groepen betreffen alle

345 mensen die zich hebben aangemeld voor het experiment dus inclusief de afmeldingen. Voor 2 aanmeldingen ontbraken BUS gegevens.

Nadere analyse (zie Tabel 2.3.2) laat zien dat er weinig verschillen zijn tussen de twee treatment groepen en de standaardgroep. Vanwege het RCT design zijn deelnemers aselect toegedeeld aan deze drie groepen juist om de groepen qua samenstelling zoveel mogelijk vergelijkbaar te maken. Dat wordt randomisatie genoemd. Deze lijkt dus redelijk goed gelukt te zijn.

Tabel 2.3.2 Representativiteit onderzoekspopulatie (inclusief afmelder) naar achtergrondkenmerken bij start van experiment, Oss, 2017-2018 (in %)

		-1	-2	-3	-4	-5
Achtergrondkenmerken bij start (BUS)	**Categorie**	**Deelnemers**	**Eigen regie**	**Intensieve begeleiding**	**Standaard-groep**	**Referentie-groep**
Geslacht	Man	50.6	47.8	51.8	52.1	49
	Vrouw	49.4	52.2	48.2	47.9	51
Leeftijd	gemiddeld	40.3	40.5	40.5	42.5	43.5
	<27 jaar	14.6	14.4	14.9	14.5	16.3
	27-34 jaar	25.2	25.2	28.1	22.2	13.5
	35-44 jaar	19.6	18.9	20.2	19.7	18.7
	45-54 jaar	25.4	24.3	25.4	26.5	24.7
	55 jaar (tot 65)	15.2	17.1	11.4	17.1	26.9
Opleidingsniveau	Laag	50.1	50	57.4	43.2	49.3
	Midden	26.6	29.6	24.1	26.1	16.8
	Hoog	10.1	9.9	8.1	8.7	4
	Geen opl/onbekend	13.2	13	10.4	16.2	29.8
Migratieachtergrond	Buitenlands	7	4.5	9.7	6.8	19.4
	Nederlands	93	95.5	90.3	91,2	80.6
Leefvorm	Alleenstaande ouder	20.8	22.5	21.2	18.8	15.4
	Alleenstaande	69	70.3	67.5	69.2	67.1
	Samenwonend	10.2	7.2	11.4	12	17.5
Bijstandsduur	Tot 1 jaar	48.1	50.5	50.4	43.7	26.8
	1 tot 3 jaar	29.9	30.6	28.7	30.3	25.5
	3 jaar of meer	22	18.9	20.9	26.1	47.8
Aantal		343	111	115	117	1503

Bron: BUS-gegevens Oss, 1 oktober 2019
Noot: Achtergrondkenmerken van deelnemers in de onderzoeksgroep (kolom 4) en in de twee experimentgroepen (kolom 2-3). T1: Eigen regie-Zelf aan het stuur; T2: Intensieve begeleiding-Op maat; Controlegroep=Standaardgroep; Referentiegroep (niet-deelnemers).

Door selectie (toelatingscriteria) en zelfselectie in het experiment (wie melden zich aan en wie melden zich af) zijn er evenwel verschillen ontstaan tussen de experimentgroepen en de referentiegroep, alle mensen die niet zich niet hebben aangemeld of deelnemen aan het experiment. Verschillen blijken te zijn ontstaan in verband met de duur van de uitkering, enigszins met leeftijd (in de drie treatmentgroepen zijn ouderen ondervertegenwoordigd en jongeren oververtegenwoordigd), mannen zijn licht oververtegenwoordigd evenals mensen met een academische opleiding. Vooral naar duur is echter

sprake van een sterke oververtegenwoordiging van korte duren en ondervertegenwoordiging van lange duren (>3 jaar). De bijstandsgerechtigden met duren van 1-3 jaar blijken daarentegen evenredig te zijn vertegenwoordigd in de treatment groepen Dit betekent dat een vergelijking van de uitkomstmaten en zeker ook de uitstroomkansen naar werk tussen de vier groepen die sterk afhankelijk is van de duur van de uitkering, rekening moet houden met deze onder -en oververtegenwoordiging naar vooral duur van de uitkering maar ook naar nationaliteit voorafgaand aan het experiment. Dit hebben we gedaan met behulp van logistische regressie en matching technieken. In de logistische regressie voor de uitstroomkans naar betaald werk corrigeren we voor verschillen in duur door de duur bij de start van het experiment op te nemen in de regressievergelijking. In de matching procedure gaan we op zoek naar mensen in de bijstandspopulatie van niet-deelnemers die zoveel mogelijk lijken op mensen die aan het onderzoek deelnemen en dus ook qua duur uitkering, leeftijd, geslacht, nationaliteit en dergelijke vergelijkbaar zijn.

2.4 Product- en procesevaluatie

Het voorliggende onderzoek bestaat zoals eerder aangegeven twee onderdelen: de product –en procesevaluatie (zie Muffels & van der Klink 2017). De *productevaluatie* is erop gericht na te gaan wat beter werkt voor de cliënt; de nieuwe regelluwe aanpak of de standaardaanpak met een strenge toetsing van verplichtingen en de toepassing van maatregelen en kortingen. Over een periode van 2 jaar zijn deelnemers aan het experiment gevolgd. In het onderzoek wordt niet alleen gekeken naar of mensen eerder betaald werk vinden maar ook wat het met mensen doet, met hun vertrouwen, gezondheid, hun welbevinden, hun zelfvertrouwen en hun zelfredzaamheid. Daartoe vulden cliënten aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in. De eerste noemen we de nulmeting en de laatste de eindmeting. De uitkomsten van de analyses op deze vragenlijsten zullen een antwoord geven op eerder gestelde onderzoeksvragen. Wanneer deelnemers uitstromen naar werk of bijvoorbeeld verhuizen, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. De bereidheid om dat te doen zal evenwel naar verwachting geringer zijn bij de uitstromers waardoor de non-response naar verwachting zal toenemen naarmate meer mensen uitstromen. In het onderzoek worden degenen die zich hebben afgemeld na de start van het experiment ook na afmelding administratief gevolgd. Dat gebeurt via de gemeentelijke administratie (BUS) alsook door koppeling van de gegevens van deze oorspronkelijke deelnemers aan de bestanden van het CBS. Dit is voor een valide en zuivere meting van de uitkomsten van het RCT experiment noodzakelijk. De wijze van gegevensverzameling is getoetst aan de AVG privacy regels en aan de ethische procedures voor 'informed consent' welke gelden voor wetenschappelijk onderzoek. De persoonsgegevens (NAW) zijn door de gemeente geanonimiseerd met behulp van een anoniem identificatienummer waarbij alleen de gemeente en het CBS toegang tot de persoonsgegevens hebben. Daartoe is een bewerkingsovereenkomst met de gemeente gesloten.

De *procesevaluatie* is erop gericht na te gaan hoe de projectleider(s) en consulenten/ (werk)coaches als professional het proces van implementatie van het experiment in de gemeente percipiëren en waarderen, en welke meer -of minderwaarde de consulenten/coaches ervaren van de andere bejegening van de cliënt voor de uitoefening van hun taken in de nieuwe aanpak. Daartoe zijn er regelmatig focusgroep gesprekken gehouden met de projectleider en de coaches. In totaal zijn er 4 focusgroep gesprekken gehouden. Het eerste gesprek vond plaats op 6 december 2017 nadat het experiment enkele maanden van start was

gegaan en het eerste gesprek met de cliënt is gehouden. Het tweede gesprek heeft op 3 oktober 2018 plaatsgevonden, het derde gesprek volgde op 12 juni 2019 en het vierde gesprek op 30 oktober 2019. Ten tweede vulde de coach aan het begin van het experiment, na 1 jaar en na 2 jaar een online vragenlijst in over de afspraken en bevindingen van de coaching gesprekken. Tezamen met het eerste focusgroep gesprek vormt dit de nulmeting van het procesevaluatiedeel. Ten slotte hebben de consulenten aan het einde van het experiment (november 2019) een afsluitende vragenlijst ingevuld over hun eigen ervaringen gedurende het experiment. Door het uitvoeren van deze procesevaluatie kan inzicht worden verkregen in de effectiviteit van de interventie, maar ook de mate waarin ineffectiviteit is terug te voeren tot een 'theory failure' (beleidstheorie) of een 'implementation failure' (uitvoering). In de volgende paragrafen zal de dataverzameling van zowel de product- als procesevaluatie worden toegelicht.

2.5 Dataverzameling product- en procesevaluatie

2.5.1 Productevaluatie: vragenlijsten

De vragenlijsten bevatten vragen over de onderwerpen gezondheid en welbevinden, zelfredzaamheid, zoek- en sollicitatiegedrag, ervaren keuzemogelijkheden (capabilities), sociale participatie, sociaal vertrouwen, vertrouwen in overheid en financiële situatie en armoede. Wanneer deelnemers uitstromen, bijvoorbeeld naar werk, blijven ze deelnemer en worden ze gevraagd de vragenlijsten in te vullen. In de praktijk blijkt dat ze zich dan sneller afmelden. Het gevolg hiervan is dat de respons afneemt naarmate het experiment vordert. Het doel van de eerste vragenlijst (de zogenaamde nulmeting) was om de situatie van de cliënt vast te stellen aan het begin van het experiment. Het invullen ervan gebeurde tijdens of direct na het eerste gesprek met de deelnemer aan het experiment. De tweede en derde vragenlijst (ook wel eerste en tweede vervolgmeting) zijn 1 jaar en 2 jaar na de start van de begeleiding verstuurd. Omdat er tot 1 juli 2018 nieuwe deelnemers zijn ingestroomd in het experiment, gebeurde het versturen van de vervolgvragenlijsten in tranches. Hierbij werden in drie tranches de deelnemers in volgorde van aanmelding geselecteerd voor invulling van de tweede en derde vragenlijst. Er is gekozen voor twee vervolgvragenlijsten om effecten over de tijd heen in beeld te kunnen brengen. Deelnemers in de referentiegroep zijn niet benaderd voor een vragenlijst.
De deelname aan het experiment is zoals eerder opgemerkt vrijwillig, men mag te allen tijde het experiment weer verlaten maar er wordt wel van een deelnemer verwacht dat deze meedoet aan het onderzoek en de vragenlijsten zorgvuldig invult. Daarover zijn afspraken gemaakt met de deelnemer, hoewel het in de praktijk soms moeilijk is gebleken om een deelnemer aan deze afspraak te houden.

Invulling deelnemer vragenlijst (productevaluatie)
Een persoonlijke benadering bleek uitermate belangrijk om deelnemers te stimuleren de online vragenlijst voor de nulmeting in te vullen. Zoals hiervoor toegelicht heeft dat veel energie gevraagd. Het vraagt overredingskracht en uitleg van de consulent om de cliënt te overtuigen de vragenlijst volledig in te vullen. Vanaf de start van het experiment op 1 oktober 2017 jl. hebben de deelnemers aan het experiment de online vragenlijst ingevuld. De bereidheid bij de deelnemers om de tweede en vooral ook de derde vragenlijst in te vullen bleek sterk te zijn afgenomen na de nulmeting ondanks een intensieve campagne gericht op invulling door de deelnemer opgezet door de gemeente Oss. Veel vragen gaan over werk zoeken en vinden hetgeen voor de mensen die na 2 jaar nog geen werk hebben gevonden ondanks de

andere bejegening en begeleiding geen haalbare kaart is waardoor de motivatie om de tweede en derde vragenlijst in te vullen veel geringer bleek. De gemeente Oss heeft met het oog op het belang voor het onderzoek van een goede invulling alles op alles gezet om mensen te overreden de vragenlijst in te vullen. Dat heeft uiteindelijk geresulteerd in een respons van ruim 62% voor de beide laatste deelnemersvragenlijsten tegenover 85% voor meting 1. Uiteindelijk is een respons bereikt van bijna 70% hetgeen bij deze als 'moeilijk bereikbare' onderzoeksgroep bekend staande groep als 'zeer goed' mag worden bestempeld zeker indien in beschouwing wordt genomen dat 58% van de deelnemers na 2 jaar om vele redenen is uitgestroomd uit de bijstand vanwege verhuizing, herzien recht op uitkering, overgang naar andere uitkeringen maar ook door het vinden van werk. In totaal heeft 31% in de periode van 2 jaar werk gevonden (inclusief uitstroom naar zelfstandigheid) en heeft 27% van de deelnemers de uitkering beëindigd om andere redenen. De 627 door hen "online" ingevulde vragenlijsten[5] zijn door de onderzoekers nader geanalyseerd. Voordat we de resultaten verder bespreken wordt hieronder inzicht geven in de bereikte aantallen ingevulde vragenlijsten per treatment groep.

Procesevaluatie: consulenten vragenlijsten

Ten behoeve van de procesevaluatie zijn er ook periodiek vragenlijsten ingevuld door de consulenten die de deelnemers in begeleiding hebben. Voor elke deelnemer die de consulent in begeleiding heeft, vult een consulent een vragenlijst in. Net als bij de deelnemersvragenlijsten zijn deze consulent vragenlijsten drie keer afgenomen: aan het begin, na 1 jaar en na 2 jaar begeleiding (maar wel in volgorde van aanmelding van een deelnemer). In deze vragenlijsten zijn vragen gesteld over het proces van begeleiding, zoals de doelen die zijn ingezet in overleg met de deelnemer, de activiteiten die zijn ondernomen en de trajecten die zijn aangevraagd, de motivatie van de deelnemer voor de begeleiding door de consulent, een inschatting van de consulent over de kansen op de arbeidsmarkt van de deelnemer en de leereffecten van het experiment voor zowel de consulent als de deelnemer. Wanneer een deelnemer uitstroomt, bijvoorbeeld vanwege het vinden van betaald werk of vanwege verhuizing, dan is bij ontbreken van contact het lastig voor de consulent om de vragenlijst in te vullen.

Tabel 2.5.1: Aantallen ingevulde vragenlijsten voor de deelnemers en consulenten, Oss, 2017-2019

	Deelnemers					Consulenten		
Groepen	CV1	CV2	CV3	Totaal	PCEV1	PCEV2	PCEV3	Totaal
Eigen Regie	89	80	64	233	98	98	97	293
Intensieve Begeleiding	89	70	67	226	101	101	101	303
Standaardgroep	75	41	49	165	88	100	105	293
Totaal (N=301)	253	191	180	624	287	299	303	889
Bruto response %	84%	63%	60%	69%	95,3%	99,3%	100%	98%

Noot: CV=Clientvragenlijsten; PCEV=Procesevaluatievragenlijsten

Bron: Gegevens vragenlijsten, Tilburg University, 1 oktober 2019

[5] De meeste deelnemers hebben de vragenlijst online ingevuld. De deelnemers kregen een door de Universiteit gemaakte persoonlijke weblink opgestuurd waarmee zij de vragenlijst in de thuissituatie online konden invullen. Alleen wanneer de deelnemer de voorkeur gaf aan invulling tijdens of na een gesprek met de consulent bij de gemeente werd die mogelijkheid gefaciliteerd. Sommigen hebben een schriftelijke vragenlijst ingevuld die vervolgens door de onderzoekers zelf online is ingevuld.

In nagenoeg alle gevallen zijn desondanks de consulenten in staat gebleken de vragenlijst in te vullen. Dit is een uitstekend resultaat en te danken aan de proactieve houding en inzet van de consulenten en de projectleiders. Dit is consciëntieus door hen gedaan. In totaal zijn 889 vragenlijsten ingevuld, hetgeen een respons percentage betekent van 98%. In Tabel 2.5.1 wordt een overzicht gegeven van de bereikte aantallen ingevulde vragenlijsten voor de nulmeting en de procesevaluatie alsmede het response percentage.

2.5.2 Administratieve BUS data

Zoals opgemerkt zullen in het voorjaar van 2020 de gegevens geupload worden naar het CBS om ook vergelijkingen met andere gemeenten mogelijk te maken. In deze rapportage maken we gebruik van de gemeentelijke BUS bestanden om bijvoorbeeld de, voorover deze aanwezig zijn, uitstroomkansen naar werk te bepalen maar ook om te corrigeren voor compositieverschillen tussen de treatment groepen voor zover deze aanwezig zijn. Deze bestanden bevatten informatie over demografische kenmerken van de bijstandspopulatie, de reden van beëindiging van de uitkering vanwege uitstroom uit de uitkering, de arbeidsgeschiedenis, de loonwaarde, het aantal en type contactmomenten en de ondernomen re-integratieactiviteiten afkomstig uit de registratiesystemen van de gemeente Oss. Daarnaast worden de data geupload naar het Centraal Bureau voor de Statistiek (CBS) om deze te verrijken met gegevens over werk (loopbaan) en inkomen, uitkeringen en andere achtergrondkenmerken.[6] Hierdoor is het mogelijk om de loopbanen van de deelnemers te reconstrueren en ook gedeeltelijke uitstroom naar flexibele banen, (kleine) deeltijdbanen en vaste banen te onderzoeken. Door het gebruik van CBS-data zijn de gegevens tussen de experimenten met de Participatiewet bovendien beter vergelijkbaar.

3 Resultaten product -en procesevaluatie

3.1 Productevaluatie en uitkomstmaten

In de productevaluatie gaat het niet alleen om de uitkomstmaat uitstroom naar betaald werk maar ook om de verbetering van de gezondheid, het welbevinden, het zelfvertrouwen, de keuzemogelijkheden, de sociale participatie en de zelfredzaamheid van de cliënt. Eerst bespreken we de resultaten van de analyse van de uitkomstmaat volledige uitstroom naar betaald werk. Hieronder verstaan we dat bijstandsgerechtigden vanwege uitstroom naar betaald werk volledig onafhankelijk worden van de uitkering. Bij een loon ter hoogte van het minimumloon betekent dit dat men minimaal 27 uur gaat werken (70% van 38-40 uur) aangezien de uitkering voor een alleenstaande 70% bedraagt van het volwassen minimumloon.

[6] In alle gebruikte databestanden werden persoonsgegevens bij voorbaat door middel van een sleutel vervangen door een betekenisloos administratienummer (pseudonimiseren). Middels dit nummer kunnen databestanden uit verschillende bronnen en over de tijd heen aan elkaar worden gekoppeld. Omdat geen persoon toegang heeft tot zowel de databestanden als de sleutel, is informatie niet naar deelnemers herleidbaar.

Deeltijdse uitstroom en uitstroom naar tijdelijk werk

Er zijn in de BUS data uitsluitend gegevens beschikbaar over volledige uitstroom uit de uitkering maar niet over deeltijduitstroom of over uitstroom naar tijdelijk werk. De gemeenten beschikken over uitkeringsgegevens van mensen in de bijstand op grond waarvan een goede inschatting gemaakt zou kunnen worden van de aantallen mensen die deeltijds in plaats van voltijds uitstromen en voor hoeveel uren. Deze info kon niet worden gebruikt. In plaats daarvan zullen we in de definitieve rapportage gebruik maken van de gegevens in de CBS microdatabestanden. De data van Oss zullen daartoe worden geupload naar de CBS omgeving. De tweede en derde cliëntvragenlijst bevatten wel enige informatie van de deelnemer over zijn of haar situatie op het moment van invulling van de vragenlijst (zie Tabel 3.1.1). Deze geeft evenwel geen betrouwbaar beeld van de uitstroom naar deeltijd -of tijdelijk werk. Van de deelnemers is 1 en 2 jaar na de start bekend is of ze een voltijdse, deeltijdse of tijdelijke baan hebben of dat ze andere activiteiten hebben ondernomen zoals vrijwilligerswerk of een opleiding of studie. De respons in de tweede en derde cliëntvragenlijst bedroeg evenwel slechts 60% waardoor geen volledig beeld mogelijk is.

Tabel 3.1.1: Type sociale participatie activiteiten in vervolgmeting 1 en 2, Oss

Type activiteit	**CV-Meting 2**		**CV-Meting 3**	
	N=	**%**	**N=**	**%**
Voltijdse baan	21	11%	24	14%
Deeltijdse baan, uitzendwerk, enkele uren werk	38	20%	33	19%
Zelfstandige	2	1%	8	5%
Werkstage, proefplaatsing, overig werk (voorbereiding werk)	8	4%	7	4%
Opleiding, cursus of studie	7	4%	4	2%
Gedeeltelijke/volledige ao-uitkering	6	3%	8	5%
Sociale werkvoorziening	11	6%	10	6%
Vrijwilligerswerk/mantelzorg	21	11%	10	6%
Bijstand	71	38%	63	36%
Verhuisd	3	2%	5	3%
Ziek, overig	0	0%	5	3%
Totaal N=	188	100%	177	100%

Bron: Vragenlijsten eerste en tweede vervolgmeting deelnemers, TiU, 2019

Uit de gegevens in tabel 3.1.1 blijkt dat in de tweede en derde meting respectievelijk 25% tot 35% volledig is uitgestroomd (14+5+6) als werknemer (inclusief sociale werkvoorziening) of als zelfstandige, 4% werkt in een proefplaatsing of stage en 20% respectievelijk 19% werkt een aantal uren per week in een kleine baan, een uitzendbaan of een deeltijdbaan. Dat betekent dat ongeveer 1 op de drie à vier deelnemers naar eigen zeggen volledig aan het werk is gegaan in de tweejarige experimenteer periode. Daarnaast werkt een op de vijf in een kleine deeltijdbaan. De gunstige arbeidsmarktsituatie in combinatie met de inspanningen van de gemeente om het experiment goed uit te voeren zullen hier zeker debet aan zijn. Voor ongeveer 80% van degenen die werk vinden is dit een deeltijdse baan of uitzendwerk. Het aandeel

mensen dat vrijwilligerswerk verricht is respectievelijk 11% en 6% in meting 2 en 3. Voor degenen die werk hebben gevonden is gevraagd wat voor soort werk of contract dat was. Tabel 3. 1.2 laat zien dat in meting 2 en 3 een op de vier à vijf deelnemers die werk heeft gevonden een vast contract heeft gekregen (22-25%). Ongeveer 80% van degenen die werk hebben gevonden komen derhalve in tijdelijk banen terecht en slechts 7% gaat als zelfstandige aan de slag. Een op de vijf vindt een uitzendbaan (17%). In de analyses op de CBS data ten behoeve van de definitieve rapportage zal op basis van de inkomensinformatie in de CBS bestanden worden onderzocht in hoeverre er verschillen zijn in deeltijduitstroom en uitstroom naar flex tussen de twee treatmentgroepen en de standaardgroep en referentiegroep. Dat is met de beschikbare BUS-data vooralsnog niet mogelijk.

Tabel 3.1.2: Type contract in meting 1 en 2, Oss

Type contract	**CV-Meting 1**		**CV-Meting 2**	
Type contract indien werk	N=	%	N=	%
Vast contract (onbepaalde tijd)	14	25,5%	13	22,4%
Tijdelijke werk (contract voor bepaalde tijd)	27	49%	31	53,4%
Oproepkracht	4	7,3%	0	0
Uitzendkracht	8	14,5%	10	17,2%
Zelfstandige/vrij beroep	2	3,6%	4	6,9%
N=	55	100%	58	100%

Bron: Vragenlijst eerste vervolgmeting deelnemers, TiU, 2019

3.1.1 Resultaten: volledige uitstroom naar werk per treatmentgroep

Op basis van de BUS-gegevens is een analyse verricht op voltijdse (minimaal 27 uur) uitstroom naar werk en daarmee uitstroom uit de bijstand. Dit is gebeurd op basis van de registratie in de BUS van de reden van beëindiging van de uitkering. In hoeverre deze registratie een betrouwbare bron is om volledige uitstroom naar werk te achterhalen, is niet onderzocht. We nemen vooralsnog aan dat de administratie van de gemeente betrouwbare informatie verschaft over uitstroom uit de uitkering. Mensen waarvan de uitkering is beëindigd vanwege ziekte of arbeidsongeschiktheid zijn evenwel ook opgenomen in deze variabele waardoor een volledig zuivere meting van de uitstroom naar betaald werk niet goed mogelijk was en de uitstroompercentages licht zijn overschat. Vergelijking van de BUS-cijfers met de administratie van volledige uitstroom die door het Vertrouwensexperiment in Oss zelf is bijgehouden laat evenwel zien dat er slechts kleine verschillen zijn tussen beide registraties in de uitstroom naar betaald werk[7].

Peildatum en mutaties in de BUS-registratie

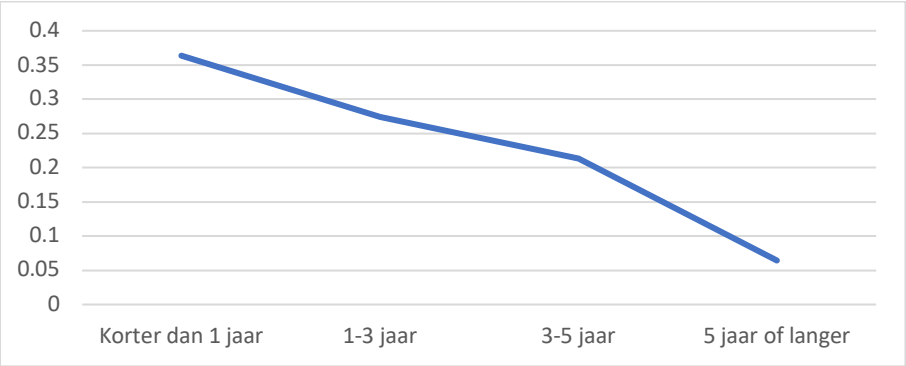
De peildatum van uitstroom is 1 oktober 2019. Alle beëindigingen van de bijstandsuitkering volgens de Participatiewet vanwege uitstroom naar ziekte arbeidsongeschiktheid of werk tot 1 oktober 2019 zijn in

[7] De cijfers van de gemeente zelf registreren een uitstroom van 99 voor de deelnemers exclusief afmeldingen terwijl wij op basis van de BUS-data per 1 oktober 2019 een uitstroom naar werk registreren van 96 voor alle deelnemers exclusief afmeldingen en 106 inclusief afmeldingen.

de analyses verwerkt. Dat betekent dat de observatie van de uitstroom naar werk een periode van twee jaar omvat. Hiervoor is gebruik gemaakt van de BUS-informatie ultimo oktober 2019 en in het bijzonder de datum van beëindiging van de uitkering welke op dat moment in de administratie is verwerkt. Gemeenten passen evenwel mutaties toe op de administratie tot 3 maanden na de eerste registratie. Dit heeft te maken met het herzien van het recht op uitkering, overgang naar andere typen bijstandsregelingen (bijvoorbeeld van alleenstaande naar samenwoning), en verandering in onder andere de persoonlijke inkomens -en vermogenssituatie van de bijstandsgerechtigde. De hier gebruikte data zijn dus voorlopige data van de gemeente Oss.

Volledige uitstroom naar werk

In de Tabellen 3.1.3 en 3.1.4 presenteren we de cijfers over de bereikte uitstroompercentages naar werk voor elk van de treatmentgroepen, de standaardgroep en de referentiegroep. Hierbij is nog niet gecorrigeerd voor compositieverschillen, ofwel verschillen in samenstelling tussen de treatment groepen. Eerder zagen we al dat kortere duren zijn oververtegenwoordigd en langere duren ondervertegenwoordigd. Dit vertekent de uitkomsten aangezien de kans op uitstroom sterk samenhangt met de duur van de uitkering. Daardoor zullen de uitstroompercentages hoger zijn dan wanneer er geen vertekening naar duur was opgetreden. Uit de regressieanalyse die we later presenteren zal blijken dat deze vertekening niet erg groot is. In Figuur 1 geven we de uitstroomkans weer voor de doelpopulatie dat zijn dus alle mensen die voldoen aan de toelatingscriteria, op 1 oktober 2017 een lopende uitkering hadden of tussen 1 oktober 2017 en 1 juli 2018 in de PW uitkering instroomden. Hiertoe behoren dus ook de 345 deelnemers aan het experiment inclusief de afmelder. Hier zien we een sterke en tot aan 5 jaar een bijna lineaire daling van de uitstroomkans naar werk bij volledige uitstroom uit de uitkering. Na 3-5 jaar neemt de kans op uitstroom nog sterker af. De gemiddelde duur in de uitkering bedraagt ongeveer 6 jaar in Oss.



Figuur 3.1.1. Uitstroomkans naar werk naar duur uitkering 2017-2019, Oss, N=1835

In Tabel 3.1.3 en 3.1.4 worden de ongecorrigeerde uitstroompercentages per treatmentgroep gegeven. Daarbij geeft tabel 3.1.3 de informatie inclusief de 44 afmelders terwijl tabel 3.1.4 de uitstroom weergeeft voor het netto aantal deelnemers exclusief afmeldingen. De resultaten gepresenteerd in Tabel 3.1.3 en 3.1.4 laten kleine effecten zien voor de vergelijking met de standaardgroep maar onverwacht grote effecten in vergelijking met de referentiegroep. Zoals gezegd zijn deze laatste cijfers vertekend naar duur en nationaliteit en moet daarvoor gecorrigeerd worden. Wel zien we dat (op basis van Tabel 3.1.3):

- De uitstroompercentages zijn gestegen vergeleken met eerdere cijfers (18% voor de totale populatie tussen 2017 en 2019 vergeleken met 15% in de periode 2015-2017). Hieraan is naar verwachting de gunstige arbeidsmarktsituatie mede debet;
- De ongecorrigeerde uitstroompercentages voor de deelnemers hoog zijn, namelijk 28% inclusief standaardgroep en zelfs 32% exclusief standaardgroep tegenover 15% voor de referentiegroep;

Tabel 3.1.3. Volledige uitstroom naar werk tussen 1 oktober 2017 en 1 oktober 2019 inclusief afmelders

Groepen	**Zittend bestand**		**Nieuwe instroom**		**Totaal**	
Treatments incl. afm.	N	%	N	%	N	%
Eigen regie	58	34.5%	56	39.3%	111	34.0%
Intensieve begeleiding	63	22.2%	61	37.7%	114	29.0%
Standaard groep	64	25.0%	51	35.3%	117	30.0%
Referentie groep	1360	14.0%	143	19.6%	1502	15.0%
Totaal	1545	16.0%	300	29.3%	1846	18.0%
Deelname per groep						
Deelnemers	185	27.0%	157	36.0%	342	28%

Noot: Zittend bestand (lopende uitkering per 1-10-2017); Nieuwe instroom tussen 1-10-2017 en 1-07-2018
Bron: BUS data gemeente Oss (uitstroom naar werk); Administratie VEX-Oss (deelnemers)

Tabel 3.1.4. Volledige uitstroom naar betaald werk tussen 1 oktober 2017 en 1 oktober 2019, exclusief afmelders

Groepen	**Zittend bestand**		**Nieuwe instroom**		**Totaal**	
Treatments excl. afm.	N	%	N	%	N	%
Eigen regie	50	38.0%	48	38.0%	98	38.0%
Intensieve begeleiding	51	20.0%	49	39.0%	100	29.0%
Standaardgroep	52	23.0%	49	37.0%	101	30.0%
Ref groep	1359	14.0%	143	19.6%	1502	15.0%
Totaal	1502	15.0%	299	29.3%	1801	18.0%
Deelnemers						
Deelnemers	153	27.0%	146	38.0%	299	32.0%

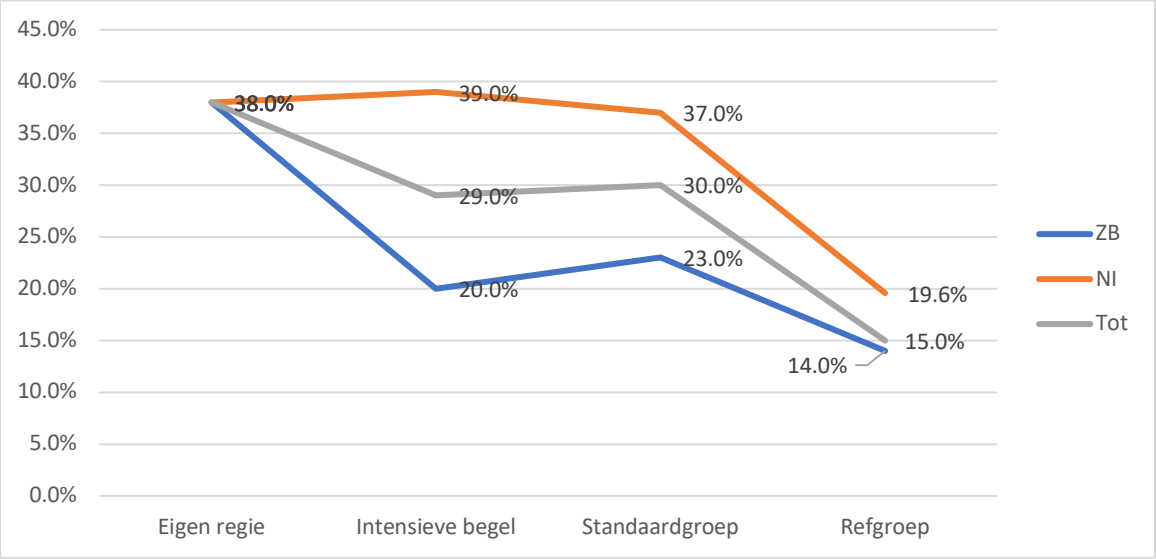
Noot: Zittend bestand (lopende uitkering per 1-10-2017); Nieuwe instroom tussen 1-10-2017 en 1-07-2018

Bron: BUS data gemeente Oss (uitstroom naar werk); Administratie VEX-Oss (deelnemers)

- De uitstroomkansen voor de nieuwe instromers (tussen 1 oktober 2017 en 1 juli 2018) aanmerkelijk groter zijn dan voor de mensen in het zittende bestand;
- De verschillen tussen de drie treatments kleiner zijn dan verwacht en betrekkelijk gering zijn;

- Het verschil met de standaardgroep klein is en bij de groep intensieve begeleiding voor het zittende bestand negatief;
- Het verschil met de referentiegroep daarentegen positief en erg groot is, 31% bij de twee treatments en 15% voor de referentiegroep (verschil is 16%).

In Figuur 3.1.2 geven we de resultaten weer voor de verschillende groepen uit tabel 3.1.4 waarin de afmelder niet zijn meegenomen. De resultaten wijken niet veel af van die in Tabel 3.1.3.



Figuur 3.1.2. Uitstroom naar betaald werk voor zittend bestand en nieuwe instroom naar treatmentgroep en referentiegroep, Oss, 2017-2019
Noot: ZB=zittend bestand; NI=nieuwe instroom
Bron: Gegevens BUS en Vertrouwensexperiment Oss; bewerking door TiU/Tranzo

De grijze lijn geeft het gemiddelde weer over alle groepen. De uitstroomkansen voor de groep nieuwe instromers zijn voor alle groepen hoger dan voor mensen in het zittende bestand. Vergelijken we de treatments dan zien we dat de ontheffingsgroep (eigen regie) een hogere uitstroom heeft gerealiseerd dan de intensieve begeleidingsgroep bij het zittende bestand. Dit is verrassend omdat we hadden verwacht dat intensieve begeleiding beter tot zijn recht zou komen bij het zittende bestand en dat eigen regie beter tot zijn recht zou komen bij de nieuwe instromers met betere perspectieven op de arbeidsmarkt. Verder valt op dat de standaardgroep met de reguliere aanpak dezelfde uitstroom heeft gerealiseerd als de intensieve begeleidingsgroep. Alle groepen inclusief de standaardgroep realiseren een hogere uitstroom dan het geval is in de referentiegroep die evenals de standaardgroep de reguliere treatment heeft gekregen. De standaard treatment is kennelijk niet goed vergelijkbaar geweest met de standaard treatment van de mensen in de referentiegroep. Dit zou kunnen wijzen op het zogenaamde “Hawthorne” effect dat de standaard treatment onder invloed van het experiment en het onderzoek meer is gaan lijken op de aanpak in de andere treatments. Zoals gezegd zijn de uitstroomcijfers geflatteerd door de oververtegenwoordiging van korte duren. Hieronder bespreken we de methode en de uitkomsten daarvan die is gebruikt om daarvoor te corrigeren.

Modelmatige schatting van de uitstroomkans

Om te corrigeren voor compositieverschillen tussen de groepen ook wat duur betreft hebben de onderzoekers bij de analyse van voltijdse uitstroom naar betaald werk gebruik gemaakt van een analysemodel dat zoveel mogelijk vergelijkbaar is gemaakt met het model dat in het landelijk overleg van onderzoekers is overeengekomen. Dat LOEP-model zal later worden gebruikt voor de analyses op de CBS-data. Het LOEP-model is afgestemd op de beschikbare informatie in de CBS-data, niet alleen de BUS-data waarvan het CBS ook gebruik maakt, maar ook andere registratiedata die het CBS gebruikt. In het model dat wij gebruiken voor Oss zijn zoveel mogelijk dezelfde variabelen gebruikt als in het LOEP-model met dezelfde categoriseringen. Bij bepaalde variabelen bleek dat moeilijk te realiseren zoals bij opleiding aangezien de gemeenten hoogst behaalde opleiding bijvoorbeeld op een eigen manier registreren welke niet volledig vergelijkbaar is met die van het CBS op basis van andere bronnen. Hierna presenteren wij de uitkomsten van het model. Eerst worden in Tabel 3.15 de gebruikte variabelen weergegeven.

Tabel 3.1.5. Aantal, gemiddelde en standaardafwijking gebruikte variabelen in het analysemodel volledige uitstroom naar werk, Oss, 2017-2019

Naam model variabele	Omschrijving	Frequentie (N=)	Gemiddelde	Standaardafwijking
uitstrwerk	Reden van uitstroom (werk, ziekte, arbeidsongeschiktheid)	1845	0.178	0.383
ovexgroep4	Onderzoeksgroep (treatments, standaardgroep, referentiegroep)	1848	1.371	0.853
duuruitk_t0	Duur uitkering PW bij start experiment [1)]	1848	4.351	5.699
nkeeruitk_t0	Aantal keer uitkering voor start experiment	1836	1.549	1.596
lft_t0	Leeftijd bij start	1848	43	13.69
geslacht_t0	Geslacht (vrouw=1; man=0)	1848	0.51	0.5
buitlnat_t0	Nationaliteit (Nederlands of buitenlands)	1848	0.17	0.376
leefv_t0	Leefvorm (alleenstaande, alleenstaande ouder, samenwonend)	1848	1.997	0.571
huisv_t0	Huisvesting (eigenaar, huurder, inwonend, overig)	1768	1.37	0.785
nizb_t0	Zittend bestand of nieuwe instroom [1)]	1848	0.831	0.376

Noot: [1)] De helft van de mensen in de treatmentgroepen komt uit het zittende bestand (dus heeft een lopende uitkering per 1 oktober 2017) maar in de referentiegroep is dit percentage groter; 90% daarvan heeft een lopende uitkering per 1 oktober 2017. De variabele nkeeruitk_t0 was niet significant en is weggelaten uit het model.
Bron: BUS-data gemeente Oss, 2017-2019

De afhankelijke variabele is voltijdse uitstroom naar betaald werk. De onafhankelijke variabelen zijn gemeten in beginsel een maand voor de start van iedere deelnemer aan het experiment tenzij voor die maand geen informatie in de BUS-data beschikbaar was. Dan is de dichtstbijzijnde maand genomen

waarvoor wel informatie beschikbaar was. De gemiddelde duur bedraagt bij aanvang van het experiment meer dan 4 jaar (6 jaar minus 2 jaar duur experiment).

Resultaten

Tabel 3.1.6 geeft de resultaten van de logistische regressie weer. Het model verklaart 15% van de variantie (pseudo R^2) en is significant. Het model corrigeert zoals gezegd voor duurverschillen tussen de groepen maar ook voor andere compositieverschilen die te maken hebben met geslacht, leeftijd, nationaliteit, duur uitkering en dergelijke. We hebben tevens een model geschat voor de vergelijking tussen de drie groepen en dus niet met de referentiegroep en dan blijkt dat dit andere model significant is maar geen significante of betekenisvolle verschillen laat zien tussen de treatment groepen en de standaardgroep op deze uitkomstmaat.

Tabel 3.1.6. Resultaten logistische regressie: geschatte uitstroomkansen naar treatment groep, inclusief en exclusief afmeldingen, plus treatment-effecten in procentpunten

Totaal	**Inclusief afmeldingen**		**Exclusief afmeldingen**		**Effect t.o.v. controle groep**	**Effect t.o.v. referentie groep**
Treatment	N=	P-werk	N=	P-werk	P-werk	P-werk
Eigen regie	111	34.2%	98	34.2%	4.20% n.s.	19.80%***
Intensieve begeleiding	114	28.9%	100	30.4%	-1.00% n.s	14.50%***
Standaardgroep	117	29.9%	101	30.0%	Ref.	15.50%**
Referentie groep	1426	14.4%	1426	14.4%	-15.50%**	Ref.
Treatments	225	32.6%	197	31.5%	n.a.	18.17%
Deelnemers	342	31.6%	298	32.3%	1.73%	17.23%
N=1768						
Pseudo R^2=0.147 Wald chi2(22) Prob > chi2 *Contrast-analyse* ER-IB ER-SG	0.147 191.6 0.000 n.s. n.s					

Noot: Logit model met controlevariabelen leeftijd, nationaliteit, geslacht, duur uitkering bij start, leefvorm bij start, huisvesting bij start. ER=eigen regie; IB=Intensieve begeleiding; SG=Standaard/controlegroep.
Bron: Bewerking BUS data, Tilburg University, 2019.

Tabel 3.1.6 geeft de geschatte uitstroomkansen en de verschillen of effecten weer in deze kansen ten opzichte van de standaardgroep en de referentiegroep, in procentpunten. Het verschil van 1,7% tussen de beide treatments en de standaardgroep is niet significant en statistisch dus nul. De laatste kolom geeft de effecten weer van de twee treatments ten opzichte van de referentiegroep. Het verschil van 15%

tussen standaardgroep en referentiegroep is wel significant of betekenisvol en positief ten gunste van de standaardgroep. Opvallend is dat ten opzichte van de ruwe data het effect van de treatments in procentpunten niet is afgenomen (17%), dus ondanks de correctie voor de compositieverschillen waaronder de duur.

Contrasten

Voor het uitgebreidere model met referentiegroep hebben we na schating van het logistische regressiemodel een contrast analyse uitgevoerd op de variabele treatment en de interactie tussen treatment en nieuwe instroom versus zittend bestand om te bezien of er significante of betekenisvolle verschillen tussen de drie treatments zijn. Deze blijken er niet te zijn. Er is dus geen verschil gevonden tussen de intensieve begeleidingsgroep en de eigen regiegroep wat betreft uitstroom naar werk.

3.1.2 De invloed van duur nader onderzocht: matching modellen

De vraag stelt zich nochtans of het model in staat is om op correcte wijze te corrigeren voor de verschillen in duren. Het effect van duur in het model met de referentiegroep als vergelijkingsgroep was sterk negatief en significant op 95% niveau. Korte en lange duren waren over -en ondervertegenwoordigd maar duren van 1 tot 3 jaar waren goed vertegenwoordigd in de treatment groepen. Indien we de geschatte kansen nu vergelijken naar duur en treatmentgroep (Fig. 3.1.3) dan blijven de verschillen tussen de treatment groepen maar ook van de standaardgroep met de referentiegroep overeind ook voor de duren van 1-3 jaar.

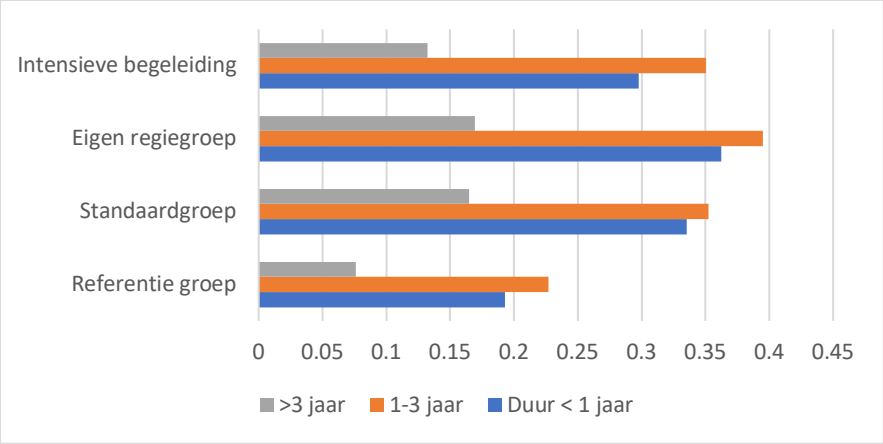


Fig. 3.1.3: Geschatte uitstroomkansen naar voltijds werk naar duur en treatment groep

Rol van consulent

Kennelijk speelt de duur minder een rol dan andere factoren die het verschil kunnen verklaren. Een van die factoren is wellicht de consulent zelf. Het eerdere logit model maar met de standaardgroep als vergelijkingsgroep en consulent nummer als verklarende variabele verklaart de verschillen in uitstroomkansen naar werk tussen de groepen twee keer zo goed als hetzelfde model zonder consulent nummer (13% in plaats van 7%). Van belang is mogelijk ook dat de consulenten in het experiment hebben gekozen voor deelname aan het experiment. Ze zijn om praktische redenen - hoewel het RCT design stricto

senso dat veronderstelt - niet aselect geselecteerd waardoor de motivatie voor het experiment een rol kan hebben gespeeld. Bovendien zijn de consulenten veelal 'dedicated' geweest aan een bepaalde treatment groep. Zelf zeggen ze dat ze binnen de gemeentelijke organisatie gezien werden als een soort 'elitegroep'. Wellicht zijn ze daardoor extra gemotiveerd aan de slag gegaan. Die extra motivatie kan wellicht deels ook de effecten verklaren.

Matching: correctie voor compositieverschillen

Ten tweede is er een 'matching' methode gebruikt om de verschillen tussen de groepen op het spoor te komen. Indien de populatie van deelnemers door (zelf)selectie niet representatief is op bepaalde aspecten, in dit geval duur en buitenlandse nationaliteit, dan kan in de data gezocht worden naar andere mensen die zoveel mogelijk lijken op de deelnemers maar die de treatment niet hebben gehad. We hebben deze methode gebruikt om de resultaten bij de deelnemers te vergelijken met een gematchte groep van mensen die qua duur en andere kenmerken op de deelnemers lijken. Op deze manier corrigeer je voor verschillen in waargenomen kenmerken zoals duur en nationaliteit. We hebben 'een-op-een' matching toegepast, dus voor iedere deelnemer is één andere persoon gezocht in de referentiegroep die op de deelnemer lijkt. Indien meer personen vergelijkbaar zijn wordt degene gekozen die het meest lijkt op de deelnemer. Dat wordt de 'dichtstbijzijnde buur' ('nearest neighbour') genoemd[8]. Voor de matching zijn zoveel mogelijk dezelfde variabelen gebruikt als in het regressiemodel. In beide methoden (nearest neighbour matching en propensity score matching) werd een betere matching bereikt met het weglaten van de variabele die aangeeft of men tot het zittende bestand of nieuwe instroom behoort en het kwadraat van de leeftijd. Ook leverde een andere opleidingsvariabele (met vier niveaus) betere 'matching' resultaten op. De vier niveaus zijn laag, midden, hoog en geen opleiding of opleiding onbekend[9]. In het geval van de wegingsmethode leverde een indeling in drie duurklassen en vijf leeftijdsklassen betere 'matching' resultaten op. De beide methoden laten toe om het effect van de interventie bij de treated zuiver vast te stellen (ATT-average treatment effect van de treated) maar ook het gemiddelde effect van de treatment voor de gehele populatie van deelnemers en niet-deelnemers. Voor de niet-deelnemers wordt dan op basis van hun kenmerken berekend wat het effect zou zijn geweest als zij ook de treatment hadden ondergaan (ATE-average treatment effect). Tot slot is het mogelijk om het effect te bepalen voor degenen die geen interventie hebben gehad ervan uitgaande dat zij die wel zouden hebben gehad (ATU -average treatment effect van de untreated). In onderzoek wordt veelal de ATT en ATE gebruikt om het effect van een interventie te meten. De ATT en ATE geven dan het verschil in geschatte uitstroomkans weer tussen de beide treatmentgroepen en de controlegroep. Omdat we geen verschil vonden van de twee treatmentgroepen met de standaardgroep wordt vooral ingegaan op de analyses gepresenteerd met de referentiegroep als vergelijkingsgroep[10]. We hebben de 'matching'

[8] De matching (zoeken en bepalen van vergelijkbare mensen) kan op basis van een wegingsfunctie van de verklarende variabelen (bijvoorbeeld middels nearest neighbour matching) of op basis van zogenaamde 'propensity scores' (de kans op treatment bepaald met behulp van een logit of probit model met verklarende variabelen) van de covariaten. Beide methoden zijn toegepast met vergelijkbare resultaten.

[9] Laag staat voor enkele klassen lager onderwijs of lager onderwijs; midden staat voor (voorbereidend) middelbaar algemeen vormend en beroepsonderwijs, hoog staat voor hoger beroeps -en wetenschappelijk onderwijs.

[10] De modellen zijn ook geschat met de standaardgroep als vergelijkingsgroep maar de resultaten bevestigen de uitkomsten van de logistische regressies; er werden geen significante verschillen gevonden tussen beide treatmentgroepen en de standaardgroep.

modellen geschat. Ten eerste wordt bevestigd dat de effecten niet significant zijn voor de vergelijking met de standaardgroep maar wel significant in vergelijking met de referentiegroep. Vergelijken we de twee treatments met de referentiegroep dan bedraagt de ATT 9.2%, en de ATE 9.4% in de wegingsmethode. In de 'propensity score' matching methode bedraagt de ATT 12.4% en de ATE 13.3%. Dat betekent dat er een 9% tot 13% hogere kans bij de treatmentgroepen is om volledig uit te stromen in vergelijking met de referentiegroep. Dat is een stuk lager dan we vonden bij de logit regressie resultaten (18%) maar niettemin een substantieel effect. De correctie voor compositieverschillen heeft zoals verwacht ertoe geleid dat het effect nu kleiner is.

De resultaten van de analyse worden in Tabel 3.1.7 weergegeven. Voordat we deze presenteren laten we zien dat het model inderdaad in staat is geweest om te corrigeren voor verschillen in duur. In Fig. 3.1.4 (links nn-matching; rechts ps-matching) staan in beide plots links de feitelijke uitstroomkansen naar werk en rechts de gecorrigeerde uitstroomkansen voor de controlegroep (referentiegroep hier) en de beide treatments dus na matching.

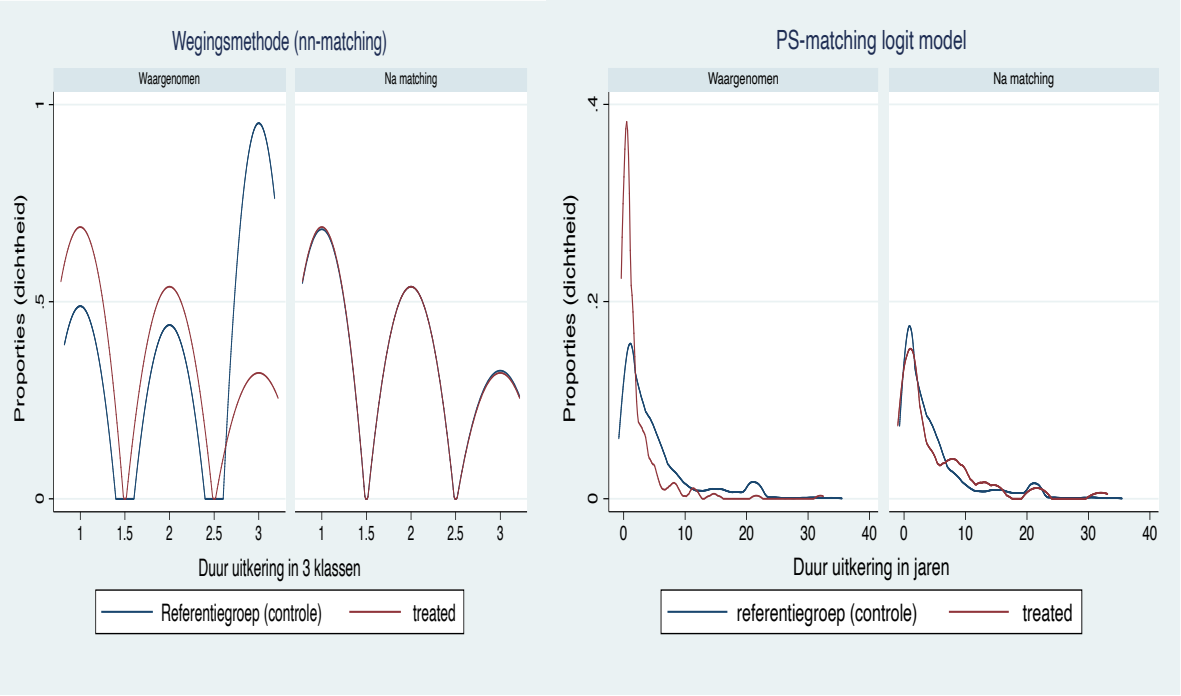


Fig. 3.1.4 Verschillen tussen controlegroep en treatments naar duur uitkering en matchingmethode

De feitelijke kansverdeling in Fig. 3.1.4 laat zien dat korte duren zijn oververtegenwoordigd en langere duren sterk ondervertegenwoordigd in de beide treatment groepen. Na matching zijn de uitstroomkansen van beide treatments naar duur nu voor 99% hetzelfde (variantie ratio) en het verschil in variantie ligt dicht bij 0[11]. Nochtans moet worden opgemerkt dat de referentiegroep in tegenstelling tot de standaardgroep niet is gerandomiseerd waardoor we niet zeker weten of de gevonden resultaten causaal geduid kunnen worden.

[11] In de technische bijlage van het later te verschijnen definitieve rapport zullen de resultaten in meer detail worden gepresenteerd.

Tabel 3.1.7 geeft de effecten weer voor de twee treatment groepen in vergelijking met de referentiegroep en de standaardgroep. Aangezien we eerder vonden dat er geen verschillen zijn tussen de groepen presenteren we de effecten voor beide groepen tezamen. Bevestigd wordt in deze cijfers dat in vergelijking met de standaardgroep de effecten van de beide treatments op de uitstroomkans niet significant of betekenisvol zijn en daardoor statistisch nul. In vergelijking met de referentiegroep zijn ze echter -ook na correctie op basis van matching- betekenisvol maar iets kleiner.

Tabel 3.1.7. Resultaten matching modellen uitstroomkans werk (voltijds)

	Wegingsmodel (NN-matching)[1)]		PS-matching		
	ATT	ATE	ATT	ATE	N=
Verschil t.o.v. referentiegroep					
Treatments (eigen regie, intensieve begeleiding)	9.4%***	9.20%***	12.40%***	13.30%***	N=1651
Deelnemers (incl. standaardgroep)	11.6%***	12.80%***	13.40%***	11.30%***	N=1768
Verschil t.o.v. standaardgroep					
Treatments (eigen regie, intensieve begeleiding)	-2.10%	1.40%	7.6%%	4.40%	N=342
Model variabelen					
Duur uitkering op t0			x	x	
Duur uitkering t0 in 3 klassen (<1 jr.; 1-3jr;>3jr)	x	x			
Leeftijd op t0			x	x	
Leeftijd t0 in klassen (10-jaars)	x	x			
Geslacht (man, vrouw)	x	x	x	x	
Buitenlandse nationaliteit (Nederlands, buitenlands)	x	x	x	x	
Leefvorm (alleenstaand, samenwonend, overig)	x	x	x	x	
Huisvesting (eigenaar, huurder, inwonend)	x	x	x	x	
Interactie duur en leeftijd	x	x			
Interactie leeftijd en nationaliteit	x	x			
Bias (duur uitkering en nationaliteit)	x	x			

Noot 1: [1)] In het model voor de 'nearest neighbourhood' matching methode is een correctie opgenomen voor bias (vertekening) op duur uitkering en nationaliteit. Beide methoden zijn geschat met Stata 15.1.

Noot 2: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een nearest neighbour matching model versus een propensity score matching logit model (N=1768)

3.1.3 Conclusie uitstroom naar werk

Samenvattend, suggereren de resultaten op de uitkomstmaat van volledige uitstroom naar betaald werk dat er geen betekenisvolle verschillen zijn tussen de drie treatmentgroepen en dat het effect van elk van de treatments ten opzichte van de standaardgroep nul is. Vergelijken we de gerealiseerde uitstroomkansen van alle drie groepen met de groep van niet-deelnemers dan zijn de verschillen echter substantieel. Gemiddeld genomen zijn de ruwe of waargenomen uitstroomkansen voor de beide treatmentgroepen 18% en voor de drie groepen tezamen 17% hoger dan voor de groep van niet-deelnemers. Dat is een fors verschil. De aangemelde deelnemers zijn aselect toegewezen aan de verschillende treatments maar door zelfselectie kan de groep van deelnemers qua samenstelling afwijken van de populatie waaruit ze is getrokken. Dat blijkt ook het geval te zijn. Vooral naar duur uitkering en nationaliteit zijn er verschillen tussen de deelnemers en de doelpopulatie ontstaan waarvoor gecorrigeerd moet worden. Om dat te doen is gebruik gemaakt van twee matching methoden een matching wegingsmethode en een model matching methode. Met beide methoden bleek het mogelijk voldoende geschikte matches te vinden voor de deelnemers onder de niet-deelnemers (100%).

Op basis daarvan berekenden we dat er een gemiddeld treatment effect van 9 tot 13% bestaat voor de beide treatment groepen tezamen op basis van de beide matching methoden. De standaard treatment is in het proces van experimenteren mogelijk erg gaan lijken op de beide andere treatments. De uitstroompercentages zijn in ieder geval hoger dan in de referentiegroep en vergelijkbaar met de scores van de treatment groepen. De wijze van ondersteuning door de consulent lijkt van groot belang. Indien we voor de 342 deelnemers waarvoor we betrouwbare informatie hebben (inclusief dus de afmelders) een model schatten om de verschillen tussen de groepen te verklaren in uitstroomkansen dan blijkt dat meer dan 50% van de verklaarde verschillen (13%) te zijn verklaard door de persoon van de consulent.

3.1.4 Welbevinden, gezondheid, keuzevrijheid, sociale participatie, zelfredzaamheid, vertrouwen en financiële situatie (armoede)

In de drie vragenlijsten voor de deelnemers (cv-nulmeting, cv-vervolgmeting1 en cv-vervolgmeting 2) zijn vragen gesteld die beogen het experiment te evalueren op de uitkomstmaten die niet direct gerelateerd zijn aan betaald werk zoals welbevinden, subjectieve gezondheid (fysieke en mentale gezondheid), sociale participatie, sociaal en institutioneel vertrouwen, zelfredzaamheid, keuzemogelijkheden, financiële situatie, materiële deprivatie en armoede. In overleg met de andere onderzoekers in LOEP zijn uit een mogelijke lijst van indicatoren waarover informatie in de vragenlijsten beschikbaar was, er negen tot 10 geselecteerd die als leidraad zullen dienen bij de rapportage van de AMvB experimenten. Dat zijn welbevinden, subjectieve gezondheid, sociale participatie, ervaren keuzemogelijkheden (capabilities), zoekintensiteit, zelfredzaamheid, vertrouwen in de ander en in instituties, financiële stress en armoede en/of deprivatie. Deze zullen ook in dit rapport worden gepresenteerd. In de beide tussenrapportages voor Oss is over een aantal indicatoren reeds verslag gedaan maar is nog geen informatie gegeven over de vergelijking naar de treatmentgroepen. Deze indicatoren passeren achtereenvolgens de revue hier.

Analysemethode

We hebben nu voor elk van deze indicatoren een regressiemodel (OLS) geschat waarmee we de kans op de uitkomst bepalen waarbij we controleren voor compositieverschillen mogelijk veroorzaakt door verschillen in duur uitkering bij de start, leeftijd, geslacht en nationaliteit tussen de treatments. Daarvoor gebruiken we de BUS informatie over leeftijd, geslacht, opleiding, nationaliteit, leefvorm, huisvesting, duur uitkering bij de start en, aantal keer in uitkering en of men behoort tot het zittende bestand op 1 oktober 2017 of tot de nieuwe instroom die na 1 oktober 2017 is ingestroomd. De BUS opleidingsinformatie is aangevuld met de informatie in de vragenlijst over hoogst behaalde opleiding. De modelmatig gecorrigeerde of geschatte cijfers zullen we hierna presenteren en dus niet de feitelijk waargenomen cijfers. Het effect van de correctie is overigens erg klein omdat door de randomisering er weinig verschillen zijn tussen de drie groepen. In de rapportage van de analyse, wordt gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde gecorrigeerde scores voor de treatment groepen op een uitkomstvariabele worden getoond per meting en in totaal. Deze gemiddelden zijn geschat op basis van de OLS regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden staat in paragraaf 3.2.3. Of een effect betekenisvol is of niet en met welke mate van betrouwbaarheid wordt aangegeven met sterretjes; drie sterretjes staat voor 99% betrouwbaarheid twee sterretjes voor 95% en een sterretje voor 90% betrouwbaarheid. Er zijn steeds twee modellen geschat een zonder interactieeffect en een met interactieeffecten tussen treatment en meting (zie Bijlage A voor de regressieresultaten). De effecten in de randtotalen geven op basis van het model zonder interacties de effecten weer van verschillen over de tijd (rijtotalen onder) of van verschillen tussen elk van beide treatments en de standaardgroep gemiddeld over de tijd (kolomtotalen rechts). We zijn met name hierin geinteresseerd (zie ook Tabel A1.1 in Bijlage A). De sterretjes bij de scores in de andere cellen geven op basis van de schattingen in het interactiemodel weer in hoeverre er een betekenisvol effect is van de treatment ten opzichte van de standaardgroep in de nulmeting bovenop het effect dus dat er was aan het begin van het experiment (zie Tabel A1.2 in Bijlage A). Aangezien aan het begin de treatment de meeste effecten nul zijn (zoals ook vastgesteld) geven deze veelal de volledige treatment effecten weer. In het model zijn maast de treatment variabele controlevariabelen opgenomen maar tevens een variabele die de persoon van de consulent weergeeft.

Gezondheid en welbevinden

In de vragenlijsten zijn vragen gesteld die te maken hebben met de gezondheid -en welzijnsbeleving van de deelnemer. Het welbevinden van de deelnemer wordt gepeild aan de hand van vragen over de tevredenheid met het leven (schaal van 0 tot 10), de mate waarin men het leven betekenisvol of zinvol acht (schaal van 0 tot 10) en de mate waarin men zich gelukkig voelt (op een schaal van 1 tot 5)[12].

In Fig. 3.1.5 presenteren we de waargenomen (WB-links) en voor compositieverschillen gecorrigeerde scores (Pwb-rechts) op welbevinden (het gemiddelde van tevredenheid, geluk en betekenisvol leven). De waargenomen en gecorrigeerde scores op welbevinden wijken nauwelijks af hetgeen laat zien dat compositieverschillen dankzij de randomisering (aselecte toewijzing aan de groepen) betrekkelijk weinig

[12] Deze drie vragen zijn vervolgens gebruikt om een gemiddelde score te berekenen van welbevinden waarbij de schaal van geluk eerst omgecodeerd werd naar 0 tot 10.

invloed hebben. In de nulmeting zijn de feitelijke scores op welbevinden nagenoeg gelijk om vervolgens toe te nemen in de metingen daarna. In de nulmeting is er dus nog geen verschil tussen de groepen hetgeen erop duidt dat er geen zelfselectie op welbevinden heeft plaatsgevonden voor het experiment. De zwarte lijn geeft de ontwikkeling in de tijd weer voor de drie groepen tezamen.

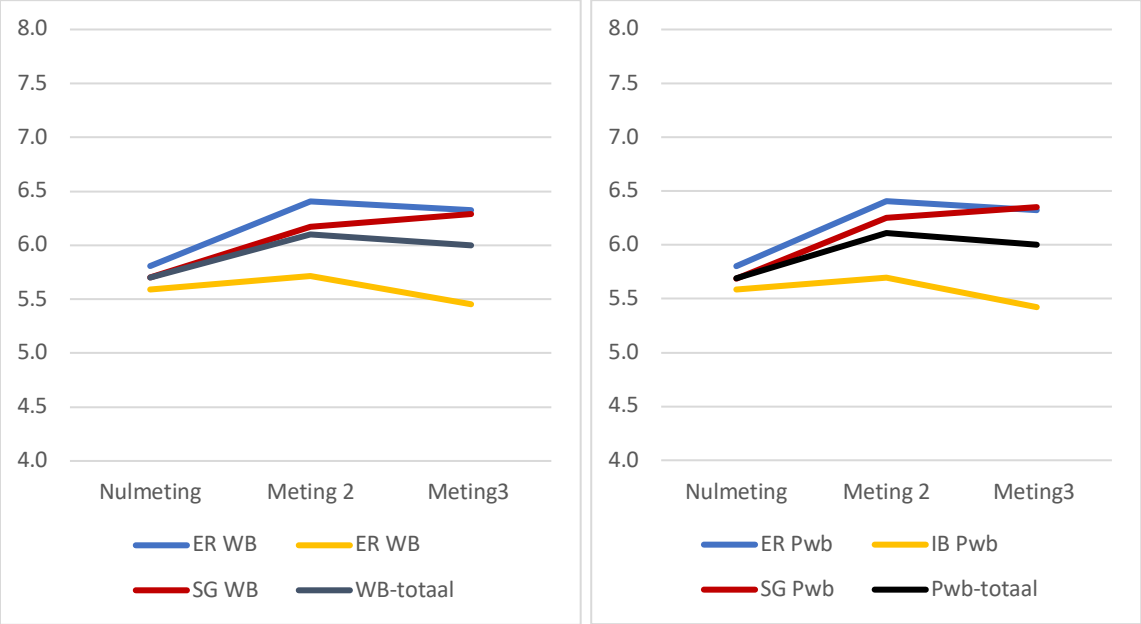


Fig. 3.1.5: Waargenomen en gecorrigeerd nivo van welbevinden (geschatte kans) in de drie metingen
Noot: WB=Welbevinden; Pwb=Geschatte kans op welbevinden; ER=Eigen Regie; SG=Standaard groep; IB=Intensieve begeleiding
Bron: Cliëntvragenlijsten, nulmeting, meting2 en 3

Figuur 3.1.5 doet vermoeden dat er een afname is met name bij de groep intensieve begeleiding ten opzichte van de andere twee treatments althans vanaf meting 2. Tussen nulmeting en meting 2 is nog sprake van een lichte stijging. De OLS regressieanalyse (zie tabel 3.1.8.) geeft hierover uitkomst. Ten eerste laat die analyse zien dat over de tijd er een significante of betekenisvolle toename is van welbevinden (zie de onderste rij met de rijtotalen). Kennelijk hebben alle treatments een positieve invloed gehad op het welbevinden van de deelnemer. Verschillen in welbevinden tussen deelnemers worden daarnaast voor een deel verklaard door de treatment maar we vinden geen betekenisvolle verschillen tussen de treatments voor de verandering in welbevinden over de tijd (zie de kolomtotalen over alle drie metingen tezamen). Met andere woorden, in alle treatments neemt het welbevinden toe over de tijd gezien, maar niet meer of minder in de ene of de andere treatment. Dus ook niet in vergelijking met de standaardgroep hetgeen wel een van de belangrijkste hypotheses was. Dat is te meer opvallend omdat op grond van de literatuurstudie voorafgaand aan het onderzoek betekenisvolle effecten verwacht werden. Heeft dit te maken met de uitvoering van de treatment of met de treatment zelf? Het welbevinden in de standaardgroep blijkt betekenisvol toe te nemen vanaf meting 2. De regressieanalyse toont voorts aan dat het welbevinden betekenisvol afneemt bij de intensieve begeleidingsgroep althans na meting 2 na aanvankelijk nog een lichte stijging door te maken. Vermoedelijk heeft dit te maken met het feit dat aan het einde van de tweejarige periode de moeilijkst te ondersteunen groep met de grootste afstand tot de

arbeidsmarkt over is gebleven. Als dit zo is dan is het dus een selectie effect en geen treatment effect. Dat het welbevinden betekenisvol toeneemt bij de standaardgroep is inderdaad opvallend en doet de vraag rijzen of de treatment hetzelfde is gebleven of is veranderd gedurende de experimentele periode. In het eerste geval zouden de welbevinden scores hetzelfde zijn gebleven als aan het begin van het experiment tenzij andere factoren zoals de economische situatie of de arbeidsmarktsituatie die in deze 2 jaar zijn verbeterd ook geleid hebben tot een stijging van het niveau van welbevinden. Helaas ontbreken cijfers over welbevinden bij de referentiegroep om dit te kunnen testen. Wel geven de cijfers over welbevinden en geluk in Nederland voor de onderkant van de inkomensverdeling geen reden om een stijging te verwachten. Het SCP constateert zelfs in een recente publicatie dat de subjectieve levenssatisfactie score van mensen met een score van 6 of lager in de afgelopen jaren (hetgeen bij bijstandsgerechtigden het geval is) is afgenomen in plaats van gestegen ondanks de gunstige economische situatie (zie Wennekers e.a. 2019.). Naar deze bevinding is verder onderzoek nodig.

Lage welbevinden scores bij bijstandsgerechtigden

Tot slot valt op dat het niveau van welbevinden relatief laag is onder de deelnemers met een score van 5,5 tot 6; de gemiddelde score in Nederland is net onder 8 voor tevredenheid en geluk. Ook de verschillen tussen de deelnemers in de scores zijn relatief klein. De onderzochte groep van bijstandsgerechtigden scoort ruim twee punten lager dan gemiddeld in Nederland (zie SCP) hetgeen als een groot verschil kan worden gekenschetst. Eerder is geconstateerd dat de respons op de vragenlijsten is afgenomen na de nulmeting van 85% naar 60%. Dit heeft vermoedelijk te maken met een geringere motivatie om de vragenlijsten nog in te vullen na uitstroom uit de bijstand. Dit kan het beeld enigszins hebben gekleurd.

Invloed consulent op welbevinden

We hebben zoals gezegd ook een variabele in het model opgenomen die de ondersteuning door de consulent indiceert. Voegen we deze variabele toe aan het model dan stijgt de verklaarde variantie met 45% van 9,5% naar 17%. Verschillen in de ondersteuning door de consulent hebben dus invloed op het welbevinden van de deelnemer en verklaren voor een deel ook verschillen tussen de deelnemers in de verschillende treatments.

Tabel 3.1.8: Gemiddelde scores welbevinden gecorrigeerd voor compositieverschillen tussen treatmentgroepen (0-10)[1;2], 2017-2019

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	5,8	88	6,4	80	6,3	64	6,2	232
Intensieve begeleiding	5,6	89	5,7	69	5,4**	66	5,6	224
Standaardgroep	5,7	74	6,3**	40	6,4***	48	6,0	162
Totaal	5,7	251	6,1***	189	6,0**	178	5,9***	618

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo R^2=0.17;* Bijdrage verklaarde variantie door consulent=45%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

De consulent doet er dus toe. De invloed van de treatments op welbevinden veranderen evenwel niet (behalve bij subjectieve gezondheid) indien we de consulentvariabele toevoegen aan het model. Dat betekent dat als we een treatment effect vinden dat dat een echt treatment effect is en geen effect van de persoon van de consulent.

Subjectieve gezondheid

Een tweede indicator die we gebruikt hebben is een gemiddelde van subjectieve gezondheid en mentale gezondheid. De indicator combineert de algemene vraag over subjectieve gezondheid (Wat vindt u van uw gezondheid?, met als antwoord categorieën: uitstekend, zeer goed, goed, matig, slecht) met enkele vragen over mentale gezondheid. In totaal zijn er vier vragen gesteld over de mentale gezondheid: de mate waarin men zich kalm en rustig voelt, zich somber voelt en neerslachtig, zich droevig voelt, zich erg nerveus voelt en zich gelukkig voelt. De positief gestelde items zijn omgecodeerd zodat alle items in dezelfde (negatieve) richting wijzen. Voor 'de mate van' is gebruik gemaakt van een Likert schaal (altijd, vaak, soms, bijna nooit, nooit). Hiervan is een 10-puntsschaal geconstrueerd van 0-10 voor 'slechte' mentale gezondheid die vervolgens weer is omgezet naar een positieve tien-puntsschaal van mentale gezondheid.

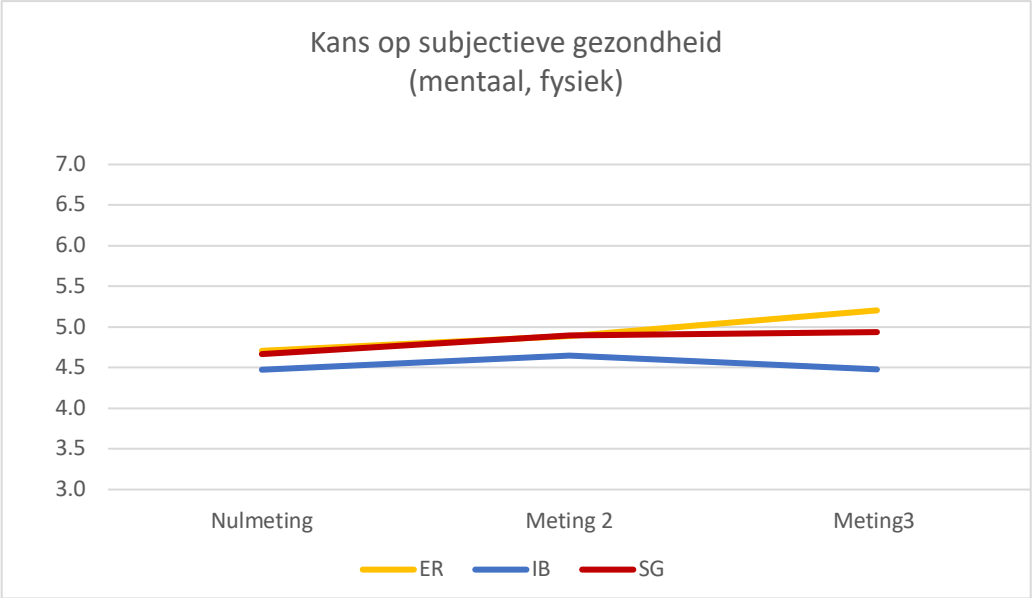


Fig. 3.1.6. Mate van fysieke en mentale gezondheid naar treatment en meting

Uit de figuur wordt duidelijk dat de gezondheidssituatie van de deelnemers te wensen overlaat. De gemiddelde score ligt rond de 5 dus krijgt een onvoldoende indien we het niveau als een rapportcijfer voor de gezondheid opvatten. De groepen gaan over de tijd iets meer uit elkaar lopen wat gezondheid betreft. De standaardgroep zit wat niveau van gezondheid betreft tussen de intensieve begeleidingsgroep (laagste score) en de eigen regie groep (hoogste score) in. Bij aanvang scoort opvallenderwijze de intensieve begeleidingsgroep iets maar wel significant slechter op gezondheid vergeleken met de beide andere groepen (zie Tabel A1.2 in Bijlage A). Over alle metingen gezien blijft dat effect significant negatief

(zie totaal kolom). Het kan duiden op selectiviteit in de afmeldingen: het aantal afmeldingen uit het zittende bestand was in deze groep relatief groot. Het kan ook te maken hebben met item non-response in de vragenlijsten. Het verschil is weliswaar significant maar wordt insignificant in een model zonder consulentnummer. De gezondheid blijft vervolgens gelijk bij de intensieve begeleidingsgroep maar stijgt ietsje bij de beide andere groepen. Deze stijging is niet betekenisvol. De subjectieve gezondheid is dus gemiddeld genomen over de tijd gelijk gebleven bij de treatment groepen in vergelijking met de standaardgroep. De veronderstelling was dat de beide treatments ten opzichte van de standaardgroep een positieve bijdrage aan ervaren gezondheid zouden kunnen leveren. In Tabel 3.1.9 presenteren we de resultaten. Op een na zijn alle verschillen in de tijd en tussen groepen wat de subjectieve gezondheid betreft niet betekenisvol. Deze hypothese wordt dus niet bevestigd. Kennelijk zijn om de gezondheid te bevorderen andere instrumenten nodig. Uit de literatuur bleek dat intensieve begeleiding en coaching middels "Individual Placement Support" en "Employment Support", dat zijn maatgerichte programma's voor mensen met fysieke of mentale gezondheidsproblemen en specifiek gericht op hun bijzondere problematiek, soelaas kan bieden om hen naar werk te begeleiden. Deze specifieke programma's zijn evenwel in dit experiment niet toegepast.

Tabel 3.1.9. Gemiddelde scores subjectieve gezondheid (mentaal, fysiek) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	4,7	88	4,9	80	5,2	64	4,9	232
Intensieve begeleiding	4,5***	89	4,6	69	4,5	66	4,5**	224
Standaardgroep	4,7	74	4,9	40	4,9	48	4,8	162
Totaal	4,6	251	4,8	189	4,9	178	4,7	618

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo* R^2*=0.162;* Bijdrage verklaarde variantie door consulent=41%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Keuzevrijheid

Voor veel mensen is keuzevrijheid een belangrijk goed. De Nobelprijswinnaar Stiglitz heeft in zijn boek "De prijs van ongelijkheid" (2010) de zere vinger gelegd op het probleem van de gebrekkige 'opportunities' die mensen aan de onderkant van de inkomensverdeling in Amerika ervaren en hebben; met andere woorden er zijn voor veel mensen weinig kansen en mogelijkheden om de eigen situatie te verbeteren en de dingen te doen die men voor het eigen leven belangrijk vindt. In zijn visie is dat slecht voor het welbevinden van de mensen zelf maar ook voor de economie. Sen (1985) benadrukt dat mensen de kansen of mogelijkheden (keuzeopties) moeten hebben of krijgen om de dingen te doen in het leven die men belangrijk vindt. Het gaat zowel bij Sen als bij Stiglitz niet om gelijkheid van uitkomsten maar om gelijkheid van kansen en mogelijkheden. Om die te bieden is een ongelijke behandeling soms nodig om die gelijke startposities te kunnen realiseren terwijl in beleid veelal het gelijkheidsbeginsel leidend is. Ook de toegang tot de arbeidsmarkt voor mensen met beperkingen belemmert. Het experiment biedt de

mogelijkheid om die kansen en mogelijkheden voor mensen aan de onderkant te bieden (maatwerk) of in ieder geval hen te leren die kansen zelf te scheppen en voor zover aanwezig te grijpen. Er is in het experiment meer vrijheid voor de consulent om aan de vraag van de deelnemer voor ondersteuning te voldoen of zelf met nieuwe oplossingen of instrumenten te komen niet allen in het domein van de P-wet voorzieningen maar ook in het bredere sociale domein, zo ook in Oss. Er zijn ook minder verplichtingen voor de deelnemers van beide treatments in het experiment of beter gezegd er wordt op een andere manier met verplichtingen omgegaan waardoor de betrokkenen meer keuzevrijheid ervaart en heeft.

In de vragenlijst is aan het aspect keuzevrijheid aandacht besteed doordat vragen zijn gesteld over de dingen die men belangrijk vindt in het leven en die men op de een of andere manier wel of niet kan doen omdat men daar de mogelijkheden wel of niet voor heeft (keuzemogelijkheden). Het zijn 7 vragen gebaseerd op Sen's capability benadering (1985) die beogen inzicht te krijgen in de keuzevrijheid ofwel in de mate waarin mensen de mogelijkheden hebben om de doelen die men stelt voor het eigen leven en die men belangrijk en waardevol vindt ook daadwerkelijk in het werk of anderszins te realiseren. De lijst zou een inzicht moeten geven in de keuzemogelijkheden en beperkingen die mensen ervaren in het leven in de bijstand. In de vragenlijst zijn zeven items voorgelegd aan de deelnemers. Ten eerste wordt gevraagd of ze deze items belangrijk vinden en ten tweede of ze de mogelijkheden hebben om dit te kunnen doen in hun eigen situatie. De zeven items zijn:

- Meebeslissen over belangrijke dingen in werk en leven
- Goede contacten met anderen hebben
- Eigen doelen stellen
- Een voldoende inkomen hebben
- Iets waardevols bij te dragen aan het leven van anderen
- Om te kunnen leren en nieuwe dingen te doen (kennis en vaardigheden ontwikkelen en gebruiken)
- Om iets waardevols bij te dragen aan het leven van anderen.

Gevraagd wordt op een vijfpunts Likert schaal naar de mate waarin men dit belangrijk vindt en in de eigen situatie ook kan doen: helemaal niet, niet, soms wel, soms niet, vaak of altijd.

Capability indicator

In overleg met de andere onderzoekers in LOEP is een soort capability indicator ontwikkel die aangeeft in welke mate mensen in staat zijn dingen die men belangrijk vindt te realiseren. Daarbij zijn twee varianten onderscheiden. De eerste variant berekent een score op basis van uitsluitend de items die men vaak of altijd belangrijk vindt. Items die men niet of minder belangrijk vindt worden niet meegenomen en krijgen een gewicht 0. De tweede variant weegt de score op de mate dat men deze zeven dingen kan doen met de mate waarin ze belangrijk worden geacht. Daardoor worden alle items meegenomen in de score, ook items die men minder belangrijk vindt. Deze laatste krijgen dan wel een lager gewicht. De antwoord categorieën (helemaal niet tot altijd) worden omgezet in een score van 0 tot 10. Vervolgens wordt een somscore bepaald van alle items gewogen met de mate van belangrijkheid (score 1=helemaal niet tot 5=altijd). De laatste score wordt gedeeld door 5 om gewichten te krijgen variërend van 0,2 tot 1. Deze tweede variant wordt hier gebruikt.

In Fig. 3.1.7 worden de resultaten weergegeven. In vergelijking met de standaardgroep liggen de geschatte scores voor beide treatments tot aan meting 3 iets hoger, met name bij de eigen regiegroep.

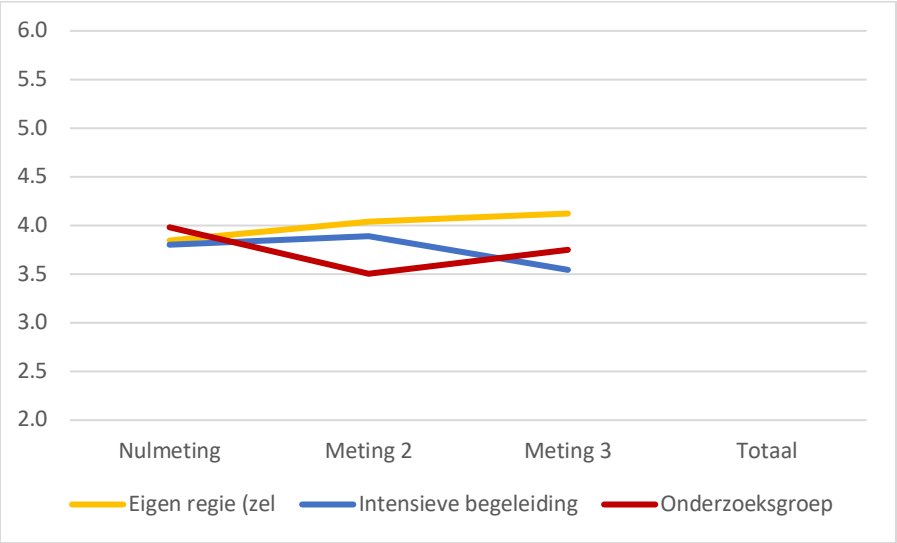


Fig. 3.1.7. Gecorrigeerde keuzemogelijkheden naar treatment en meting

Daarna nemen de ervaren keuzemogelijkheden de intensieve begeleidingsgroep af en bij de standaardgroep weer toe. Opvallend is opnieuw deze terugval na meting 2 bij de treatment intensieve begeleiding. Dit kan mogelijk verklaard worden doordat mensen met de kleinste afstand tot de arbeidsmarkt zijn uitgestroomd en de vragenlijst niet invullen. Of dat bij deze groep meer het geval is dan bij de anderen is getoetst. In tabel 3.1.10 worden de resultaten van de regressieanalyse weergegeven. De verschillen tussen de treatments blijken bij toetsing niet significant te zijn. Ook over de tijd zijn de verschillen tussen de groepen niet significant.

Tabel 3.1.10. Gemiddelde scores keuzemogelijkheden (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting		Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	3.8	88	4.0	80	4.1	64	3.8	232
Intensieve begeleiding	3.8	89	3.9	69	3.5	66	3.8	224
Standaardgroep	4.0		3.5	40	3.7	48	4.0	162
Totaal	3.9	251	3.9	189	3.8	178	3.9	618

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo* R^2*=0.097;* Bijdrage verklaarde variantie door consulent=34%.

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020

Het niveau van keuzemogelijkheden is gemiddeld bij alle treatmentgroepen laag te noemen met een gemiddeld niveau van 4 op een schaal van 0 tot 10. Qua keuzemogelijkheden scoren de deelnemende bijstandsgerechtigden dus een ruime onvoldoende. Nadere analyse laat zien dat vooral op het ontbreken van mogelijkheden om de kennis en vaardigheden te gebruiken en een voldoende inkomen te hebben bij deze bijstandsgroepen laag wordt gescoord. Om te toetsen of de veronderstelling juist is dat selectieve respons de daling na meting 2 bij de intensieve begeleidingsgroep verklaart hebben we een interactie tussen uitstroom naar werk en de treatments als verklarende variabele toegevoegd. Dan blijkt dat deze een positief en significant effect heeft op de ervaren keuzemogelijkheden maar alleen bij de eigen regiegroep. De veronderstelling dat deelnemers vanwege de treatments in vergelijking met de standaardgroep ook meer keuzemogelijkheden ervaren wordt dus niet bevestigd. Vooral bij de eigen regiegroep zien we een stijging ten opzichte van de standaardgroep maar ook deze is niet betekenisvol.

Sociale participatie

De bevordering van sociale participatie is voor veel gemeenten een belangrijk doel van het experiment. Zelfs indien de uitstroom naar betaald werk niet haalbaar is voor de deelnemer vanwege zijn of haar problemen en beperkingen, kunnen het experiment en de treatments er wel toe bijdragen dat de sociale participatie van de betrokkenen wordt gestimuleerd. Uit de literatuur over geluk en welbevinden is bekend dat sociale participatie van belang is om verschillen in geluk tussen mensen en veranderingen in geluk van mensen te kunnen verklaren (zie Wenneker e.a., 2019; Headey et al. 2010)

In de vragenlijsten zijn daartoe vragen gesteld over de mate waarin men zich opgenomen voelt in de samenleving dus de perceptie van sociale integratie in de samenleving. Daarnaast zijn vragen gesteld over de sociale participatie, dat is het aantal uren per week dat men besteed aan vrijwilligerswerk en mantelzorg, en er zijn vragen gesteld over het sociale netwerk door te vragen naar het aantal keer per maand dat men contact heeft met buurtgenoten, kennissen en vrienden en familie. Deze vragen zijn nu gebruikt om twee sociale participatie indicatoren te construeren.

- *Perceptie sociale integratie*: De eerste gaat uit van de beleving van de bijstandsgerechtigde deelnemer over de mate waarin hij of zij zich welkom en opgenomen voelt in de samenleving;
- *Sociale participatie:* De tweede baseert zich op de uren besteed aan vrijwilligerswerk en mantelzorg (maatstaf voor sociale participatie) en het aantal keren dat men per maand contact heeft met anderen (buren, kennissen-vrienden, familie ofwel het sociale netwerk). De scores voor de uren besteed aan vrijwilligerswerk/mantelzorg en sociale participatie/netwerk zijn ieder voor zich eerst omgecodeerd naar een schaal van 1 tot 10 en vervolgens is het gemiddelde over beide indicatoren berekend[13].

De gemeente acht de bevordering van de sociale participatie een belangrijk doel van het experiment. Voor een belangrijk deel van de cliënten is dit ook het hoogst bereikbare doel vanwege de grote afstand tot de

[13] De verdeling van uren besteed aan vrijwilligerswerk en mantelzorg was erg scheef verdeeld. Veel mensen scoorden 0 (geen vrijwilligerswerk) terwijl anderen ofwel minder dan 10 uur per week eraan besteedden ofwel veel meer. We hebben de verdeling van de scores omgezet naar een verdeling in 10 gelijke groepen en het gemiddelde voor elke groep vervolgens toegerekend. Op deze manier werd een schaal van 0 tot 10 verkregen. Voor het aantal contacten hebbe we somscores bepaald over de drie typen van contacten en vervolgens genormaliseerd ([x-min(x)/max(x)-min(x)]*10) van 0 tot 10.

arbeidsmarkt als gevolg van gezondheidsproblemen of een lage opleiding. Meer dan 70% heeft geen startkwalificatie en 60% heeft fysieke of mentale gezondheidsproblemen zoals uit de eerste tussenrapportage bleek (zie Muffels et al. 2018).

Kijken we nu naar de beleving van sociale integratie en hoe deze zich ontwikkeld heeft tijdens het experiment (Fig. 3.1.8 links) dan zien we tussen meting 1 en 2 dat deze licht is gestegen voor de eigen regiegroep en de standaardgroep en constant is gebleven voor de intensieve begeleidingsgroep. Na meting 2 neemt de ervaren sociale integratie echter sterk af bij deze groep. De verschillen over de tijd en tussen de afzonderlijke treatmentgroepen en de standaardgroep zijn evenwel niet significant. Daarom presenteren we hieronder alleen de regressieresultaten voor sociale participatie.

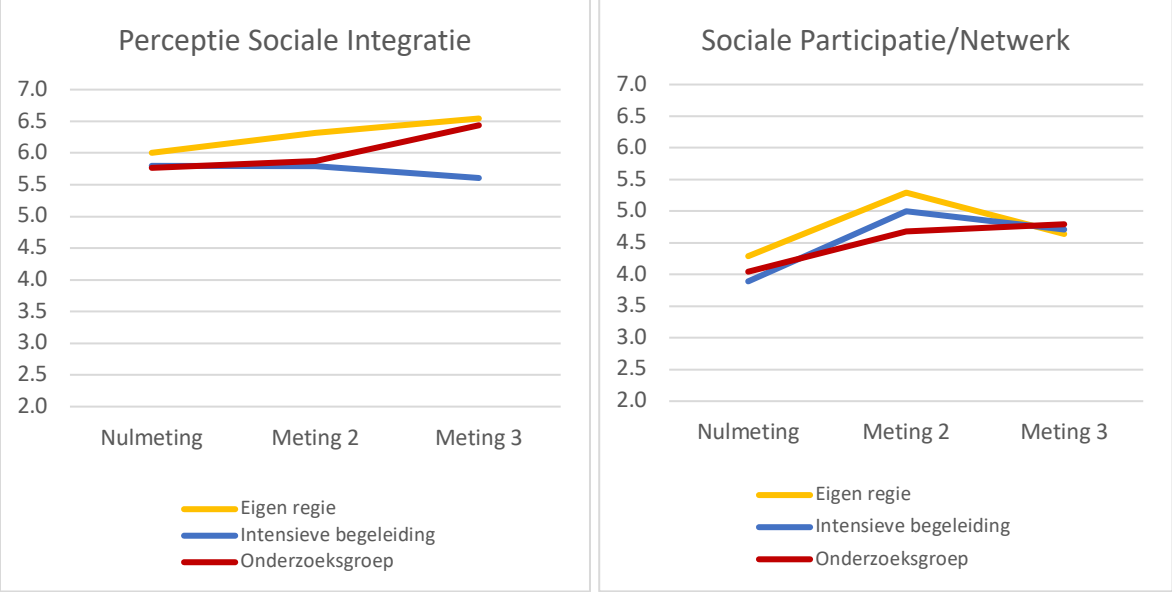


Fig. 3.1.8. Gecorrigeerde indicator perceptie sociale integratie (links) en participatie sociale netwerken (rechts) naar treatment en meting

In de figuur (rechts) over sociale participatie zien we evenwel een betekenisvolle stijging van sociale participatie bij alle groepen van meting 1 naar meting 3. Opvallend is dat ook bij de standaardgroep sprake is van een stijging; we vinden hier een treatment effect in meting 3. Na meting 1 valt het niveau van sociale participatie bij de beide treatment groepen wat terug (met name bij eigen regiegroep) maar blijft hoger dan het basisniveau terwijl deze terugval er niet is bij de standaardgroep. Indien de standaard treatment inderdaad hetzelfde is gebleven, hetgeen in de opzet van het experiment de intentie was, is er een andere verklaring nodig voor deze stijging. Het gevonden treatment effect duidt er echter op dat de standaard treatment is veranderd over tijd. In Tabel 3.1.11 worden de regressieresultaten voor sociale participatie weergegeven. Voor sociale participatie vinden we een positief treatment effect voor de standaardgroep in meting 3 (zie Tabel A1.1 in Bijlage A) en de intensieve begeleidingsgroep in meting 1 en over alle metingen heen (kolom totaal). Dit laatste is opvallend omdat het niveau van sociale participatie bij aanvang van het experiment significant lager was in de intensieve begeleidingsgroep (selectieeffect). Tot slot vinden we een positief treatment effect voor eigen regie in de kolom totaal.

Tabel 3.1.11. Gemiddelde scores sociale participatie (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	4,3	88	5,3	81	4,6	64	4,7*	232
Intensieve begeleiding	3,9**	89	5,0	69	4,7	66	4,5**	224
Standaardgroep	4,0	74	4,7	40	4,8**	48	4,4	162
Totaal	4,1	251	5,1***	190	4,7**	178	4,6***	618

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse (N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo R^2=0.17;* Bijdrage verklaarde variantie door consulent=12%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020

Zelfredzaamheid

De treatments zijn in belangrijke mate gestoeld op het idee dat het bevorderen van de zelfredzaamheid van de deelnemers uiteindelijk een belangrijk middel is om sociale participatie te verbeteren. De training die door Roland Blonk is aangeboden is gebaseerd op de theorie van gepland gedrag van Fishbein & Ajzèn (2010). In de training van de consulenten is door Blonk (2019) een zoekgedrag model gebruikt welke gebaseerd is op die theorie en die als basis diende voor de zogenaamde 'skills' methodiek waarin coaches/consulenten vaardigheden leren om de zelfredzaamheid van de deelnemers te bevorderen. Dat had niet alleen betrekking op de 'eigen regie' treatment maar ook op de 'intensieve begeleiding' treatment hoewel de wijze van coaching in deze laatste treatment anders, wellicht wat 'meer sturend' is als in de eerste. Maar ook in de eigen regie treatment wordt de deelnemer aangesproken op zijn of haar activiteiten en gedrag. In de vragenlijst is aan de mate van zelfredzaamheid daarom ook aandacht besteed. Er wordt onder andere gevraagd aan de deelnemer of men zichzelf zelfredzaam acht in verband met het zoeken en vinden van werk bijvoorbeeld. De lijst van capabilities eerder behandeld geeft ook inzicht in de belemmeringen en kansen om zelfredzaam te zijn, dat wil zeggen de dingen te kunnen doen die men belangrijk vindt voor het eigen leven. Wat zelfredzaamheid betreft op het gebied van werk zijn in de vragenlijst een viertal vragen gesteld daarover. Dat zijn:

- Ik kan betaald werk vinden als ik er echt moeite voor doe
- Ik denk dat ik in de toekomst betaald werk zal vinden
- Ik denk dat ik bij sollicitaties een goede indruk kan maken
- Ik denk dat ik een baan kan vinden die past bij mijn opleiding en ervaring.

De antwoordcategorieën vormen een 5-punts Likert schaal van 'geheel mee oneens' tot 'geheel mee eens'. Deze vier items zijn vervolgens omgezet in een schaal van 0 (geheel oneens) tot 10 (geheel eens) en er is een gemiddelde bepaald over de vier items. Dat noemen we de indicator zelfredzaamheid. Vervolgens is het standaard regressiemodel gedraaid om de voor verschillen in samenstelling gecorrigeerde kansen op zelfredzaamheid voor elk van de groepen te bepalen. In Figuur 3.1.9 worden de resultaten in beeld gebracht.

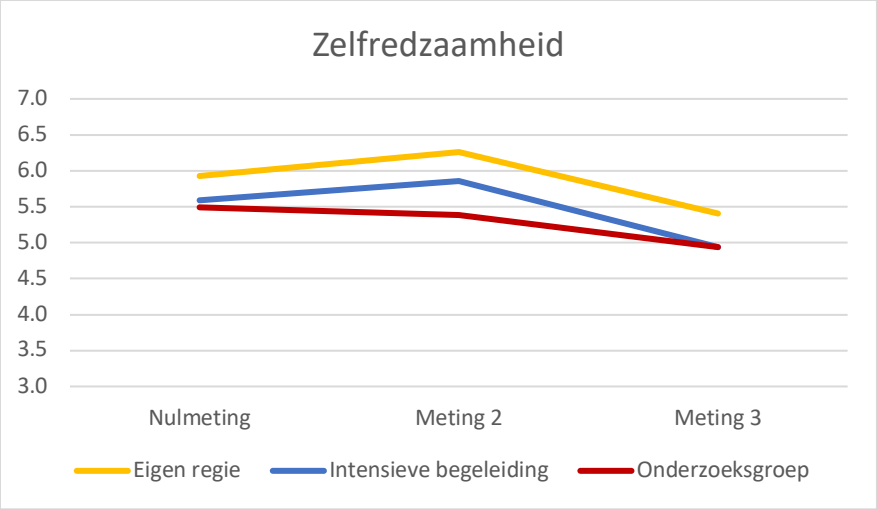


Fig. 3.1.9. Gecorrigeerde indicator zelfredzaamheid bij zoeken naar en vinden van werk naar treatment en meting

De zelfredzaamheid neemt in het eerste jaar over alle treatments tezamen significant toe om na meting 2 betekenisvol te dalen onder het basisniveau. Eerder vermoedden we dat selectie vanwege uitstroom naar werk mogelijk een rol heeft gespeeld.

Voor de beide treatment groepen ligt het niveau van zelfredzaamheid steeds hoger dan voor de standaardgroep waarvoor het niveau over de gehele periode van 2 jaar daalt. Het verschil tussen beide treatment groepen in vergelijking met de standaardgroep is evenwel niet significant zo blijkt uit de regressieanalyse (zie Tabel 3.1.12). Het is niet significant tussen de groepen maar wel significant en positief in de tijd in het eerste jaar evenals significant negatief in het tweede jaar. Zoals gezegd kan selectie in meting 3 de resultaten hebben beïnvloed aangezien op dat moment de groep met de kleinste kansen op werk is overgebleven en de mensen die zijn uitgestroomd de vragenlijst niet hebben ingevuld. Voegen we uitstroom naar werk toe aan de regressie dan is het effect positief en significant op 99% niveau. Causaal gezien leidt zelfredzaamheid tot uitstroom en niet andersom hetgeen betekent dat de meest zelfredzame deelnemers het eerst uitstromen. In interactie met de treatments is het effect uitsluitend significant bij de eigen regie groep waarin zelfredzaamheid het uitgangspunt van de treatment is. De meest zelfredzame mensen stromen dus het eerst uit maar in vergelijking met de standaardgroep vooral bij de eigen regiegroep. Er is dus geen zelfstandig effect van treatment op zelfredzaamheid maar alle treatments leiden wel tot toename van zelfredzaamheid althans in het eerste jaar. In jaar 2 is er wellicht door selectie een negatief treatment effect. Tabel 3.1.12 geeft het gemiddelde van de geschatte kansen per treatment en meting en de significantieniveaus.

Tabel 3.1.12 Gemiddelde scores zelfredzaamheid (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N=
Eigen regie	5.9	88	6.3	81	5.4	64	5.9	233
Intensieve begeleiding	5.6	89	5.9	69	4.9	66	5.5	224
Standaardgroep	5.5	74	5.4	40	4.9	48	5.3	162
Totaal	5.7	251	5.9*	190	5.1**	178	5.6***	619

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo* R^2*=0.28;* Bijdrage verklaarde variantie door consulent=12%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020

Vertrouwen

Het experiment is vertrouwensexperiment genoemd omdat het vertrekt vanuit het idee dat het opbouwen van een vertrouwensrelatie met de deelnemer uiteindelijk effectiever kan zijn om het gewenste doel te bereiken, namelijk bevordering van sociale participatie, dan de standaardaanpak met een strenge toepassing van de verplichtingen en bij niet nakoming sanctionering van de uitkering. De toetsing van rechtmatigheid blijft intact maar de toetsing aan de andere (informatie, sollicitatie, tegenprestatie) verplichtingen wordt op een meer relaxte manier vormgegeven in nauwe afstemming en overleg met de deelnemer. Het is zeker geen 'loslaten' maar leren de eigen verantwoordelijkheid te nemen en zelfredzaam te zijn met het aanpakken van de problemen die men heeft en ervaart. Uit de experimentele economische literatuur weten we bijvoorbeeld dat het geven van vertrouwen leidt tot wederkerigheid, tot extra inspanning en verbeterde motivatie en productiviteit in reactie op het gestelde vertrouwen (Frey & Jegen 2001; Bohnet et al., 2001). Uit de psychologische motivatietheorie weten we dat extrinsieke motivatie ten koste kan gaan van de intrinsieke motivatie en dat sturing op intrinsieke motivatie beter kan werken().Vandaar ook de inbouw van de training van de consulenten om hen vaardigheden te leren die hen helpen mensen te ondersteunen in zelfredzaamheid. Oss is als niet-AMvB gemeente niet gemachtigd ruimere vrijlatingsbepalingen voor extra verdiensten dan in de wet staat toe te staan. Dat is voorbehouden aan de artikel 83 gemeenten. Daardoor is er geen treatment die het aspect belonen in plaats van straffen kan toetsen op zijn werkbaarheid. Er vindt wel beloning plaats maar dan in de vorm van intensieve begeleiding en maatwerk, een luisterend oor en aandacht. Doordat de caseload van de consulent verlaagd is van 125 naar 75 tot 90, is er meer tijd voor de begeleiding van de deelnemer. Er is tijd om een 'relatie' en of persoonlijke band met de deelnemer op te bouwen omdat er meer tijd voor contact is en meer coaching op basis van vertrouwen. Het wantrouwen onder de deelnemers ten opzichte van de gemeente is groot zo hebben alle consulenten ervaren. Om daar verandering in te brengen kost tijd (zie verslag Focusgroep gesprekken hierna). Vertrouwen komt te voet en gaat te paard. Dus het vertrouwen is ook weer snel weg indien het vertrouwen wordt geschaad; dat geldt overigens voor beide partijen.

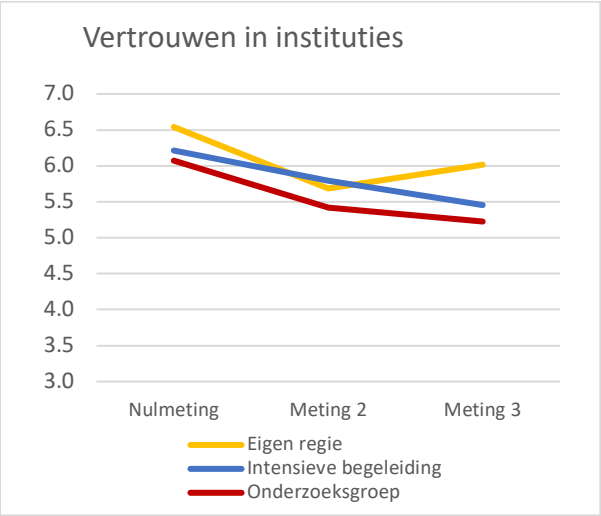
Over dit aspect van vertrouwen hebben we vragen gesteld in de vragenlijst. Ten eerste betreft dat een vraag over vertrouwen in de instituties. De vraag luidde in welke mate men vertrouwen had in:

- De landelijke overheid (regering, parlement)
- De gemeente
- De afdeling werk en inkomen van de gemeente
- De eigen consulent/werkcoach.

De antwoord categorieën vormde een vier-puntsschaal van 'helemaal geen vertrouwen', 'niet zo veel vertrouwen' 'een beetje vertrouwen' tot 'veel vertrouwen'. Deze is omgezet naar een 0-10 schaal. Daarnaast is gevraagd naar het vertrouwen dat men in andere mensen heeft. Dat wordt in de literatuur aangeduid als sociaal vertrouwen. Het wordt gezien als een vorm van sociaal kapitaal, mensen met veel vertrouwen hebben veelal een rijk sociaal netwerk dat hen ook helpt om vooruit te komen in het leven. Sociaal vertrouwen of vertrouwen in andere mensen is gemeten met een vraag die afkomstig is uit het "European Social Survey" of men vindt dat de meeste mensen te vertrouwen zijn of dat men niet voorzichtig genoeg kan zijn op een schaal van 0 (geen vertrouwen) tot 10 (volledig vertrouwen).

In Fig. 3.1.10 wordt de ontwikkeling in vertrouwen volgens de twee maatstaven per treatment groep weergegeven. Opvallend is dat het vertrouwen in de instituties significant is afgenomen maar het vertrouwen in anderen significant is toegenomen tussen de nulmeting en meting 2. Na het eerste jaar neemt het sociaal vertrouwen weer af maar het blijft wel op een hoger niveau. Zowel de afname (meting 2 en 3) bij institutioneel als de toename bij sociaal (vanaf meting 2) is significant. Het niveau van sociaal vertrouwen is lager dan het niveau van institutioneel vertrouwen zeker aan het begin.

We vinden bij de intensieve begeleidingsgroep een significant lager niveau van sociaal vertrouwen vergeleken met de standaardgroep welke reeds bij de start van het experiment bestaat en blijft bestaan daarna. We zagen eerder al dat er in deze groep enige selectivitiet is opgetreden.



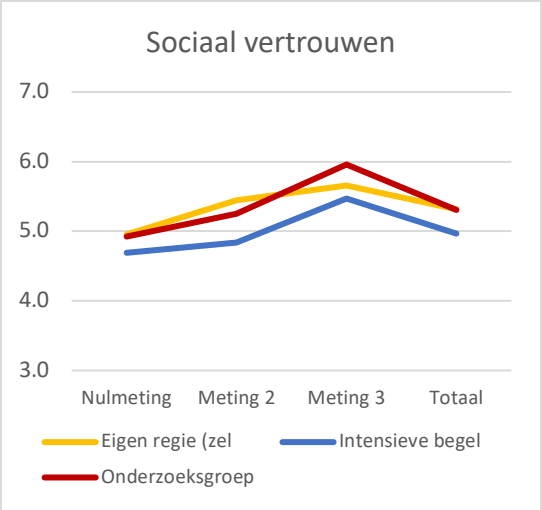


Fig. 3.1.10 Gecorrigeerde indicatoren vertrouwen in instituties en sociaal vertrouwen naar treatment en meting

In Tabel 3.1.13 worden de cijfers voor zowel institutioneel en sociaal vertrouwen in tabelvorm weergegeven.De significante effecten zijn vooral te vinden in de randtotalen welke verschillen over de tijd en tussen intensieve begeleiding en standaard treatment (negatief effect) laten zien.

Tabel 3.1.13 Gemiddelde scores institutioneel en sociaal vertrouwen (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Institutioneel vertrouwen				Sociaal vertrouwen				
	Nulmt	**Mt2**	**Mt 3**	**Tot.**	**Nulmt**	**Mt 2**	**Mt3**	**Totaal**	**N=**
Eigen regie	6.5	5.7	6.0	6.1	5.0	5.4	5.7	5.32	233
Intensieve begeleiding	6.2	5.8	5.5	5.9	4.7***	4.8	5.5	4.97***	224
Standaardgroep	6.1	5.4	5.2	5.7	4.9	5.2	6.0	5.31	162
Totaal	6.3	5.7***	5.6***	5.9	4.9	5.2	5.7***	5.19	619

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo R^2: Trust=0.11; Strust=0.20.* Bijdrage verklaarde variantie door consulent Itrust=20%; Strust=40%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020

We vinden voor zowel institutioneel als sociaal vertrouwen geen causale treatment effecten in de vergelijking met de standaardgroep behalve een selectie effect bij sociaal vertrouwen bij de groep intensieve begeleiding. De veranderingen bij de standaardgroep (daling institutioneel vertrouwen en stijging sociaal vertrouwen) wijzen niet op een treatment effect maar zijn wellicht het gevolg van andere ontwikkelingen aangezien de daling en stijging ook bij de andere groepen voorkomt.

Financiële situatie en armoede

In de niet-AMvB gemeenten is het niet wettelijk mogelijk om een vrijlatingsgroep in te richten waarbij mensen een extra deel van hun verdiensten mogen behouden. Dat is voorbehouden aan de gemeenten die zijn erkend als experimenteer gemeente. Verwacht mag worden dat voor deze groepen in deze gemeenten de financiële situatie wordt verbeterd en dat dit effecten heeft op de ervaren stress die een slechte financiële situatie en armoede met zich meebrengt. Nochtans kunnen er ook in Oss als niet-AMvB gemeente effecten zijn van een andere bejegening en aanpak bijvoorbeeld doordat meer mensen voltijds werk vinden en volledig onafhankelijk worden van een uitkering of dat schuldproblemen worden aangepakt of dat mensen gestimuleerd worden om hun huishoudboekje beter bij te houden.

Een van de veronderstellingen die we op grond van de psychologische literatuur naar voren hebben gebracht is dat door ervaren financiële stress mensen cognitief en motivationeel daar zo mee bezig zijn dat er mentaal geen ruimte is voor activiteiten die de eigen situatie op iets langere termijn zouden kunnen verbeteren zoals zoeken naar werk, het volgen van een opleiding, het inschakelen van het sociale netwerk en dergelijke.

Inkomensarmoede

In overleg met de LOEP onderzoekers zijn drie maatstaven gebruikt om de financiële situatie en de armoedesituatie in beeld te brengen. De eerste is een maatstaf welke veel is gebruikt in onderzoek naar de inkomenssituatie van huishoudens in Nederland. Daarbij wordt direct gevraagd naar de financiële situatie van het huishouden waarin men leeft. De vraag luidt als volgt:

Wat is op dit moment de financiële situatie van uw huishouden? De antwoordcategorieën zijn:

- Ik maak schulden
- Ik gebruik mijn spaarmiddelen om van te leven
- Ik kan precies rondkomen
- Ik houd een beetje geld over
- Ik houd genoeg over om te sparen
- Ik weet het niet
- Ik wil dit niet zeggen

In de vraag wordt tevens vermeld wat het begrip rondkomen inhoudt. Rondkomen is net genoeg geld hebben om de absoluut minimaal noodzakelijke kosten van levensonderhoud te kunnen betalen.

De eerste twee items zijn nu gebruikt om een persoon inkomensarm te noemen. Er zijn in de literatuur veel armoedemaatstaven ontwikkeld en ook verschillende inkomensarmoede maatstaven. De bijstandsgrens is zo'n grens die het SCP bijvoorbeeld hanteert naast een aantal andere. Een daarvan is het "niet-veel-maar-toereikend budget" dat ongeveer op het niveau van de bijstand van een alleenstaande ligt. We verwijzen naar de armoederapporten van het SCP die regelmatig verschijnen (Hoff e.a., 2019). Ook het CBS publiceert cijfers over armoede. Het SCP heeft berekend dat in 2014 ongeveer 45% van de mensen in de bijstand onder de "niet-veel-maar-toereikend" budgetgrens leven. Reente cijfers laten zien dat dit gedaald is naar 25-35% in 2019. De maatstaf voor armoede op basis van de vraag naar de financiële situatie noemen we inkomensarmoede en geeft het percentage deelnemers dat niet kan rondkomen met het huishoudinkomen en ofwel de spaarmiddelen aanspreekt om rond te kunnen komen ofwel schulden moet maken.

Deprivatiearmoede

De tweede maatstaf die we hebben gebruikt noemen we deprivatiearmoede en is gebaseerd op twee vragen waarin gevraagd wordt bij elk van de 14 items of men deze absoluut noodzakelijk vindt om te hebben of te kunnen doen en of men voldoende geld heeft om het te kunnen kopen of te kunnen doen. De lijst van 14 items is een selectie uit een lijst van 21 items die ook is bevraagd in onderzoek van de Europese Commissie, EU-SILC genaamd. We hebben er de 14 voor de Nederlandse samenleving relevante items uit geselecteerd. Dat zijn de volgende.

- 1 keer per dag groenten en fruit
- 1 maaltijd per dag met vlees kip of vis
- Twee paar schoenen
- Laptop of desktop computer
- Vrijetijdsactiviteiten buiten het gezin
- Vieren van bijzondere gebeurtenissen
- Jaarlijks 1 week op vakantie gaan in het buitenland

- Vervangen van beschadigde meubels
- Vervangen van huishoudelijke apparaten
- Kleding vervangen
- 1 keer per maand uit eten gaan
- Regelmatig deelnemen aan vrijetijdsactiviteiten
- Toegang tot internet
- Gebruik kunnen maken van openbaar vervoer

Vervolgens is het aantal items berekend dat men noodzakelijk vindt maar niet heeft. Voor de maatstaf van deprivatiearmoede zijn er vervolgens vijf essentiële items geselecteerd: 1x per dag groenten en fruit; 1 maaltijd per dag met vlees, kip of vis; vervangen beschadigde meubels; vervangen van huishoudelijke apparaten en versleten kleding vervangen. Indien iemand elk van deze items noodzakelijk vindt maar er minstens drie van deze vijf items mist beschouwen we deze persoon als arm. We noemen iemand dan deprivatiearm.

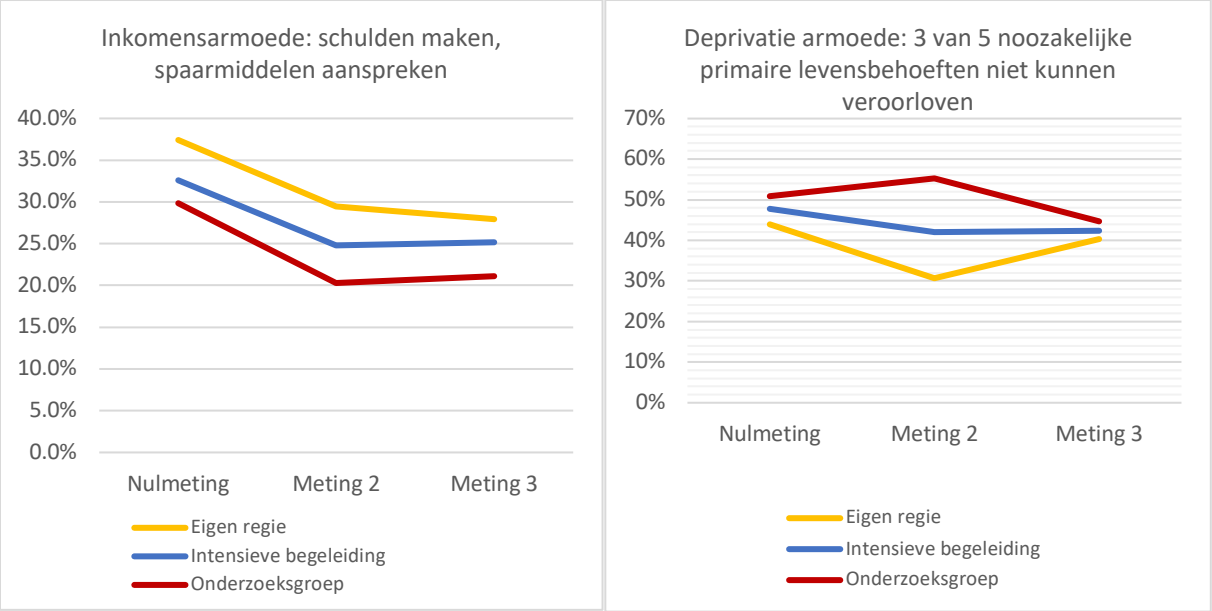


Fig. 3.1.11 Gecorrigeerde indicatoren voor financiële stress en deprivatiearmoede

Treatment effecten

We vinden geen significante effecten aan het begin, in de tijd of tussen groepen hoewel de scores tussen eigen regie en de standaardgroep in procentpunten voor zowel deprivatie als inkomensarmoede bij de start 8 tot 12% verschillen. Het ontbreken van een significant effect bij de start duidt erop dat geen zelfselectie heeft plaatsgevonden. Het slechte nieuws van deze cijfers is dat een flink deel van de bijstandsgerechtigden in inkomens -of deprivatiearmoede leeft. Naast 25 tot 35% inkomensarmoede leeft zo'n 40 tot 50% in deprivatiearmoede. Zeker ook dit laatste cijfer is verontrustend. De treatments in Oss hebben hier weinig invloed op kunnen uitoefenen mede omdat er geen inkomensvrijlatingsgroep kon worden ingericht. De daling kan het gevolg zijn van de gunstige economische situatie maar het kan ook zijn dat de hoge uitstroom naar werk (30%) in Oss hier debet aan is geweest. Echter, zoals we eerder

constateerden is deze uitstroomwinst niet beperkt tot de beide experimentgroepen maar realiseerde ook de standaardgroep gunstige uitstroomcijfers die meer dan 10% boven de uitstroomcijfers van de referentiegroep van niet-deelnemers uitstak. Voegen we uitstroom naar werk toe aan het model dan is er een significant negatief effect op deprivatiearmoede bij alle treatment groepen plus de standaardgroep.

Opvallend is nog wel dat de beide maatstaven iets heel anders lijken te meten. De standaardgroep heeft de laagste inkomensarmoede maar de hoogste deprivatiearmoede en omgekeerd de eigen regiegroep de laagste deprivatiearmoede maar de hoogste inkomensarmoede. De correlatie tussen beide is erg laag (0,07) hetgeen bevestigt dat ze iets anders meten. De inkomensarmoede meet vooral de perceptie van de eigen situatie en uit de armoedeliteratuur weten we dat mensen zich aanpassen en leren leven met het inkomen of de dingen die men heeft. De inkomensarmoede blijkt minder bij deelnemers die langer dan 3 jaar in de bijstand leven dan bij deelnemers die er nog maar kort in zitten terwijl de deprivatiearmoede juist toeneemt bij langere verblijfsduren.

Tabel 3.1.14 Gemiddelde scores inkomensarmoede en deprivatiearmoede (0-10) gecorrigeerd voor compositieverschillen tussen treatmentgroepen, 2017-2019

	Deprivatiearmoede				Inkomensarmoede				
	Nulmt	**Mt2**	**Mt 3**	**Tot.**	**Nulmt**	**Mt 2**	**Mt3**	**Totaal**	**N=**
Eigen regie	45.4%	31.4%	40.7%	39.3%	37.4%	29.5%	27.9%	33.5%	232
Intensieve begeleiding	47.8%	42.9%	43.0%	44.9%	32.6%	24.8%	25.1%	25.8%	224
Standaaard	51.6%	53.4%	43.5%	49.7%	29.8%	20.3%	21.1%	25.1%	162
Totaal	48.1%	40.2%	42.3%	44.0%	33.5%	25.8%	25.1%	28.7%	618

Noot[1]: ***p<0,01; **p<0,05; *p<0,10, getest op basis van OLS regressie analyse N=618); Model met treatment, cliëntvragenlijstmeting, consulentnummer en de controle variabelen.

Noot[2]: *Pseudo R^2: Ink armoede 0,10; Depriv.=0.07.* Bijdrage verklaarde variantie door consulent 26%;23%

Bron: Vragenlijsten nulmeting en vervolgmetingen productevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020

Samenvatting en conclusies: treatment effecten

In Tabel A1.1 in bijlage A presenteren we een overzicht van de gevonden treatment effecten ten opzichte van de standaardgroep en de effecten van de beide metingen voor alle drie groepen ten opzichte van de nulmeting op de niet-werk gerelateerde uitkomstmaten (ontwikkeling in de tijd). De meeste treatment effecten zijn niet significant. De meeste hypothesen worden daarom niet bevestigd met uitzondering van de hypotheses ten aanzien van sociale participatie. Daar vinden we positieve treatment effecten voor eigen regie en intensieve begeleiding vergeleken met de standaardgroep. De effecten zijn betrekkelijk groot zo blijkt uit tabel A1.1. Dat er geen treatmenteffect gevonden wordt bij welbevinden kan ook te maken hebben met verandering van de standaard treatment. Bij de uitkomstmaat van welbevinden vinden we immers een positief significant treatment effect over de tijd gezien bij de standaardgroep (zie Tabel A1.2 in Bijlage A). Indien de standaard treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Met andere woorden, indien de standaard treatment inderdaad is veranderd, is geen zuivere vergelijking met de treatment groepen meer mogelijk. Dan zullen

we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar.

Effecten over tijd

We vinden overigens wel positieve en significante effecten over tijd in verband met welbevinden maar dan voor alle groepen. Of dit het gevolg is van het experiment, dus een treatment effect, is vooralsnog causaal niet vast te stellen zonder zo'n vergelijkingsgroep. Het is wel opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel cijfers van het SCP suggereren dat voor mensen met een levenssatisfactiescore van 6 of lager (waaartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren is verslechterd. Het is dus aannemelijk om te veronderstellen dat we hier te maken hebben met een 'Hawthorne' effect bij de standaardgroep maar zeker weten doen we het niet en vraagt nader onderzoek. Dat de treatment van de standaardgroep is veranderd zou ook afgeleid kunnen worden uit het feit dat we positieve en significante effecten in de tijd vinden bij de standaardgroep in geval van sociale participatie (vrijwilligerswerk, sociaal netwerk). Voor sociale participatie blijven de treatment effecten overigens wel overeind.

Negatieve treatment effecten

We vinden ook twee negatieve effecten in geval van subjectieve gezondheid en sociaal vertrouwen; bij beiden betreft dit de intensieve begeleidingsgroep in vergelijking met de standaardgroep. We vermoeden dat enige zelfselectie heeft plaatsgevonden bij de afmeldingen (zittend bestand) of dat er item-nonrespons is in de vragenlijst maar dit vraagt nader onderzoek. Veel mensen in deze treatment hebben ook ten opzichte van de andere groepen bij de start van het experiment een lage score op subjectieve gezondheid en sociaal vertrouwen. De treatment heeft daar in de jaren daarna weinig aan kunnen veranderen.

Kortom...

Enkele van onze verwachtingen worden bevestigd anderen worden weerlegd op basis van deze cijfers. Met uitzondering van twee uitkomstmaten (subjectieve gezondheid en vertrouwen in de instituties) worden alle uitkomsten in de tijd gezien beter. We vinden geen causaal verband met de treatments met uitzondering van sociale participatie. Nader onderzoek moet uitwijzen of onderliggend deze stijging niet toch een treatment effect is dat niet zichtbaar wordt omdat de standaard treatment gedurende het experiment is gaan lijken op de andere treatments. We vinden ook verschillen tussen eigen regie en intensieve begeleiding. Veelal zijn de scores van eigen regie beter maar hier speelt mogelijk het selectieeffect bij intensieve begeleiding bij de aanmeldingen een rol.

Het gevonden positieve treatment effect op sociale participatie is zeker niet onbelangrijk aangezien dit een van de belangrijkste doelen was die de gemeente zich bij aanvang stelde. Eerder constateerden we al dat er positieve effecten zijn op uitstroom naar werk en bij deze uitkomstmaat konden we controleren voor het 'Hawthorne' effect. Daar vonden we sterke aanwijzingen voor zo'n effect, maar wel een die het onderzoek heeft bemoeilijkt bij de niet-werk gerelateerde uitkomstmaten.

Hierna bespreken we de uitkomsten van de procesevaluatie en analyseren we het uitvoeringsproces zelf aan de hand van de vragenlijsten gesteld aan de consulent.

3.2 Procesevaluatie: vragenlijsten voor consulenten

De procesevaluatievragenlijsten, zijn de vragenlijsten die door de betreffende consulent zijn ingevuld voor de deelnemers die deze consulent in begeleiding heeft. Een consulent heeft dus voor al zijn of haar deelnemers in beginsel drie keer de vragenlijst ingevuld tenzij de deelnemer vanwege verhuizing of afmelding is uitgestroomd. Deze vragenlijst is zoals eerder aangegeven in totaal drie keer ingevuld: de nulmeting, de eerste vervolgmeting (meting 2) en de tweede vervolgmeting (meting 3). De resultaten van deze vragenlijsten geven inzicht in het proces van begeleiding vanuit het perspectief van de consulent. Er wordt een beeld geschetst van de ontwikkeling die de deelnemer heeft doorgemaakt volgens de consulent en in hoeverre er sprake was van een effectieve begeleiding en of er zich leereffecten hebben voorgedaan voor zowel deelnemer als consulent.

3.2.1 Gehanteerde uitkomstmaten

Om een inzicht te krijgen in het proces van begeleiding en de ontwikkeling van de deelnemer door de ogen van de consulent, wordt er naar de volgende aspecten gekeken:

- Hoe beoordeelden de consulenten de **motivatie** van de deelnemers voor de treatment en de te behalen doelen en hoe veranderde dit over de tijd?
- Hoe beoordeelden de consulenten de deelnemer ten aanzien van een viertal aspecten van de begeleiding: of de deelnemer **realistische eisen** stelt ten aanzien van de (hoogst bereikbare) **doelen**; het **zelf in actie** komen, het **zelf doelen stellen** en het verwachte **effect van de treatment** gedurende het experiment.
- Hoe beoordeelden de consulenten de mate van **sociale participatie** en de **afstand tot de arbeidsmarkt** voor de deelnemer (de plaats op de participatieladder) en hoe veranderde dit over de tijd?
- Hoe beoordeelden de consulenten de **leereffecten** bij de deelnemer (mate van **zelfredzaamheid**) en bij zichzelf en hoe veranderde dit over de tijd?
- In hoeverre beoordeelden de consulenten de geboden treatment als geschikt voor de deelnemer (alleen in de derde vragenlijst bevraagd)?

Uiteraard is een belangrijk doel van het onderzoek om te bekijken in hoeverre er verschillen zijn tussen deelnemers die verschillende treatments hebben ondergaan. Voor de bovengenoemde aspecten zal dan ook onderscheid worden gemaakt tussen deelnemers van verschillende treatments. Hierdoor zal duidelijk worden of en welke verschillen er op een bepaald moment zijn tussen deelnemers en in hoeverre deze verschillen veranderen over de tijd en in hoeverre deze zijn toe te schrijven aan het effect van de treatment.

3.2.2 Analyse methode vragenlijst

Op basis van de verzamelde data van de consulent vragenlijsten is er een gewone regressieanalyse (OLS) uitgevoerd, waarbij het verband tussen de treatments en de gehanteerde uitkomstmaten wordt geschat. In deze analyse zijn een aantal onafhankelijke variabelen toegevoegd om de schatting van de uitkomstmaat (afhankelijke variabele) zo zuiver mogelijk te maken, namelijk: de meting waarin de score

is behaald op de uitkomstmaat, het nummer van de consulent die de vragenlijst heeft ingevuld en een dummy variabele die aangeeft of de deelnemer uit het zittende bestand komt, of uit de nieuwe instroom. In de rapportage van de analyse, wordt er gebruik gemaakt van een beschrijvende tabel waarin de gemiddelde scores voor de treatment groepen op een uitkomstvariabele worden getoond per meting en in totaal. Deze gemiddelden zijn geschat op basis van de eerder genoemde regressieanalyse. Een nadere uitleg over hoe de tabellen gelezen moeten worden, volgt in de volgende paragraaf.

3.2.3 Resultaten

In deze paragraaf zullen de resultaten op de vragen uit paragraaf 3.2.1 worden beschreven. De tabellen die worden getoond zijn als volgt te lezen. In de tabellen zijn de gemiddelde scores op de uitkomstvariabele (bijvoorbeeld motivatie, plaats op participatieladder, zelfredzaamheid, etc.) per treatmentgroep en per meting te zien. Deze gemiddelden zijn *geschat* op basis van een OLS regressie waarbij de treatment groepen en de metingen de onafhankelijke variabelen waren en de uitkomstvariabelen telkens de afhankelijke variabele. Daarnaast zijn er nog twee onafhankelijke variabelen gebruikt, namelijk: consulentnummer en een variabele die aangeeft of een deelnemer uit het zittend bestand of nieuwe instroom komt[14]. Er kan dus gezegd worden dat de gemiddelden die hier zijn weergegeven, de gemiddelden zijn onafhankelijk van het effect van de consulent en of iemand afkomstig is uit het zittend bestand of nieuwe instroom.[15] Hoe moeten de tabellen gelezen worden? De kolom 'gemiddelde totaal' geeft de gemiddelde score op de 3 metingen voor elke treatmentgroep. De sterretjes geven significantieniveaus weer (zie ook Bijlage B voor bijbehorende geschatte effecten). Bij drie sterretjes weten we met 99% betrouwbaarheid dat het verschil tussen de treatments (deels)[16] wordt veroorzaakt door de treatment zelf. Bij twee sterretjes weten we dit met 95% betrouwbaarheid en bij een ster met 90% betrouwbaarheid. Ontbreken de sterretjes, dan wordt het verschil waarschijnlijk door iets anders veroorzaakt dan door het effect van de treatment. De onderste rij met 'Totaal', geeft de gemiddelden over de groepen per meting, dus over de tijd. Sterretjes in deze rij laten zien dat er een significante of betekenisvolle toename of afname plaatsvindt over de tijd ten opzichte van de *nulmeting*. Kijken we naar de resultaten in de nulmeting dan kunnen we sterretjes interpreteren als betekenisvolle effecten van de treatments eigen regie en intensieve begeleiding vergeleken met de *standaard treatment* in die meting. Bij de resultaten voor metingen 2 en 3 geven de sterretjes de effecten (een betekenisvolle toe-of afname)

---

[14] Er zijn ook regressiemodellen getest waarin een meer uitgebreide lijst van onafhankelijke variabelen is opgenomen om de robuustheid van de resultaten te testen. In deze modellen werd er constant gehouden voor de volgende kenmerken van de deelnemer: leeftijd, geslacht, opleidingsniveau, soort woning, leefvorm, duur van uitkering. Uit deze analyses is gebleken dat de belangrijkste conclusies robuust zijn en dus overeind blijven na het toevoegen van deze variabelen. Om deze reden is er gekozen een beperktere lijst van onafhankelijke variabelen in het model mee te nemen, omdat dit grotendeels dezelfde uitkomsten opleverde.

[15] In totaal zijn er voor elke uitkomstvariabele twee regressiemodellen geschat: een zonder interactie tussen de treatments en meting en een met de interacties. De regressie coëfficiënten (effectmetingen) voor de effecten van de treatments en de metingen van het eerste model zijn opgenomen in Bijlage B, tabel B1.1. De regressie coëfficiënten van het tweede model met interacties zijn te vinden in Bijlage B, tabel B1.2. Voor het gemak van interpretatie laten we in deze paragraaf de geschatte gemiddelden zien, en niet de volledige uitkomsten van de regressieanalyse. De regressieanalyse (met alle eerder genoemde predictoren) is wel gebruikt als basis om de gemiddelden in tabel 3.2.1 t/m 3.2.10 te schatten.

[16] Waarom 'deels'?: als het effect van een treatment significant is, verklaart dit effect nog niet 100% van de verschillen in gemiddelden tussen groepen. Ook de variabelen consulentnummer en zittend bestand verklaren een deel en daarnaast zijn er altijd factoren die een invloed uitoefenen, maar niet worden meegenomen in het model.

weer van een treatment *bovenop* het effect in de *nulmeting* voor die treatment. Een sterretje bij de standaardgroep in bijvoorbeeld meting 2 betekent dan een betekenisvolle toe- of afname in de tijd vergeleken met de score in de nulmeting. Bij de beschrijving van de resultaten zal de nadruk liggen op betekenisvolle verschillen in scores die toe te schrijven zijn aan het effect van de treatments. In totaal zijn er 891 vragenlijsten ingevuld: 295 in meting 1, 296 in meting 2 en 300 in meting 3. Echter, in de tabellen die volgen kan het aantal vragenlijsten van dit totaal afwijken omdat er dan missende antwoorden zijn op bepaalde variabelen in de analyse. De totale aantallen staan in de tabellen vermeld.

Motivatie

In de vragenlijsten zijn de consulenten gevraagd om aan te geven in hoeverre zij de deelnemer gemotiveerd achten om zich in te zetten voor respectievelijk de eigen treatment en de (hoogst bereikbare) doelen. De resultaten worden getoond als een rapportcijfer van 0-10, waarbij 0 totaal niet gemotiveerd betekent en 10 heel gemotiveerd.

Tabel 3.2.1: *Gemiddelde scores motivatie eigen treatment (0-10)*

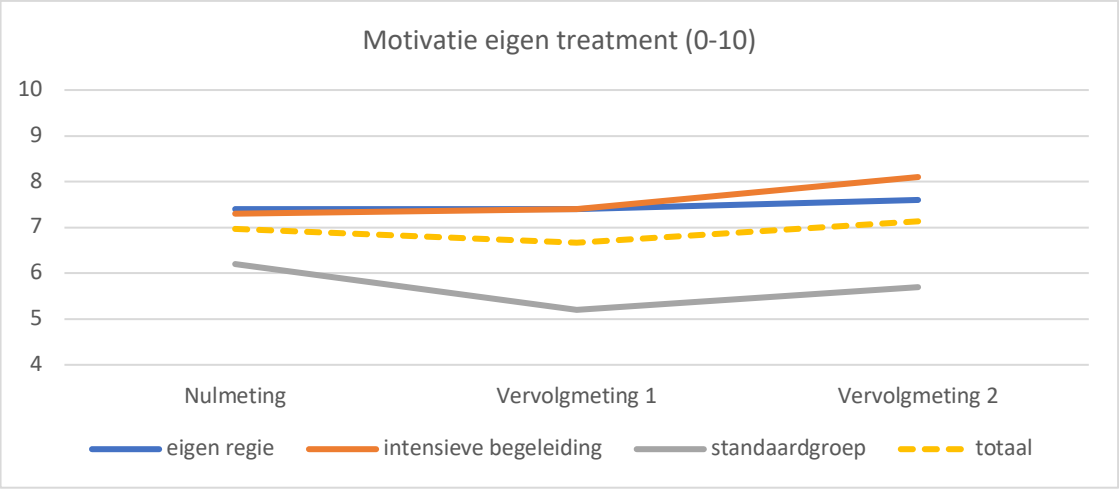
	Nulmeting	**N**	**Meting 2**	**N**	**Meting 3**	**N**	**Gemiddelde totaal**	**N totaal**
Eigen regie	7,4**	95	7,4*	90	7,6	81	7,5***	266
Intensieve begeleiding	7,3**	100	7,4*	88	8,1	88	7,6***	276
Standaardgroep	6,2	86	5,2**	74	5,7	77	5,7	237
Totaal	7,0	281	6,7	252	7,1	246	6,9	779

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=779)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

De resultaten in Tabel 3.2.1 tonen dat gemiddeld genomen de intensieve begeleidingsgroep het hoogst scoort op motivatie voor de eigen treatment (7,6), gevolgd door de eigen regie groep (7,5). De hogere score van beide groepen in vergelijking met de standaard treatment (5,7) is onder andere toe te schrijven aan het effect van beide treatments op de perceptie van motivatie van de consulent (zie ook tabel 1 in Bijlage A voor een positief effect van beide treatments). Uit figuur 3.2.1 blijkt ook dat de standaardgroep lager dan gemiddeld scoort, terwijl de andere twee groepen hoger scoren. Als we dan kijken naar de ontwikkeling over de tijd, dan zien we in de onderste rij 'Totaal' geen betekenisvolle toe- of afname over alle groepen heen. Maar als we dan kijken naar de vergelijking tussen de groepen en de ontwikkeling in de tijd, dan zien we dat in de nulmeting de treatments eigen regie en intensieve begeleiding significant hoger scoren op veronderstelde motivatie dan de standaardgroep (zie ook tabel B1.2 in Bijlage B). Tijdens de nulmeting had er nog nauwelijks begeleiding plaatsgevonden, waardoor we deze initiele verschillen in motivatie wellicht deels moeten toeschrijven aan selectie-effecten en niet aan daadwerkelijke verschillen door het effect van de treatment. In de eerste vervolgmeting is er voor de standaardgroep een betekenisvolle afname van motivatie ten opzichte van de nulmeting (zie ook tabel B1.2, Bijlage B), terwijl voor de eigen regie groep het positieve verschil met de standaardtreatment in de nulmeting nog groter wordt. In de 2^{e} vervolgmeting zien we een stijging in motivatie voor alle groepen ten opzichte van de

nulmeting, maar deze kan niet verklaard worden door de verschillen tussen treatments en het effect van tijd. Waarschijnlijk moet de verklaring gezocht worden in andere factoren. Samenvattend is het dus opvallend dat consulenten de motivatie in de nulmeting voor eigen regie en intensieve begeleiding hoger inschatten en dat het bij deze twee groepen ook iets toeneemt over de tijd, terwijl het bij de standaardgroep afneemt.



Figuur 3.2.1. *Gemiddelde score op motivatie voor eigen treatment per treatmentgroep en meting*

Daarnaast hebben de betrokken consulenten ook ingevuld wat hun inschatting is van de motivatie van de deelnemer om zich in te zetten voor de (hoogst bereikbare) doelen. In tabel 3.2.2 laten de resultaten zien dat gemiddeld genomen de treatments eigen regie en intensieve begeleiding significant hoger scoren dan de standaardgroep (zie ook tabel B1.1, Bijlage B). Daarnaast zien we over alle groepen totaal een toename in de veronderstelde gemiddelde motivatie over de tijd (van 5,3 naar 7,4). Als we dan specifiek kijken naar welke groepen er verschillend zijn van de standaardtreatment en of er een verschil is over tijd, dan zien we dat in de nulmeting de groepen eigen regie en intensieve begeleiding significant hoger scoren dan de standaardgroep (zie ook tabel B1.2, Bijlage B). Echter, kunnen deze verschillen in de nulmeting wederom deels worden toegeschreven aan selectie-effecten en niet zozeer aan verschillen in treatment, omdat de begeleiding nog nauwelijks had plaatsgevonden.

Tabel 3.2.2: Gemiddelde scores op motivatie hoogst bereikbare doel (0-10)

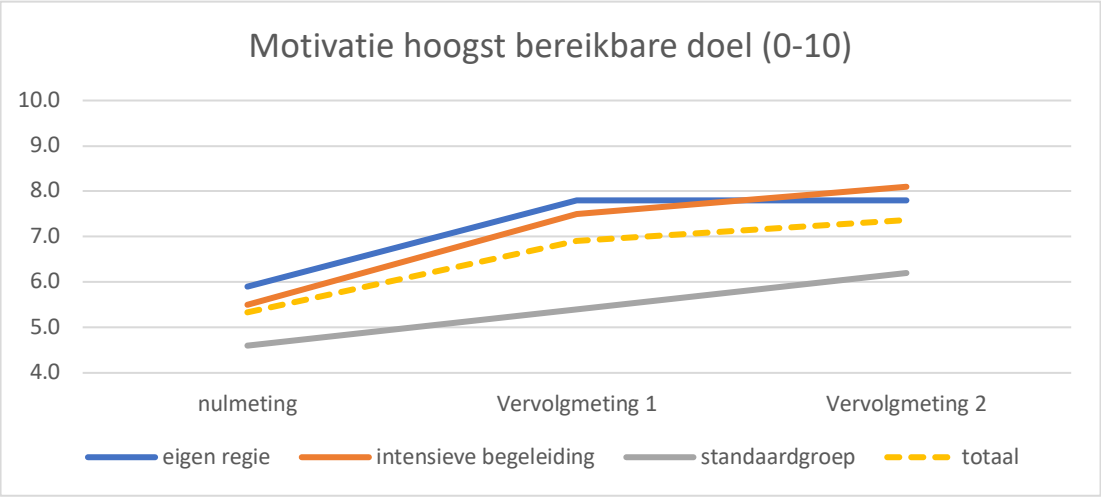
	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N Totaal
Eigen regie	5,9*	94	7,8	89	7,8	82	7,2**	265
Intensieve begeleiding	5,5**	97	7,5*	88	8,1	88	7,0**	273
Standaardgroep	4,6	85	5,4**	72	6,2***	77	5,4	234
Totaal	5,3	276	6,9***	249	7,4***	247	6,5	772

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=772)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

In de tweede meting zien we een betekenisvolle toename van motivatie voor de standaardgroep en intensieve begeleiding, maar geen significant effect van eigen regie. De treatments eigen regie en intensieve begeleiding stijgen ook in meting 3 flink in motivatie, maar deze stijging kan niet worden toegeschreven aan het effect van de treatments (zie tabel B1.2, Bijlage B). De standaardgroep neemt wel significant toe ten opzichte van de nulmeting en dit kan wel (deels) worden toegeschreven aan het effect van de treatment.

In de onderstaande figuur is deze ontwikkeling ook te zien: de standaardgroep scoort wat motivatie betreft duidelijk onder het gemiddelde, maar de motivatie neemt volgens de consulent wel toe. De intensieve begeleidingsgroep neemt relatief gezien het meeste toe over de tijd.



Figuur 3.2.2. *Gemiddelde score op motivatie voor het hoogst bereikbare doel per treatmentgroep en meting*

Realistische eisen

Het eerste aspect van de begeleiding waar we nu aandacht aan besteden is hoe realistisch de consulenten de eisen vinden die de deelnemer aan de te behalen doelen stelt. Ook hier kijken we weer naar een rapportcijfer van 0-10 waarin 0 staat voor zeer onrealistisch en 10 voor zeer realistisch.

Tabel 3.2.3: Gemiddelde scores op de mate van realistisch eisen stellen (0-10)

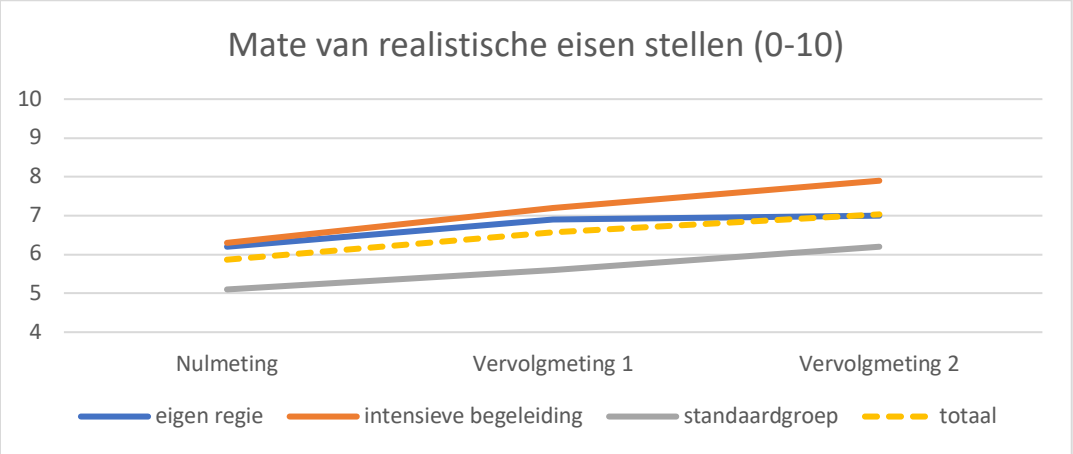
	Nulmeting	N	Meting 2	N	Meting 3	N	Totaal	N totaal
Eigen regie	6,2**	96	6,9	90	7	76	6,7**	262
Intensieve begeleiding	6,3	101	7,2	94	7,9	86	7,1	281
Standaardgroep	5,1	85	5,6*	85	6,2***	83	5,6	253
Totaal	5,9	282	6,6**	269	7,0***	245	6,5	796

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=796)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

In tabel 3.2.3 kunnen we zien dat gemiddeld genomen de eigen regie groep significant hoger scoort (6,7), dan de standaardgroep op de mate van realisme bij de deelnemer volgens de consulent. De intensieve begeleidingsgroep scoort ook hoger, maar dat verschil is niet te verklaren door het effect van de

treatment (zie ook tabel B1.1, Bijlage B). Verder zien we over alle groepen heen een betekenisvolle toename in meting 2 en 3 van 5,9 naar 7,0. Nader bekeken blijkt dat de treatment eigen regie hoger scoort dan de standaardgroep in de nulmeting en dat het verschil significant is (zie ook tabel B1.2, Bijlage B). Dat zou kunnen betekenen dat in de perceptie van de consulent de deelnemers aan de verschillende treatments reeds bij de start van het experiment een verschillend niveau van realisme laten zien. Dat zou kunnen wijzen op een selectie-effect waardoor bij aanvang van het experiment er reeds verschillen zijn die te maken hebben met zelf-selectie bij de aanmelding voor het experiment.



Figuur 3.2.3. *Gemiddelde scores op de mate van realistisch eisen stellen per treatmentgroep en meting*

Daarnaast zien we een betekenisvolle toename van verondersteld realisme voor de standaardgroep in meting 2 en 3 ten opzichte van de nulmeting. De consulent schat dus de eisen die deze groep stelt aan de te behalen doelen steeds realistischer in. Voor de andere twee treatment groepen, met name voor eigen regie zien we een toename over de tijd, maar deze kan niet worden toegeschreven aan het effect van de treatment (zie tabel B1.2, Bijlage B). Figuur 3.2.3 laat deze ontwikkeling zien.

Zelf in actie komen

Een ander aspect van de begeleiding is de mate waarin de consulent de deelnemer in staat acht om zelf in actie te komen om de gestelde doelen te behalen. In onderstaande tabel zien we de resultaten. Hier is weer een schaal van 0-10 gebruikt, waarbij 0 staat voor zeer slecht en 10 voor zeer goed.

We zien dat gemiddeld genomen de eigen regiegroep hoger scoort op het -volgens de consulent- zelf in actie komen, vergeleken met de standaardgroep en dat dit verschil is toe te schrijven aan het effect van de treatment (zie ook tabel B1.1, Bijlage B). Bij intensieve begeleiding is het verschil niet toe te schrijven aan het effect van de treatment en daarom niet significant. In de onderste rij zien we een betekenisvolle toename van het zelf in actie komen over tijd, van 4,8 naar 6,5 (zie ook tabel B1.1, Bijlage B). In de nulmeting kan het verschil tussen de eigen regie groep en de standaardgroep (deels) worden verklaard door het verschil in treatment, echter ook hier kan weer sprake zijn van een selectie-effect bij de start van het experiment.

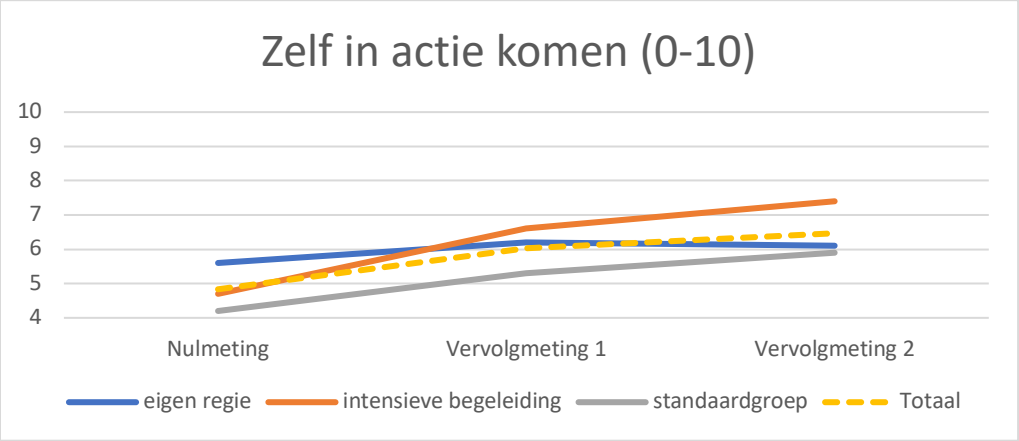
Tabel 3.2.4: Gemiddelde scores op zelf in actie komen (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	5,6**	88	6,2	91	6,1**	85	6,0**	264
Intensieve begeleiding	4,7	99	6,6	96	7,4	90	6,2	285
Standaardgroep	4,2	83	5,3**	87	5,9***	84	5,1	254
Totaal	4,8	270	6,0***	274	6,5***	259	5,8	803

Noot: ***p<0,01; **p<0,05;*p<0,10, getest op basis van een multiple regressie analyse (N=803)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Voor de standaardgroep is er een betekenisvolle toename vanwege de treatment over de tijd van 4,2 naar 5,9 in meting 3 (zie ook tabel B1.2, Bijlage B). Voor eigen regie zien we in meting 3 een significant effect van treatment, echter dit effect blijkt negatief (zie tabel B1.2, Bijlage B), terwijl het gemiddelde toch toeneemt. Als gevolg van de treatment neemt het zelf in actie komen dus iets af ten opzichte van de nulmeting, maar andere factoren zorgen ervoor dat er uiteindelijk toch een stijging plaatsvindt. In figuur 3.2.4 is te zien dat vooral de intensieve begeleidingsgroep volgens de consulenten een sterke groei doormaakt. De regressieanalyse maakt echter duidelijk dat deze toename niet aan het effect van de treatment kan worden toegeschreven en dus een andere verklaring heeft (zie tabel B1.2, Bijlage B).



Figuur 3.2.4. *Gemiddelde scores op zelf in actie komen per treatmentgroep en meting*

Zelf doelen stellen

Een derde aspect van de begeleiding is in hoeverre deelnemers zelf in staat zijn doelen te stellen, volgens de consulenten. Ook hier is weer een score van 0 tot 10 gebruikt, waarbij 0 betekent geheel niet in staat zelf doelen te stellen en 10 geheel wel.

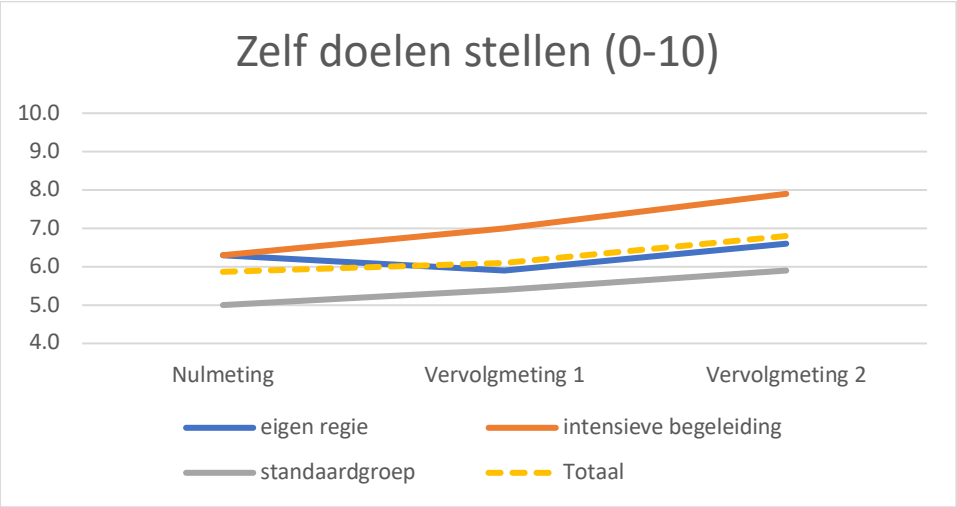
Tabel 3.2.5: Gemiddelde scores op zelf doelen stellen (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	6,3**	95	5,9	91	6,6*	83	6,3	269
Intensieve begeleiding	6,3	101	7,0	95	7,9	89	7,1	285
Standaardgroep	5,0	86	5,4	86	5,9**	91	5,4	263
Totaal	5,9	282	6,1	272	6,8***	263	6,3	817

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=817)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

In bovenstaande tabel zien we dat gemiddeld genomen de intensieve begeleidingsgroep het hoogst scoort, gevolgd door de eigen regie en de standaardgroep. Deze verschillen zijn echter niet (deels) toe te schrijven aan het effect van de treatments (zie tabel B1.1, Bijlage B). Daarnaast zien we een betekenisvolle toename in het zelf doelen kunnen stellen voor alle groepen tussen de nulmeting en meting 3 (zie tabel B1.1, Bijlage B). Nadere analyse toont echter dit alleen voor de standaardgroep een toename is (deels) vanwege het effect van treatment en dat voor de eigen regie groep dit eigenlijk een negatief effect van treatment is ten opzichte van het effect van eigen regie in de nulmeting (zie tabel B1.2 , Bijlage B). De stijging voor de eigen regie groep moet hier dus worden toegeschreven aan andere oorzaken. De sterke stijging voor de intensieve begeleidingsgroep kan ook niet worden toegeschreven aan het effect van de treatment. Wel zien we nog een betekenisvol verschil tussen de eigen regiegroep en de standaardgroep in de nulmeting. Dit kan zoals eerder gesteld wijzen op zelf-selectie bij de aanmelding. De eigen regie groep scoort duidelijk hoger, maar dit verschil wordt kleiner over de tijd, doordat de standaardgroep een betekenisvolle positieve ontwikkeling doormaakt. Dit is ook te zien in onderstaande figuur.



Figuur 3.2.5. *Gemiddelde scores op zelf doelen stellen per treatmentgroep en meting*

Verwacht effect van treatment op cliënt

Ook is de consulenten gevraagd hoe groot zij verwachten dat het effect van de treatment is voor de klant op een schaal van 0 tot 10. Hierbij staat 0 voor totaal geen effect en 10 voor een zeer groot effect.

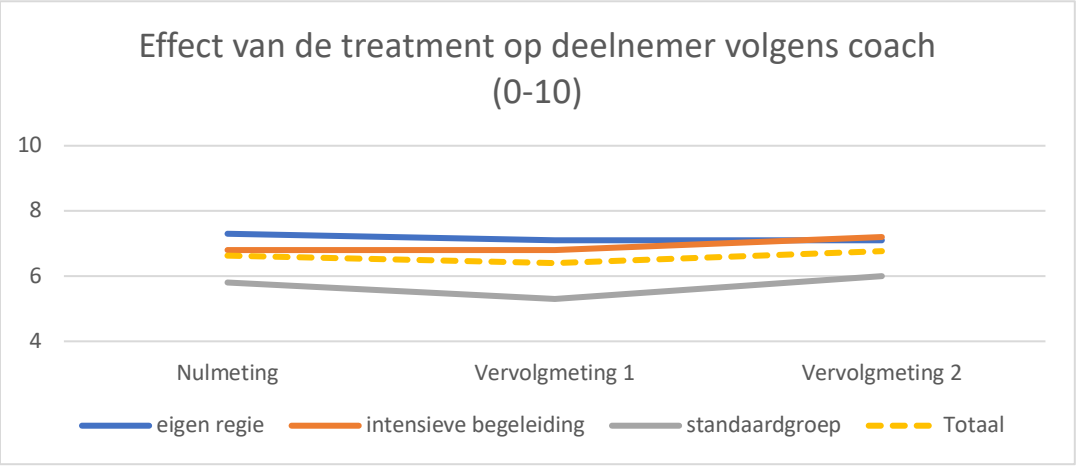
Tabel 3.2.6: Gemiddelde scores op het verwachtte effect van de treatment (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	7,3***	98	7,1	90	7,1*	81	7,2***	269
Intensieve begeleiding	6,8**	101	6,8	91	7,2	84	6,9**	276
Standaardgroep	5,8	86	5,3	88	6	88	5,7	262
Totaal	6,6	285	6,4	269	6,8	253	6,6	807

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=807)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Wat betreft de resultaten voor het verwachte effect van de treatment, valt op dat de consulenten voor de eigen regie en de intensieve begeleidingsgroep dit effect gemiddeld hoger inschatten dan die van de standaardgroep (zie tabel B1.1, Bijlage B). Deze verschillen zijn betekenisvol (deels) te verklaren door het effect van de treatments. Dit is vrij logisch, omdat dat natuurlijk een nieuwe aanpak is waar hoge verwachtingen voor zijn ten opzichte van de standaardgroep. Ook zien we dat dit verschil zich vooral voordoet aan de start van het experiment, en later niet of nauwelijks meer verandert. Over de tijd (de onderste rij) vindt er dus geen betekenisvolle ontwikkeling plaats in de verwachting van de consulenten. In meting 3 zien we een kleine afname in het gemiddelde voor de eigen regie groep (ook in tabel B1.2, Bijlage B), maar nog steeds is dit gemiddelde significant hoger dan dat van de standaardgroep. Figuur 3.2.6 laat ook zien dat er weinig verschil over tijd te zien is, maar dat wel de standaardgroep duidelijk onder het gemiddelde ligt.



Figuur 3.2.6. *Gemiddelde scores op het verwachtte effect van de treatment per treatmentgroep en meting*

Sociale participatie

Wat betreft sociale participatie, zijn consulenten gevraagd om een inschatting te geven van de mate hiervan voor de klant. Deze vraag is weer gescoord van 0 tot 10, waar 0 staat voor participeert totaal niet en 10 staat voor participeert totaal wel.

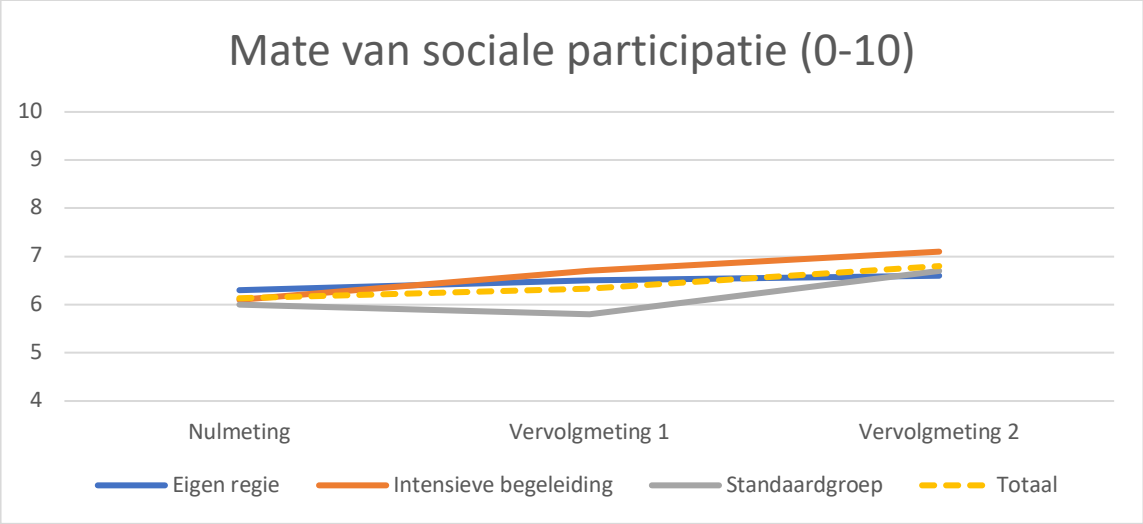
Tabel 3.2.7: Gemiddelde scores op sociale participatie (0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	6,3	98	6,5*	90	6,6	83	6,5***	271
Intensieve begeleiding	6,1	101	6,7*	87	7,1	86	6,6	274
Standaardgroep	6	86	5,8	89	6,7***	83	6,2	258
Totaal	6,1	285	6,3*	266	6,8***	252	6,4	803

Noot: ***p<0,01; **p<0,05;*p<0,10, getest op basis van een multiple regressie analyse (N=803)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Bovenstaande tabel laat zien dat de gemiddelden voor alle groepen niet erg verschillend zijn; allemaal net onder of boven de 6,4. Toch is de hogere score voor de eigen regie groep vergeleken met de standaardgroep (deels) betekenisvol en toe te schrijven aan het effect van die treatment. Verder zien we wel een duidelijke ontwikkeling over de tijd over alle groepen heen, namelijk er vindt een stijging plaats van 6,3 in de nulmeting naar 6,8 in meting 3 (zie ook tabel B1.1 , Bijlage B). In de nulmeting zien we al dat de eigen regie groep hoger scoort dan de standaardgroep en dat dat (deels) komt door het effect van die treatment. Een verdere toename van dat positieve effect ten opzichte van de standaardgroep zien we in meting 2 (zie tabel B1.2, Bijlage B). Daar zien we ook een betekenisvolle toename voor de intensieve begeleidingsgroep. De score van de standaardgroep neemt betekenisvol toe in meting 3, vergeleken met de nulmeting, wat (deels) kan worden verklaar door het effect van de treatment (zie tabel B1.2 , Bijlage B). Al met al is de mate van sociale participatie voor alle groepen matig en neemt gedurende het experiment maar bescheiden toe. Zo laat figuur 3.2.7 dit ook zien.



Figuur 3.2.7. *Gemiddelde scores op sociale participatie per treatmentgroep en meting*

Participatieladder

Naast de sociale participatie is de consulenten ook gevraagd een score te geven voor de afstand van de deelnemer tot de arbeidsmarkt op een schaal van 1 tot 6. Dit wordt ook wel de participatieladder of tredeladder genoemd. 1 staat hier voor een zeer grote afstand tot de arbeidsmarkt, terwijl 6 staat voor het verrichten van betaald werk.

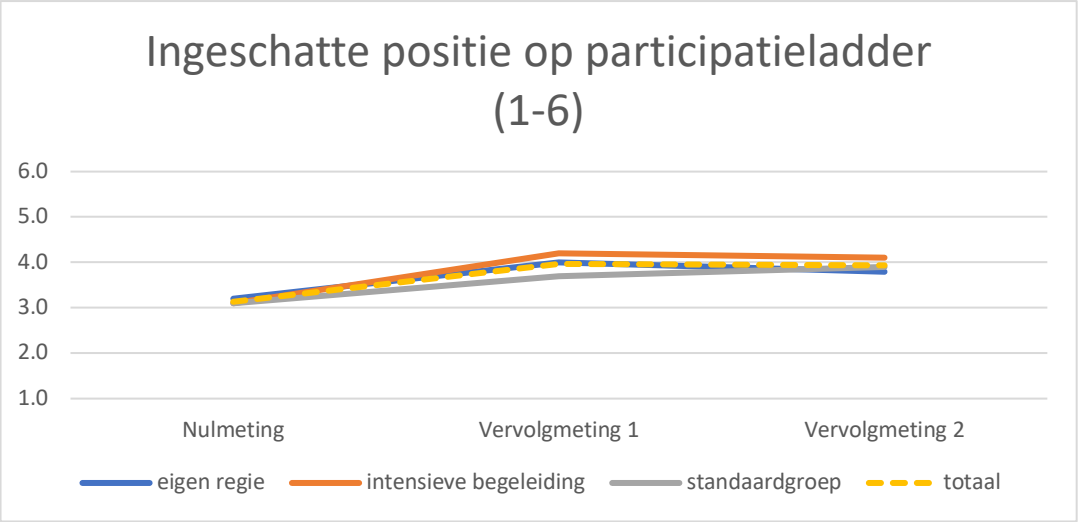
Tabel 3.2.8: Gemiddelde positie op participatieladder (1-6)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	3,2	94	4,0	92	3,8	89	3,7	275
Intensieve begeleiding	3,1	101	4,2	95	4,1	90	3,8	286
Standaardgroep	3,1	85	3,7**	93	3,9***	92	3,6	270
Totaal	3,1	280	4,0***	280	3,9***	271	3,7	831

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=831)
Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Het gemiddelde over alle groepen ligt op 3,7 en alle treatment groepen wijken hier maar weinig of niet vanaf. Wel is te zien dat over alle groepen bezien deze afstand kleiner wordt naarmate het experiment vordert: er is een betekenisvolle toename van 3,1 in de nulmeting naar 3,9 in meting 3 (zie tabel B1.1, Bijlage B). Gaan we echter dieper in op de verschillen tussen de groepen en over tijd, dan zien we dat deze toename over de tijd alleen voor de standaardgroep is toe te schrijven aan het effect van de treatment. Voor de andere twee groepen zijn er andere oorzaken voor de stijging en de kleine afname in meting 3. Een van de redenen van zowel de stijging over de tijd als de afname in meting 3 kan zijn dat naarmate het experiment vordert alleen degenen overblijven met een grote afstand tot de arbeidsmarkt. De deelnemers met een kleine afstand tot de arbeidsmarkt zijn inmiddels uitgestroomd en aan het werk. Dit beïnvloedt daardoor de perceptie van de consulenten over de arbeidsmarktperspectieven van degenen

die in begeleiding blijven en dus niet zijn uitgestroomd naar werk. In termen van afstand tot de arbeidsmarkt verschillen de groepen niet veel en scoren ze relatief laag op de schaal. Dat heeft voor een belangrijk deel te maken met de geringe kansen op werk voor met name de deelnemers met ernstige mentale of fysieke gezondheidsproblemen en vooral uit het zittende bestand. Alleen de standaardgroep laat althans volgens de consulenten zelf een positieve ontwikkeling zien gedurende het experiment die toe te schrijven is aan de treatment.



Figuur 3.2.8. *Gemiddelde positie op participatieladder per treatmentgroep en meting*

Leereffecten

Het begeleiden van de deelnemers met behulp van de treatments is een leerproces voor zowel de deelnemer als voor de consulent. Om dit leerproces in kaart te brengen gedurende het experiment, is aan de consulenten de vraag gesteld in hoeverre zij vinden dat de deelnemer leert om zelfredzaam te zijn en in hoeverre zij zelf leren van de treatment om zo hem of haar beter te begeleiden. Beide vragen hebben een score van 0 tot 10 gekregen, waarbij voor het leereffect van de deelnemer 0 staat voor totaal niet zelfredzaam en 10 voor volledig zelfredzaam en voor het leereffect van de consulent staat 0 voor niets bijgeleerd en 10 voor heel veel bijgeleerd.

Leereffecten deelnemer: zelfredzaamheid

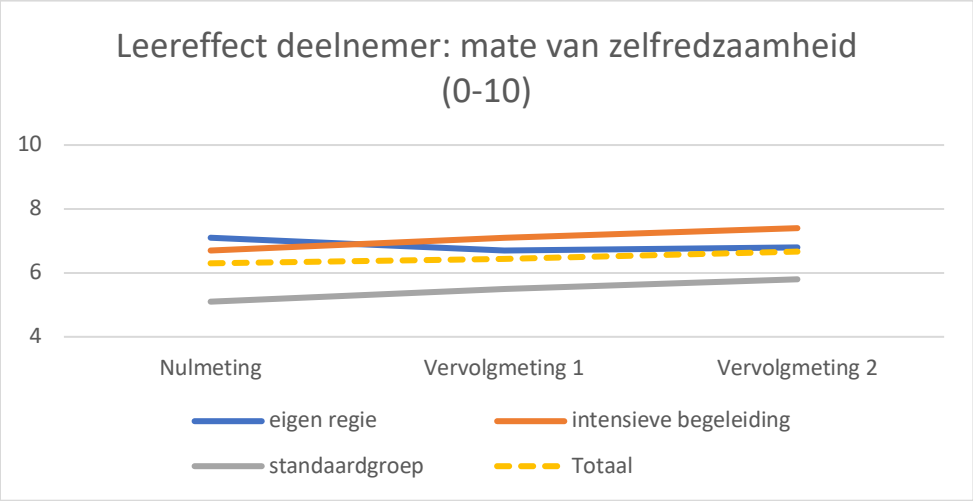
Tabel 3.2.9: Gemiddelde ingeschatte mate van zelfredzaamheid (leereffect deelnemer)(0-10)

	Nulmeting	N	Meting 2	N	Meting 3	N	Gemiddelde totaal	N totaal
Eigen regie	7,1***	98	6,7**	90	6,8**	81	6,9***	269
Intensieve begeleiding	6,7**	101	7,1	88	7,4	86	7,1**	275
Standaardgroep	5,1	86	5,5	79	5,8	84	5,5	249
Totaal	6,3	285	6,4	257	6,7	251	6,5	793

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=793)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

In tabel 3.2.9 is te zien dat gemiddeld genomen de treatments eigen regie en intensieve begeleiding hoger scoren op veronderstelde zelfredzaamheid dan de standaard treatment groep en dat dit verschil (deels) is toe te schrijven aan het effect van de treatments (zie tabel B1.1, Bijlage B). Uiteindelijk was dat ook een hoofddoel van deze treatment, om deelnemers zelfredzaamheid aan te leren. Opvallend is dat er geen betekenisvol verschil over de tijd is te zien over alle groepen heen. Nader bekeken blijkt wel dat in de nulmeting zowel de treatments eigen regie en intensieve begeleiding hoger scoren dan de standaard treatment hetgeen opnieuw kan wijzen op een zelf-selectie effect bij de aanmelding of een overschatting als gevolg van de aanvangsverwachtingen van de consulenten van deze treatments. Vervolgens zien we dat de eigen regie groep in de twee vervolgmetingen een betekenisvolle afname laat zien volgens de consulenten in de mate van zelfredzaamheid, wat wordt veroorzaakt door de treatment zelf (zie tabel B1.2, Bijlage B).



Figuur 3.2.9. *Gemiddelde ingeschatte mate van zelfredzaamheid (leereffect deelnemer) per treatmentgroep en meting*

Een mogelijke verklaring die hiervoor al is geopperd is, dat de meest zelfredzame deelnemers uitstromen naar werk, en er een groep overblijft die minder zelfredzaam is en meer tijd nodig heeft om dit te ontwikkelen. Dat punt kwam ook naar voren uit de focusgroep gesprekken die in het hoofdstuk hierna worden besproken. Al met al scoort de standaard treatment het laagst van alle groepen op zelfredzaamheid en heeft de intensieve begeleidingsgroep in de perceptie van de consulenten de hoogste score aan het einde van het experiment. Echter kan dit verschil tussen de nulmeting en meting 3 voor de treatment intensieve begeleiding niet geheel aan het effect van de treatment worden toegeschreven. De ontwikkelingen in de ingeschatte mate van zelfredzaamheid voor de drie groepen zijn ook duidelijk te zien in figuur 3.2.9.

Leereffecten consulent

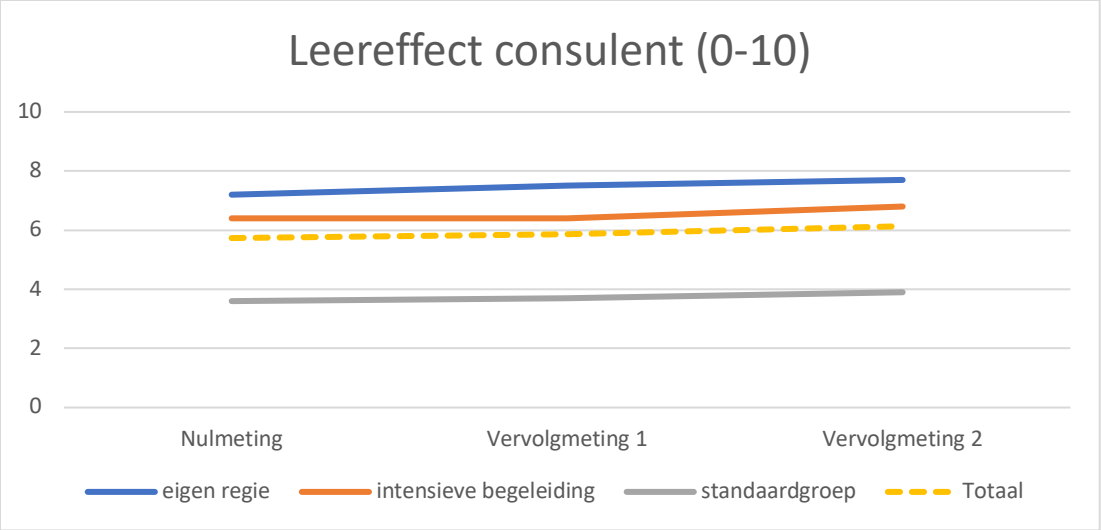
Vervolgens kijken we naar het leereffect voor de consulent. Wat meteen opvalt uit tabel 3.2.10 is dat gemiddeld genomen de treatments eigen regie en intensieve begeleiding significant hoger scoren dan de standaardtreatment en dit effect is toe te schrijven aan het effect van de treatments (zie tabel B1.1, Bijlage B). De standaardtreatment scoort een ruime onvoldoende, terwijl de andere twee treatments dik voldoende scoren. Vooral in de treatment eigen regie lijken de consulenten veel te leren. Over de tijd (in de rij 'Totaal') zien we echter geen betekenisvolle toe- of afname. Kijken we naar de tussenliggende cellen (zie ook tabel B1.2 , Bijlage B), dan vinden we dat de treatments eigen regie en intensieve begeleiding in de nulmeting significant hoger scoren op het veronderstelde eigen leereffect, dan de standaardtreatment. Dit duidt erop dat de consulenten betrokken bij de beide treatments aan het begin van het experiment meer verwachten van het experiment wat de eigen leereffecten betreft dan de consulenten in de standaardgroep hetgeen ook logisch is aangezien voor de laatsten er niet veel verandert. We zien ook dat dit effect van treatments niet toe- of afneemt in meting 2 en 3, want die effecten van treatment zijn niet significant (zie ook tabel B1.2, Bijlage B). In figuur 3.2.10 is duidelijk het verschil tussen de treatments te zien, met de standaard treatment die ver onder het gemiddelde scoort.

Tabel 3.2.10: ***Gemiddelde score op leereffect van consulent (0-10)***

	Nulmeting	**N**	**Meting 2**	**N**	**Meting 3**	**N**	**Gemiddelde totaal**	**N totaal**
Eigen regie	7,2***	98	7,5	98	7,7	97	7,5***	293
Intensieve begeleiding	6,4***	101	6,4	101	6,8	101	6,5***	303
Standaardgroep	3,6	86	3,7	95	3,9	102	3,7	283
Totaal	5,7	285	5,9	294	6,1	300	5,9	879

Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=879)

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.



Figuur 3.2.10. *Gemiddelde score op leereffect van consulent per treatmentgroep en meting*

Vergelijking eigen regie en intensieve begeleidingsgroep

In voorgaande tabellen is niet beschreven of de verschillen tussen de groepen eigen regie en intensieve begeleiding betekenisvol konden worden toegeschreven aan het effect van treatment, omdat we steeds een vergelijking hebben gemaakt tussen deze twee groepen en de standaardgroep. Dat was onze primaire vergelijking. Echter, om meer inzicht te bieden in het verschil tussen de eigen regie en de intensieve begeleidingsgroep, hebben we een kort overzicht gemaakt in onderstaande tabel. Daarin is steeds de groep intensieve begeleiding vergeleken met de eigen regie groep en of eventuele verschillen (deels) kunnen worden toegeschreven aan verschillen in treatment. Het gaat hier om de kolom 'gemiddelde totaal' in de tabellen 3.2.1 t/m 3.2.10. Er is gebleken dat er betekenisvolle verschillen zijn in de veronderstelde motivatie voor de eigen treatment en het hoogst bereikbare doel, waarbij intensieve begeleiding op het eerste concept iets hoger scoort en op de tweede iets lager, vanwege de treatment. De verschillen zijn zeer klein en slechts significant op 10% niveau (met 90% betrouwbaarheid). Daarnaast vinden we betekenisvolle verschillen in het effect van treatment bij de gepercipieerde positie op de participatieladder en het leereffect van de consulent. Bij de eerste scoort de intensieve begeleidingsgroep iets hoger en bij de tweede de eigen regie groep. Deze verschillen kunnen we met 95% betrouwbaarheid aannemen. Samenvattend kan worden gezegd dat de verschillen in effect van deze twee treatments op de onderzochte concepten minimaal zijn en er grotere verschillen bestaan in vergelijking met de standaardgroep.

Tabel 3.2.11 ***Geteste verschillen in score op concepten tussen intensieve begeleiding en eigen regie***

Concept	Verschil intensief vs. eigen regie
Motivatie eigen treatment	*
Motivatie hoogst bereikbare doel	*
Realistische eisen	-
Zelf in actie komen	-
Zelf doelen stellen	-
Verwacht effect treatment	-
Sociale participatie	-
Participatieladder	**
Leereffect deelnemer	-
Leereffect consulent	**

Noot: ***p<0,01; **p<0,05; *p<0,10

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Geschiktheid treatment

Ten slotte gaan we in deze rapportage in op een vraag die in de tweede vervolgmeting is gesteld aan de consulenten over de treatments. In deze meting is gevraagd in hoeverre de consulenten de treatment geschikt vonden voor deze klant, met een score 0 voor geheel ongeschikt en 10 voor totaal geschikt. In onderstaande tabel is te zien dat de eigen regie groep en de standaardgroep ongeveer hetzelfde scoren en er ook geen betekenisvol verschil is tussen deze groepen.

Tabel 3.2.12: ***Rapportcijfer voor geschiktheid treatment***

Treatment geschikt (0-10)?	Gemiddelde score	N
Eigen regie	6,8	88
Intensieve begeleiding	8**	87
Standaardgroep	6,7	96
Totaal	7,2	271

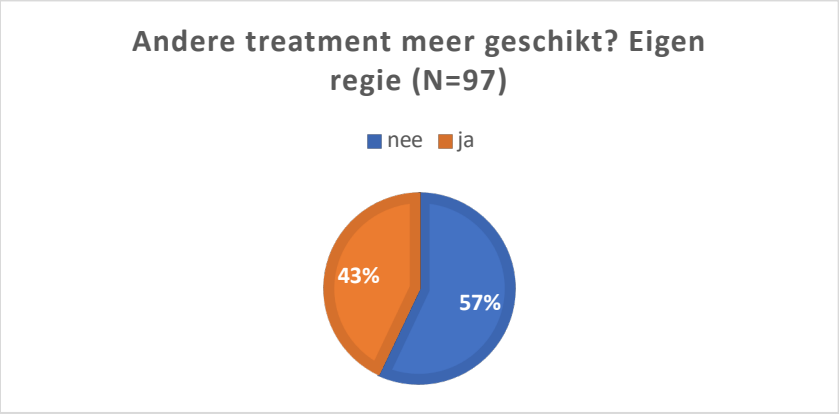
Noot: ***p<0,01; **p<0,05; *p<0,10, getest op basis van een multiple regressie analyse (N=271)

Bron: Vragenlijst tweede vervolgmeting procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

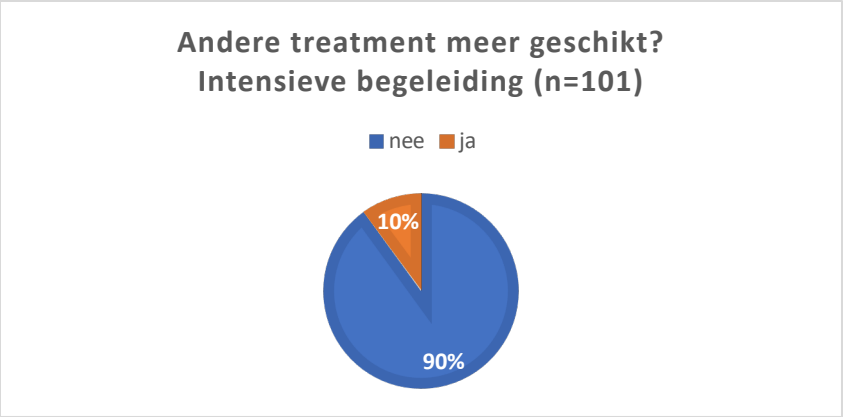
Beiden scoren ze een dikke voldoende op geschiktheid. Echter, de intensieve begeleidingsgroep is positiever met een gemiddelde score van een 8 voor geschiktheid. Ook is deze score significant hoger dan die van de standaardgroep. De consulenten van de intensieve begeleidingsgroep zijn dus positiever over de geschiktheid van de treatment ten opzichte van de andere twee groepen.

Vervolgens is na deze vraag de vraag gesteld of de consulenten vinden dat een andere treatment meer passend was geweest voor die deelnemer. Hieronder staan de resultaten samengevat in figuren. Meteen valt op dat bij de intensieve begeleiding het overgrote deel van de consulenten die treatment het meest

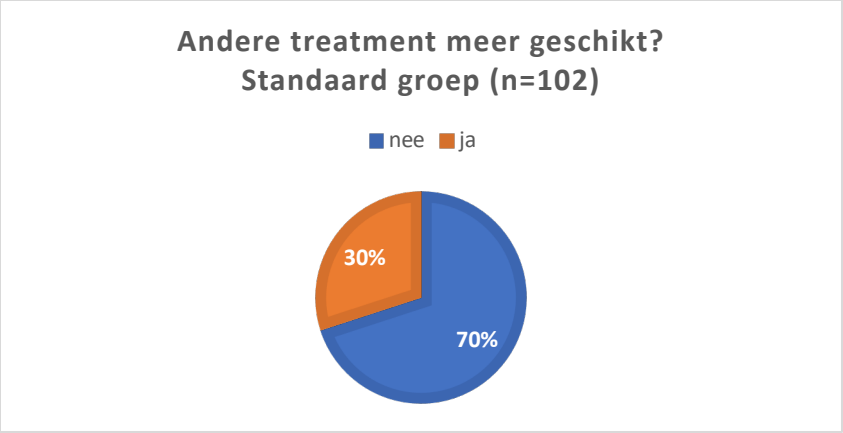
geschikt vonden en dat zij slechts in 10% van de gevallen een andere treatment zouden hebben gekozen. Bij de eigen regie groep is het meer verdeeld en zouden de consulenten in 43% van de gevallen een andere treatment hebben gekozen. Ten slotte laat de figuur voor de standaardgroep zien dat ze in 30% van de gevallen een andere treatment zouden hebben gekozen. Samenvattend lijkt het erop dat de consulenten van de eigen regiegroep het minst overtuigd waren van de geschiktheid van de treatment voor de deelnemer, terwijl de consulenten van de intensieve begeleidingsgroep het meest overtuigd waren dat de treatment de juiste was.



Figuur 3.2.11. *Percentage ja/nee voor treatment eigen regie*

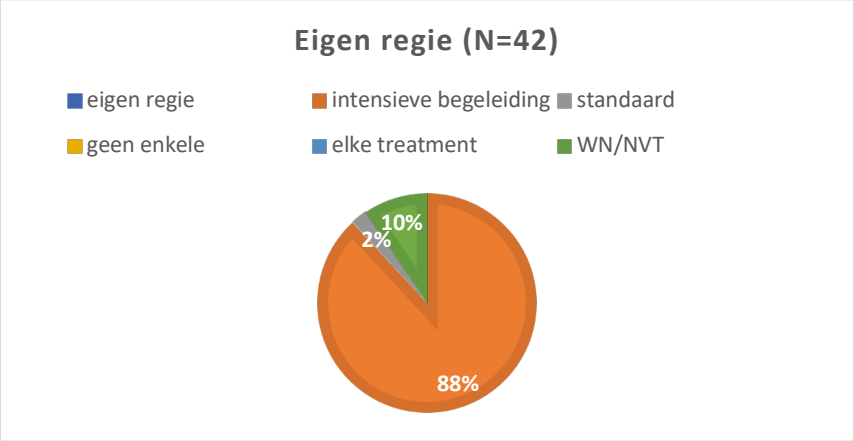


Figuur 3.2.12. *Percentage ja/nee voor treatment intensieve begeleiding*



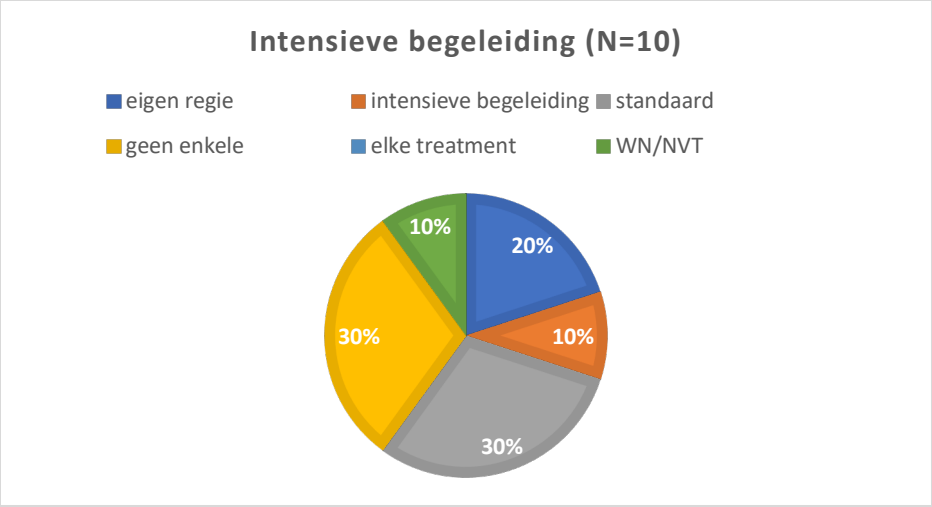
Figuur 3.2.13. *Percentage ja/nee voor treatment standaard groep*

Ter verdieping van de vorige vraag, is de consulenten ook gevraagd welke treatment ze dan zouden hebben gekozen als ze ja hadden geantwoord op de vorige vraag. Hieronder staan de resultaten samengevat. Voor 42 deelnemers binnen de groep eigen regie zou de consulent een andere treatment hebben gekozen, namelijk: in 88% van de gevallen voor intensieve begeleiding, in 10% van de gevallen weet de consulent het niet en voor 2% zou de standaardtreatment geschikt zijn geweest.



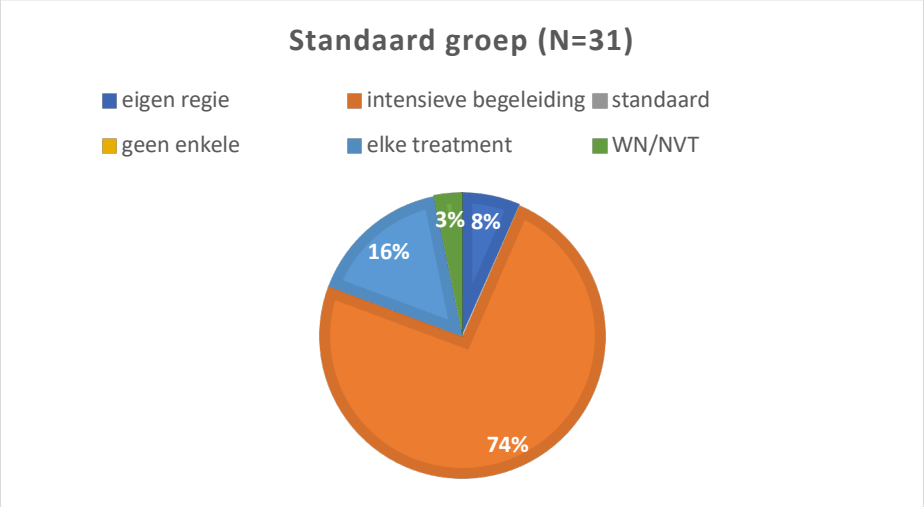
Figuur 3.2.14. *Indien andere treatment meer geschikt, welke is dan meer geschikt?*

Voor de intensieve begeleidingsgroep zou voor slechts 10 deelnemers een andere aanpak zijn gekozen, namelijk: in 30% van de gevallen voor de standaard treatment, voor 30% zou geen enkele treatment geschikt zijn, voor 20% had elke treatment geschikt geweest en voor de resterende 20% weet de consulent het niet of zou nog intensievere begeleiding moeten plaatsvinden. Opvallend is dat de antwoorden hier veel meer divers zijn dan bij de eigen regie groep. Wellicht is dit te verklaren doordat er veel tijd en energie is gestoken in de intensieve begeleidingsgroep en dat dat als het maximale van begeleiding wordt gezien door de consulenten, waardoor het onduidelijker is wat er moet gebeuren als dat zelfs niet werkt voor bepaalde deelnemers.



Figuur 3.2.15. *Indien andere treatment meer geschikt, welke is dan meer geschikt?*

Ten slotte zien we dat er voor 31 deelnemers in de standaardgroep een andere vorm van begeleiding gekozen zou zijn en ook hier zijn de antwoorden minder divers dan bij de intensieve begeleidingsgroep. Zo zien we dat voor 74% er zou zijn gekozen voor intensieve begeleiding, voor 16% had elke treatment geschikt geweest en voor 8% zou eigen regie passend zijn geweest. Voor 3% van de gevallen weet de consulent het niet.



Figuur 3.2.16. *Indien andere treatment meer geschikt, welke is dan meer geschikt?*

Samenvattend lijkt het erop dat de consulenten de treatment van intensieve begeleiding hoog waarderen. Deze krijgt het hoogste rapportcijfer voor geschiktheid en als er een andere treatment meer geschikt had geweest, dan is er bij de consulenten van de eigen regie groep en de standaardgroep een duidelijke voorkeur voor intensieve begeleiding. Echter, de intensieve begeleidingsgroep laat wel zien, dat wanneer die treatment niet geschikt is, het complex wordt geacht wat er in die gevallen moet gebeuren. Een derde van de consulenten weet het dan niet, of zou teruggrijpen naar de standaard manier van werken.

3.2.4 Conclusies

In deze paragraaf volgen een aantal conclusies uit de voorgaande analyses. Uit de resultaten bleek dat er voor een aantal uitkomstmaten verschillen bestonden tussen de treatmentgroepen, die ook te verklaren zijn aan de hand van het effect van de afzonderlijke treatments. Zo hebben we gezien dat de treatments eigen regie en intensieve begeleiding hoger scoren dan de standaardgroep op de volgende uitkomstmaten: *de motivatie voor de eigen treatment, de motivatie voor het inzetten voor het hoogst bereikbare doel, het verwachtte effect van de treatment, het leereffect van de deelnemer (zelfredzaamheid) en het leereffect van de consulent*. Daarnaast bleek dat de eigen regie groep ook hoger scoort dan de standaardgroep op *het stellen van realistische eisen, het zelf in actie komen en sociale participatie.* Hier lijken we uit op te kunnen maken dat de deelnemers van de eigen regie en intensieve begeleidingsgroep in de ogen van de consulent meer gemotiveerd zijn voor de treatment en om het doel te behalen en dat het verwachtte effect groter is, zowel om het doel te bereiken als de leereffecten voor beide partijen. Verder achten de consulenten de deelnemers van de eigen regie groep realistischer te zijn in het stellen van eisen aan het te bereiken doel, beter zelf in actie te komen en meer sociaal te participeren. Samenvattend blijkt in de perceptie van de consulent vooral de eigen regie groep voordelen te hebben ten opzichte van de standaardbegeleiding.

Bij een nadere analyse waar we hebben gekeken naar de gemiddelde score op de uitkomstmaten voor de verschillende treatments per meting, kwam naar voren dat de grootste verschillen tussen de treatmentgroepen bestaan tijdens de nulmeting. Omdat er ten tijde van de nulmeting nog maar een kort traject van begeleiding heeft plaatsgevonden (1 tot 4 maanden) en het verwachtte treatment effect laag is, zullen deze verschillen waarschijnlijk deels toe te schrijven zijn aan bijvoorbeeld zelf-selectie van deelnemers aan het begin van het experiment (de meest gemotiveerde deelnemers starten met het experiment of komen naar een afspraak met de consulent), of bijvoorbeeld omdat de consulenten zelf hogere verwachtingen hebben van de treatments eigen regie en intensieve begeleiding. Dat enthousiasme heeft mogelijk een positief effect op de inschatting door de consulent van de motivatie en het gedrag van de deelnemer.

De veranderingen die we over tijd zien zijn als volgt: volgens de consulenten neemt de motivatie voor de eigen treatment in de standaardgroep significant af, maar neemt juist de motivatie voor het hoogst bereikbare doel toe. Dit laatste geldt ook voor de intensieve begeleidingsgroep, als gevolg van de treatment. Bij de treatments eigen regie en intensieve begeleiding steeg de motivatie voor de eigen treatment wel.

Verder steeg de standaardgroep in de -door de consulenten veronderstelde- mate van het stellen van realistische eisen als gevolg van deze treatment en kwamen zij meer zelf in actie gedurende het experiment. Bij de eigen regie groep zagen we juist in de nulmeting een positief effect van deze treatment op het zelf in actie komen, maar in de derde meting werd dit gecompenseerd door een negatief effect. Een dergelijke afname van het positieve effect van de treatment eigen regie zagen we ook bij het zelf doelen stellen, het verwachtte effect van de treatment en het leereffect van de deelnemer. Deze resultaten kunnen te maken hebben met het feit dat uiteindelijk de meest kansrijke en zelfsturende groep eerder in het experiment uitstroomde en de groep die overbleef minder goed in staat bleek tot het voeren van eigen regie. Ook dat laatste blijkt uit het resultaat dat consulenten voor deelnemers waarvoor de

treatment eigen regie niet goed blijkt te werken, de voorkeur hebben voor een intensievere vorm van begeleiding.

Verder werd de standaardgroep beter in het zelf stellen van doelen en nam de sociale participatie toe. Dit laatste zagen we ook bij de eigen regie en intensieve begeleidingsgroep. Alle groepen groeiden dus in de door de consulent veronderstelde mate van sociale participatie, wat een belangrijk doel was van het experiment. Verder verbeterde volgens de consulent de standaardgroep de positie op de participatieladder gedurende het experiment. Dit was niet het geval voor de andere twee treatments.

Hoewel de gepercipieerde uitkomsten in de standaard begeleidingsgroep dus op veel vlakken verbeterden, bleef de gemiddelde score voor deze groep op veel uitkomstmaten onder het gemiddelde. Opvallend is wel dat de andere twee treatmentgroepen volgens de consulenten minder veranderingen over de tijd laten zien die toe te schrijven zijn aan de treatment, dan in de standaardtreatment, terwijl deze geacht zou worden om niet te veranderen gedurende het experiment. Kennelijk heeft er toch een ontwikkeling over de tijd plaatsgevonden in de aanpak van de begeleiding van de standaardgroep. Hiervoor wordt ook bewijs gevonden in het verslag van de focusgroepgesprekken onder stap 3.

Wat betreft de beoordeling van de treatments op de mate van geschiktheid voor de deelnemer, lijken veel consulenten de voorkeur te geven aan de intensieve begeleidingsgroep, ook wanneer zij een andere treatment zouden hebben gekozen voor de deelnemers waarvoor de treatments eigen regie en de standaard treatment minder geschikt bleken. Echter, blijkt het voor de consulenten lastiger te bepalen wat er moet gebeuren als ook de treatment intensieve begeleiding niet geschikt blijkt te zijn.

3.3 Focusgroep gesprekken met consulenten

Het focusgroep gesprek is onderdeel van de procesevaluatie welke gericht is op het proces van experimenteren met een andere benadering en uitvoering van de Participatiewet. Hoe gaan de direct uitvoerenden (consulenten en projectleider) maar ook de gemeente als organisatie om met de implementatie en ontwikkeling van het experiment? Het proces van experimenteren is een proces van verandering vooral ook in de benadering van de cliënt door de consulenten.
In deze focusgroep gesprekken wordt besproken wat de kansen en eventuele belemmeringen zijn van de andere bejegening in de relatie met de deelnemer en welke resultaten men verwacht en al dan niet bereikt. In het bijzonder wordt ingegaan op de handelwijzen van de consulent in de verschillende interventies en in welk opzicht ze van elkaar dienen te verschillen. Betekent ‘eigen regie’ dat men de deelnemer volledig moet loslaten of juist niet omdat het ‘leren van zelfredzaamheid’ meer contacten met de cliënt vraagt?
Ten tweede wordt ingegaan op de casuïstiek van de begeleiding waarbij specifieke problemen in de begeleiding zoals fraude, schulden en ongemotiveerd of onaangepast gedrag aan de orde wordt gesteld om ervan te leren. Mede daarvoor is ook een training aangeboden waarbij consulenten onder begeleiding van een professionele coach leren hoe ze een vertrouwensrelatie met de deelnemer kunnen opbouwen en hoe ze de zelfredzaamheid van deelnemers kunnen bevorderen.
Ook wordt ingegaan op de context waarbinnen het vertrouwensexperiment plaatsvindt, dat is de gemeentelijke organisatie, de politiek, de media en alle betrokkenen daaromheen. Deze gesprekken

vinden plaats na 6 tot 9 maanden in het bijzijn van de projectleider en de strategische beleidsadviseur die het experiment beleidsmatig en uitvoeringstechnisch mede coördineert.
In totaal zijn er 4 focusgroep gesprekken gehouden. Het eerste gesprek vond plaats op 6 december 2017 nadat het experiment enkele maanden van start was gegaan en het eerste gesprek met de cliënt was gehouden. Het tweede gesprek heeft op 3 oktober 2018 plaatsgevonden, het derde gesprek volgde op 12 juni 2019 en het vierde gesprek op 30 oktober 2019.

3.3.1 Topics

Tijdens de focusgroep gesprekken is gebruik gemaakt van een PowerPoint presentatie als leidraad voor het gesprek. In deze PowerPoint presentatie werden 5 stappen onderscheiden die ontleend zijn aan de stappen in het proces van experimenteren. Deze 5 stappen zijn:

- Stap 1. De beleidsuitgangspunten en het doel van het experiment
- Stap 2. Achterliggende theorie/filosofie en wetenschappelijke vragen
- Stap 3. Van theorie naar RCT-design voor een niet-AMvB experiment
- Stap 4. Van design naar implementatie: de drie treatment groepen; de relatie met de cliënt; de eigen rol als coach etc.
- Stap 5. Van implementatie naar uitkomsten: de lokale en politieke context, de randvoorwaarden en de uitkomsten voor de cliënt.

Aan de hand van deze stappen zijn er in de PowerPointpresentatie topics opgenomen die meer inzicht bieden in de ervaring van de consulenten met deze stappen.

3.3.2 Analyse methode

Elk van de vier focusgroep gesprekken zijn door middel van een mobiele telefoon opgenomen, met toestemming van de deelnemers aan het gesprek. Na opname zijn deze gesprekken anoniem getranscribeerd en vervolgens bewerkt tot een samenhangend verslag waarin de structuur van de stappen is aangehouden. In de praktijk lopen de stappen echter tijdens het gesprek soms door elkaar heen, waardoor bepaalde onderwerpen op meerdere plaatsen terug kunnen komen. Vervolgens zijn de verslagen van deze drie focusgroep gesprekken met elkaar vergeleken, met als uitgangspunt de structuur van de verschillende stappen en is er een analyse gemaakt van wat de standpunten op deze stappen waren en hoe deze veranderden gedurende het proces. Ook is er in de analyse nadrukkelijk aandacht besteed aan zogeheten ‘implementatie fouten’, wat inzicht geeft over de mate waarin de treatments daadwerkelijk zijn uitgevoerd, zoals ze theoretisch bedoeld waren, of dat de praktijk van uitvoering er anders uit zagen, wat dan de implementatie fouten zijn.

3.3.3 Resultaten

Tijdens de vier focusgroep gesprekken die zijn gehouden, zijn er verschillende stappen besproken. In praktijk valt wat wordt besproken vaak onder meer dan een stap, omdat de onderwerpen elkaar deels overlappen. In deze analyse is er voor gekozen toch zoveel mogelijk de opgenomen gesprekken te structureren aan de hand van de stappen in de gebruikte PowerPoint presentatie.

Stap 1: beleidsuitgangspunten en doel experiment

In deze eerste stap is besproken of de consulenten de uitgangspunten van het experiment en het doel (nog) helder voor ogen hebben. Daarnaast is er gevraagd of er in de praktijk al duidelijk wordt welke

aanpak, belonen of bestraffen, beter werkt en voor welke groepen. In focusgroep 1 en 2 is deze stap het meest uitgebreid besproken.
Tijdens het eerste gesprek komt naar voren dat er in Oss veel steun en enthousiasme is voor het experiment onder de consulenten. Dat geldt voor zowel de uitgangspunten, de filosofie en de uitwerking. Deels kwam dit voort uit onvrede over de oude manier van werken waarin er onvoldoende maatwerk was, te weinig middelen en tijd en een negatief beeld van de cliënt. De nieuwe aanpak vraagt volgens de consulenten ook wel wat van de cliënten, namelijk een andere denkwijze en bewustwording bij de deelnemer dat zij iemand mogen zijn en hun eigen identiteit mogen verwerven.
De consulenten geven aan dat er een aantal belangrijke verschillen zijn tussen de eigen regie en intensieve begeleidingsgroep enerzijds en de standaardbegeleiding anderzijds. Zo is er bij de eigen regie en intensiveringsgroep meer tijd en aandacht voor de cliënt en tijdens het experiment wordt er niet van consulent gewisseld, waardoor de vertrouwensband kan worden opgebouwd met de deelnemer. Consulenten merken ook dat de aandacht en het vertrouwen wat ze aan de deelnemers geven weer terugkomt in vertrouwen naar hen toe. Een ander positief punt is dat cliënten ook wel eens wat mogen uitproberen en bij mislukking zitten er dan niet direct nadelige financiële gevolgen aan vast zitten voor de cliënt. Ook is de toon in de communicatie met deelnemers bijvoorbeeld via brieven erg veranderd en informeler geworden. Dit is een groot verschil met de oorspronkelijke dienstverlening. Deelnemers zijn hierdoor minder angstig en wantrouwend, waardoor zij meer open zijn in gesprekken. Er is dan meer ruimte om te bespreken waarom iets niet is gelukt. Toch wordt tijdens het derde focusgroep gesprek duidelijk dat de standaardprocedure nog steeds is dat de consulent op basis van het gesprek moet *beoordelen* of er een sanctie moet worden toegepast. Volgens de consulenten zou dat moeten veranderen omdat de term 'beoordelen' een machtsverhouding impliceert en de insteek eigenlijk 'het er samen over te hebben' zou moeten zijn. De vraag wordt gesteld of burgers in de participatiewet nog wel ruimte hebben om zichzelf te zijn in de afhankelijkheidsrelatie die ze aangaan met de overheid bij het krijgen van een uitkering.
Ondanks de positieve veranderingen die de consulenten noemen, blijft er tijdens het eerste focusgroep gesprek een gevoel bestaan dat de organisatie zelf graag vooruit loopt op de resultaten en de focus legt op de uitstroom naar werk, terwijl er meer oog zou moeten zijn voor wat het doet met de deelnemer en zijn of haar beleving.

Stap 2 achterliggende theorie
In deze tweede stap lag de focus op de achterliggende theorie van het experiment en hoe dat werkt in de praktijk. Op deze stap is met name in het eerste en tweede focusgroep gesprek ingegaan.
In de achterliggende theorie van het experiment is het opbouwen van een vertrouwensband met de cliënt een belangrijk aspect om een cliënt beter te kunnen begeleiden. In de praktijk blijkt dat het opbouwen van deze vertrouwensband niet altijd makkelijk is en tijd vergt vanwege het grote wantrouwen van de cliënt naar de gemeente toe. Consulenten geven aan dat het uiteindelijk wel echt beter werkt in de begeleiding van de cliënt. In eerste instantie zijn cliënten achterdochtig en wantrouwend en vragen zich af of de andere bejegening wel echt is, vooral als zij al eerder een maatregel hebben gehad. Later zijn ze verbaasd en als ze merken dat er naar hen wordt geluisterd, dan verdwijnt het wantrouwen en ontstaat het gesprek. Er wordt hiermee volgens de consulenten een nieuw model van coaching ontwikkeld. Wel wordt er benoemd dat het met sommige cliënten nog lastiger is om een relatie op basis van vertrouwen

aan te gaan, bijvoorbeeld doordat zij bepaalde persoonlijkheidsstoornissen hebben of een andere culturele achtergrond. Voor hen zou een andere aanpak wellicht beter zijn. In sommige gevallen wordt er de consulenten zelf, maar ook door andere betrokken partijen twijfel geuit over of het goed is om een klant in het experiment te houden. In zulke gevallen wordt de contactregistratie wel bijgehouden, maar wordt geaccepteerd dat er geen arbeidsperspectief is en dat er geen vragenlijst wordt ingevuld.
In het derde focusgroep gesprek komt naar voren dat de meeste cliënten die op dat moment nog in begeleiding zijn ondanks hun problemen wel voor ogen hebben waar ze heen willen. Consulenten geven ook aan dat de deelnemers binnen het experiment de ruimte voelen om hun plan aan te passen en dat is positief want daardoor groeien ze in eigen regie.

Stap 3: design van de treatments
Tijdens de gesprekken is er in deze stap aandacht besteed aan het design van de verschillende treatments en de mate waarin deze treatments op de beoogde manier in de praktijk uitvoerbaar zijn.

Onderscheidend vermogen van de treatments
Tijdens het eerste focusgroep gesprek wordt duidelijk dat de consulenten bij de uitvoering van de treatments tegen een aantal problemen aanlopen. Zo blijkt dat cliënten die in de eigen regie groep zitten meer begeleiding vragen dan verwacht, waardoor er ook vaker afspraken worden ingepland. Vooral in de beginfase van het experiment hebben deze cliënten meer behoefte aan contact, omdat hen moet worden aangeleerd om zelfredzaam te zijn. Het onderscheid met de intensieve begeleidingsgroep komt hierdoor onder druk te staan, hoewel dit standpunt in de latere focusgroepen wordt bijgesteld. Dan zeggen de consulenten juist duidelijk voor ogen te hebben wat het verschil is tussen de treatmentgroepen.
Het tweede probleem wat werd aangekaart is dat de standaard treatment onderhevig is aan invloeden vanuit het experiment, de organisatie en ontwikkelingen in de samenleving, waardoor het zuiver meten van de effecten bemoeilijkt kan worden. Als aanvulling hierop wordt in het tweede gesprek benoemd dat het voor consulenten erg moeilijk is de verleiding te weerstaan om ook in de reguliere treatment de omgang met de deelnemer toch iets aan te passen, omdat er veel positieve dingen gehoord worden over de eigen regie en de intensieve begeleidingsgroep. De uitvoeringspraktijk van de reguliere aanpak is volgens de consulenten zelfs compleet anders dan 2 jaar geleden en lijkt steeds meer op de andere twee treatments. In het experiment is hiermee rekening gehouden, omdat er niet alleen vergeleken wordt tussen de standaardgroep en de eigen regie en intensiveringsgroep, maar ook met de referentiegroep, die niet deelneemt aan het experiment.
Consulenten merken tijdens het eerste gesprek ook op dat de training die zij hebben ontvangen niet specifiek was toegespitst op de verschillende treatments, maar meer op zelfredzaamheid in het algemeen. Consulenten zouden graag zien dat er meer aandacht wordt besteed aan het verkrijgen van inzicht in de verschillen tussen de treatments.
Een ander punt is dat de manier van begeleiden volgens de consulenten sterk afhangt van de cliënt, maar ook van de consulent. Om het experiment helemaal zuiver te laten verlopen, hadden consulenten eigenlijk ook willekeurig toebedeeld moeten worden, maar in de praktijk is rekening gehouden met de wensen, motivatie, capaciteiten en persoonlijkheden van de consulenten.

Ten slotte lopen consulenten tegen het probleem aan dat zij meer integraal moeten handelen binnen het experiment, maar dat men niet altijd deskundig is op die andere vakgebieden. In zo'n geval zou iemand moeten worden doorverwezen naar andere meer deskundige begeleiders.

Ervaringen per treatment

- Eigen regie

Bij eigen regie blijkt vertrouwen een centraal onderwerp. Niet alleen vertrouwen van de consulent in de cliënt maar ook van de cliënt in zichzelf. In het experiment is eigen regie niet bedoeld als het volledig loslaten van de cliënt, maar de cliënt juist motiveren om een eigen plan te maken en zelfredzaam te worden. Dat is een leerproces geweest voor zowel cliënt als consulent. Consulenten moesten hiervoor de 'eigen arrogantie' van 'beter weten hoe het moet' loslaten en accepteren dat de cliënt het op zijn eigen manier en tempo gaat doen. Daarmee blijft het nog steeds wel maatwerk, omdat de ene cliënt meer ondersteuning nodig heeft dan de andere. Maar de basis dat de deelnemer wordt gevraagd om mee te denken is positief. Het is dan ook bij deze groep van groot belang dat het traject niet wordt ingevuld voor de deelnemer zonder eigen inbreng. Hierbij mag de consulent wel aangeven wat de mogelijkheden zijn, zo lang hij/zij de cliënt het initiatief laat nemen om die mogelijkheden te benutten.

- Intensieve begeleiding

Consulenten vinden de intensieve dienstverlening heel duidelijk. Volgens hen is daar ook weinig aan veranderd sinds de start van het experiment. Consulenten geven aan dat het goed is om regelmatig contact te hebben en te houden met de cliënt. Zo wordt een vertrouwensband opgebouwd en ook de consulent zelf leert de klant beter kennen en krijgt meer begrip voor de situatie, maar past zich ook eerder aan het tempo van de cliënt. Er wordt meer vraaggericht gewerkt en er kan beter worden ingespeeld op de wensen en mogelijkheden van de deelnemer. Er is meer tijd om de deelnemer aandacht te geven en het persoonlijke gesprek aan te gaan.

- Standaard begeleiding

Zoals eerder genoemd, hebben de consulenten aangegeven dat de reguliere begeleiding sterk veranderd is vergeleken met 2 jaar ervoor. Toch heeft een consulent bij de standaardtreatment vooral als taak de uitkering te beoordelen (slechts de inkomenskant dus) en vaak werken er meerdere consulenten aan eenzelfde traject voor een deelnemer. Ook is er vaak binnen dit traject een overdracht moment waarbij de klant van de consulent naar een werkconsulent gaat en hierdoor is de klant vaak een tijdje uit beeld bij de consulent die deel neemt aan het experiment. De consulenten van de standaardtreatment voelen wel de ruimte om anders met mensen om te gaan, maar dat was eigenlijk ook voor het experiment al zo. Wel is er vergeleken met de ander treatment groepen minder tijd om met de deelnemers te spreken.

Stap 4 de treatment groepen, relatie cliënt, rol consulent

In stap 4 is een groot aantal onderwerpen besproken. Er is besproken of de treatments in de praktijk effectief zijn voor de deelnemers, hoe coaches hun eigen rol zien, welke positieve en negatieve ervaringen de coach heeft met het uitvoeren van de treatments, welke behoefte er is aan ondersteuning van de consulenten, verwachtingen ten aanzien van het management en de organisatie, de eigen motivatie en werkgeluk van de consulent en de leereffecten van klant en consulent.

Effectiviteit treatment

Wat betreft de effectiviteit van de treatments voor de deelnemers, werd eerder al besproken dat veel consulenten al veranderingen zien in de mate van vertrouwen van de cliënten, de mate waarin zij eigen regie pakken en de mate waarin zij open en eerlijk kunnen zijn tijdens de gesprekken. Ook blijkt uit het tweede focusgroep gesprek dat zelfs cliënten die aanvankelijk het experiment wilden gebruiken als excuus om niets te hoeven doen, uiteindelijk door het ontstaan van de vertrouwensband toch in actie kwamen. Ook deelnemers waarvoor het doel werk moeilijk te bereiken is, worden makkelijker benaderbaar, omdat het doel naar beneden kan worden bijgesteld, passend bij wat zij kunnen.

Tijdens het derde focusgroep gesprek is er gesproken over in hoeverre er een plan met de deelnemer wordt gemaakt. De consulenten geven aan wel plannen te maken samen met de klant, maar deze worden niet zwart-op-wit vastgelegd. Dat geeft volgens de consulenten ruimte om het plan aan te passen als dat nodig is en mee te laveren met de klant. Er worden wel afspraken gemaakt met deelnemers en als ze zich daar niet aan houden, dan wordt gevraagd waarom ze zich niet aan hun eigen afspraak houden, zonder daar verdere consequenties aan te verbinden (bij eigen regie en intensivering). De consulenten maken wel rapportages waarin de afspraken staan en in elk contact met de klant wordt er opnieuw aandacht besteed aan het plan. De consulenten zijn van mening dat het niet onderleggen van het plan met een beschikking of contract, de klant juist meer eigen regie geeft. Zo kunnen ze zelf een eerlijk antwoord formuleren als iets niet is gelukt bijvoorbeeld en iets nieuws bedenken. Ook denken de consulenten dat een schriftelijk plan niet nodig is als de consulenten het goede voorbeeld geven van commitment en vertrouwen in de klant. Soms is het voor de consulenten nog wel moeilijk om niet teveel te gaan sturen.

Ook wordt tijdens dit derde gesprek opgemerkt dat consulenten soms wel de maatregelen hebben gemist, vooral in gevallen waarin er geen stappen werden gezet. Er wordt door de consulenten verschillende gedacht over maatregelen. De een vindt het een teken van onmacht en dat het niets oplost, terwijl de ander zegt dat het effect van een maatregel afhankelijk is van het individu die hem opgelegd krijgt. Samengevat heeft het geven van een maatregel alleen zin als er verwacht wordt dat het effect heeft.

Tijdens het vierde en afsluitende focusgroep gesprek geven de consulenten aan dat cliënten waar voorheen niet veel in geïnvesteerd zou zijn omdat ze het profiel zorg hadden, ook juist door de begeleiding nu resultaat hebben behaald op uitstroom. Een kanttekening is wel dat deze mensen voorheen ook wel uitstroomden maar dan om andere redenen dan voor werk.

De rol van de consulent

Uit het eerste gesprek met de consulenten komt naar voren dat men een soort ‘clash van rollen’ ervaart als gevolg van het experiment. De nieuwe rol van de consulent is geheel anders dan de oude rol waar vooral de nadruk lag op de rechtmatigheidscontrole en staat op gespannen voet daarmee. Er wordt geprobeerd daar een balans in te vinden door open te communiceren en het gesprek met de deelnemer

aan te gaan. In de praktijk blijkt het lastig in te schatten hoe deze nieuwe rol moet worden ingevuld bij fraude of ongewenst gedrag. Daarbij is het belangrijk om deze spanning te signaleren en binnen de groep en de organisatie aan te kaarten.
Tijdens het uitvoeren van de nieuwe rol ervaren consulenten dat ze deze manier van werken prettiger vinden omdat ze van mens tot mens kunnen spreken en ze ervaren minder druk. Ook voelen deelnemers zich trots op zichzelf voor de stappen die ze zelf zetten. Consulenten ervaren dit proces als een leerproces waarin voortdurend aanpassing en afstemming nodig is. Ze waarderen de samenwerking in een team met enthousiaste mensen.
Toch blijft er rondom de treatment eigen regie onduidelijkheid bestaan over de rol van de consulent, hoe ver reikt deze? Vooral in lastige situaties waarin cliënten niet zelfredzaam zijn en de afstand tot werk nog ver is, wat kan en moet de consulent dan doen in zo'n situatie? Men is van mening dat het gesprek met de klant moet gaan over welke stappen er nodig zijn en welke andere instanties daarvoor mogelijk moeten worden ingeschakeld. Een consulent hoeft dus niet alles zelf te doen, maar moet wel op de hoogte blijven van wat er gebeurt en waar mogelijk een bijdrage leveren. Daarnaast is er onduidelijkheid over hoe lang de periode mag zijn dat een klant binnen eigen regie niets laat horen. Als de consulent dan toch het initiatief moet nemen voor het gesprek, dan lijkt de treatment weer meer op intensieve begeleiding. Eigen regie betekent echter niet dat de klant wordt losgelaten maar juist ook in gesprek blijft met de consulent als er nog geen sprake is van zelfredzaamheid.
Tijdens het derde focusgroep gesprek is er met de consulenten gesproken over hoe en of zij hun eigen gevoel uiten naar de klant over hoe het proces verloopt. De consulenten geven aan dat ze dat wel doen en dat de klant het gevoel vaak ook wel snapt en dan ook open kan vertellen wat er aan de hand is. Dan komen er soms problemen aan het licht waardoor de consulenten voor een dilemma komen te staan over hoe ver hun eigen rol reikt. De rol van de consulent komt dan meer in de buurt van een maatschappelijk werker en niet meer alleen de rechtmatigheid en doelmatigheid. Soms is het nodig om een andere partij in te schakelen, zoals de GGZ, maar ook dan kan de consulent ondersteunend zijn en een luisterend oor bieden. Op deze manier kunnen bepaalde drempels en problemen worden weggenomen waardoor er een beter toekomstperspectief ontstaat. Belangrijk is dat de klant gesprekspartner is en blijft en er een open gesprek kan worden gevoerd.

Tijdens het derde gesprek wordt er geconcludeerd dat er een aantal aspecten steeds naar voren komt als er wordt gesproken over goede begeleiding en dat zijn tijd, aandacht en breder kijken dan alleen de snelste weg naar werk. Maar daarbij is het wel zoeken naar welke dienstverlening er bij de sociale dienst hoort. Ook de interne samenwerking en omgang met deelnemers is een punt van aandacht.

Behoefte aan training/intervisie
Wat betreft de training en intervisie sessies die de consulenten hebben gehad, is al eerder benoemd dat de training nuttig was, maar er ook behoefte was om deze meer toe te spitsen op de afzonderlijke treatments. Ook is er in Oss het initiatief genomen om de training frequenter te gaan geven, zodat de ideeën vers blijven, omdat de theorie na een tijd toch wat vervaagd. Men geeft aan dat de gesprektraining van Roland Blonk voor de consulenten de basis geweest om anders in de spreekkamer te zitten.

Arbeidsvreugde
De mate van arbeidsvreugde wordt als positief beoordeeld door de consulenten, mede vanwege de onderlinge betrokkenheid bij de aanpak. De consulent van de standaardtreatment ervaart wel enige afstand met de consulenten van de andere treatments, maar dat beïnvloedt de arbeidsvreugde niet. Verder wordt er een lagere werkdruk ervaren, voornamelijk voor de eigen regie groep. Het proces is er wel een van vallen en opstaan en wellicht is het experiment te kort van duur om het helemaal goed te doen.

Leereffecten
Wat betreft de leereffecten voor de consulenten en de deelnemers, geven de consulenten aan veel te hebben geleerd over zichzelf en ook over de deelnemers van het experiment. Het inzicht is verkregen dat het beter is om de deelnemer zelf met een plan te laten komen. Aan het einde van het eerste jaar is de verwachting dat het tweede jaar een verdiepingsslag wordt van de begeleiding. Het is wel de vraag wat er gaat gebeuren met de groep cliënten die minder goed in staat is tot eigen regie en de mensen weer terug instromen in de bijstand na uitstroom. De consulenten zijn wel van mening dat het experiment effect heeft: er is sowieso een betere verstandhouding bereikt met de klant.

Als er wordt gevraagd naar wat consulenten nog meer hebben geleerd, dan wordt er genoemd dat het beter werkt als hij of zij gas terug neemt en zich aanpast aan het tempo van de klant. Daarnaast werkt het ook beter om te richten op de vraag vanuit de klant, in plaats van het aanbod van trajecten. In de reguliere aanpak werden consulenten er op gestuurd om mensen aan te melden voor bepaalde trajecten en ook vanuit de accountmanagers en de vacatures die zij bieden. Nu wordt er vanuit de vraag van de klant gewerkt. Dit heeft tot gevolg gehad dat de uitgaven voor trajecten sterk gedaald is binnen de treatments eigen regie en intensieve begeleiding. Toch wordt er wel benoemd dat ondanks dat het goed is om op de vraag te richten, het wel belangrijk is dat de consulent het aanbod goed kent, om zo de mogelijkheden te kunnen voorleggen aan de klant. Verder leren de consulenten ook veel van elkaar en de aanpak van de ander.

Consulenten geven ook aan zelf trajecten vorm te geven en meer onderzoeken wat mogelijk is als de klant zelf met een idee komt. Ook door het positieve in het proces te bekrachtigen, raakt de klant gemotiveerd om zelf ook op onderzoek uit te gaan. Doordat de consulent minder bezig is met het nadenken over de mogelijkheden en het invullen voor de klant, blijft er meer ruimte over om te signaleren wat er gebeurt. Hierbij is het volgens de consulenten wel van belang dat de organisatie de ruimte geeft om 'buiten de lijntjes te kleuren'. Daar moeten dan echter wel andere afdelingen uit het sociale domein ook in meegenomen worden.

Tijdens het derde gesprek komt naar voren dat het in de praktijk voor de treatments eigen regie en intensieve begeleiding het eerste halfjaar heel hard werken was om in te zien dat de aanpak heel anders is dan wat er al werd gedaan. In de standaardbegeleiding was en is de gesprekvoering ten opzichte van de nieuwe aanpak volledig anders volgens de projectleider. Daarnaast zijn bij de standaardbegeleiding nog steeds veel deelnemers niet goed in het vizier, bijvoorbeeld omdat ze nog niet gekoppeld zijn aan een werkconsulent. Er is nu wel een nieuw instroomproces opgezet bij de werkconsulenten, maar ook daar is

er discussie over welke vragen er gesteld moeten worden tijdens de gesprekken. Binnen deze instroomrondes is professioneel handelen belangrijk en daarom hebben de mensen die de intake doen scholing, training en intervisierondes gehad.

Tijdens het vierde gesprek komt naar voren dat de consulenten benieuwd zijn of hun eigen mensbeeld wat ze hadden van de deelnemers is veranderd gedurende het experiment, maar ook of het beeld wat de deelnemers hebben van de consulenten is veranderd. Uit de voorlopige resultaten van de consulent vragenlijsten kwam naar voren dat de consulenten de leereffecten van zowel de deelnemer als van zichzelf een magere voldoende gaven. Desgevraagd geven de consulenten aan dat een reden daarvoor was dat ze hun leereffecten per klant moesten aangeven en dat is vaak lastig, juist omdat ze leren van de begeleiding van hun deelnemers in het algemeen. Daarnaast hebben ze de leereffecten laag ingeschat omdat ze erachter kwamen dat ze juist nog veel te leren hadden en daar nog steeds niet mee klaar zijn. Er is meer inzicht gekomen in wat de consulenten zelf nog moeten leren. Ze zijn zich meer bewust geworden van hun eigen handelen en wat er wel en niet in hun eigen rol past. En ze hebben meer begrip voor de aanpak van collega's gekregen.
De consulenten geven aan dat hun eigen beeld van de klant niet echt is veranderd, maar dat ze de klant wel beter hebben leren kennen. Juist als er frustratie ontstaat bij de consulent over de trage vooruitgang, dan helpt het om in contact te blijven en meer begrip te krijgen voor de klant.

Stap 5 de lokale en politieke context, de randvoorwaarden en de uitkomsten voor de cliënt.
In stap 5 is er gesproken over de context waarbinnen het experiment plaatsvond. De omstandigheden waarin consulenten hun werk deden, bijvoorbeeld de hoeveelheid caseload maar ook de interne organisatie van het experiment en de rol van het management. Daarnaast welke rol de politieke en lokale context heeft gespeeld.

Organisatie
Tijdens het eerste focusgroep gesprek wordt een integrale aanpak als een belangrijk aspect benoemd om de filosofie van het experiment goed uitvoerbaar te maken. Bij een integrale aanpak worden verschillende instrumenten uit verschillende domeinen ingezet voor de klant. Een ander aspect wat belangrijk is om maatwerk te kunnen leveren is een goede samenwerking met accountmanagers van organisaties. Volgens de consulenten is het de taak van een accountmanager om de arbeidsmarkt voor te bereiden op een doelgroep die aan het veranderen is. Namelijk, de mensen in begeleiding zijn niet direct bemiddelbaar en daarin moeten accountmanagers een stap richting consulenten zetten. Daarnaast is de beschikbaarheid van een plek die gefinancierd wordt van belang, zodat iemand aan de slag kan. Wel is er binnen het experiment veel meer ruimte voor maatwerk. Aan de andere kant zijn er ook mensen waarvoor een standaardtraject heel goed kan werken.
Tijdens het tweede gesprek komt naar voren dat de organisatie nog wel moet wennen aan het experiment. Zo wordt gesteld dat de consulenten die deelnemen aan de experimenten gezien worden als een 'elite groep'.
Tijdens het vierde gesprek wordt meer uitgebreid teruggeblikt op de context waarin het experiment zich heeft afgespeeld en welke ontwikkelingen zich binnen de organisatie hebben voorgedaan. Daarnaast worden er aanbevelingen gedaan voor de toekomst van de Participatiewet.

De consulenten geven aan dat er binnen de organisatie een aantal grote veranderingen zichtbaar is. Echter, omdat er verschillende pilots tegelijk lopen is het moeilijk te zeggen welke het gevolg zijn van het experiment. Het algehele beeld wat wordt geschetst van de organisatie is dat die erg bereid is de eigen dienstverlening te verbeteren. De organisatie probeert binnen het sociaal domein meer te werken vanuit de vraag van de burger. Het experiment is een verdieping van hoe je een klant leert om eigen regie te nemen. Dat begint met de klant vragen te stellen. Maar dat gebeurt niet alleen bij de klant, maar ook binnen de organisatie zelf steeds meer. Zo stellen managers steeds vaker vragen, in plaats van dat ze een opdracht geven. Aan de ene kant is dat een goede ontwikkeling, maar aan de andere kant is ook het kader en de begrenzing weg, waardoor er een soort continue instabiele factor is waarin iedereen een richting moet zoeken. De uitkomsten van het experiment kunnen daarin meer richtinggevend zijn.
Een andere vraag waar de organisatie mee bezig is, is hoe de rol van de consulent eruit moet zien over een of twee jaar. Om daarmee aan de slag te gaan is er binnen de organisatie een soort leertafel opgericht waarin verschillende expertises bij elkaar worden gebracht om zich over deze vraag te buigen. Zo zijn er verschillende sterke teams gevormd, en is er voor een aantal van deze teams ook team coaching. Aan de ene kant geven al deze ontwikkelingen de consulenten nieuwe energie, maar de veelheid ervan is soms ook vermoeiend en er wordt uitgekeken naar een periode waarin alles zich wat stabiliseert.
Het gevoel van de consulenten is dat er een cultuurverandering gaande is in de organisatie, maar ook in de organisaties die het raakt en misschien zelfs ook in de werkgeversbenadering. Het gevoel is ook dat de organisatie faciliteert in het goed kunnen uitvoeren hiervan en dat ook zij gedurende het experiment er meer vertrouwen in kregen.
Een belangrijke factor hierin is ook dat waar er voorheen vooral over uitstroom gesproken werd en er weerstand was tegen de keuzes die werden gemaakt vanuit regelluw, er nu nooit meer over uitstroom wordt gesproken. Er is vanuit het management veel meer interesse voor wie de klant is en wat die nodig heeft.
Een kanttekening is echter wel dat deze werkwijze gevolgen heeft voor het aantal taken wat de consulent heeft en hoe dat werkbaar blijft en hoe dat het beste georganiseerd kan worden. Een aanbeveling die daaruit naar voren komt, is dat het fijn is om met minstens twee mensen aan een caseload te werken. Op die manier is er meer continuïteit in de caseload omdat men elkaars afwezigheid kan opvangen. Nu worden deelnemers verplaatst naar een andere consulent als ze in een ander profiel terecht komen. Een oplossing daarvoor kan zijn om wijkgericht te gaan werken, en daar verschillende disciplines samen te brengen is een groter team. Een gebied verandert minder snel dan een profiel en daardoor krijgt de cliënt met minder verschillende gezichten te maken. Daarnaast sluit deze manier van werken ook aan op de zorg in de wijk en vergemakkelijkt het de samenwerking met de sociale teams. De uitdaging is dan wel hoe je de werkwijze in verschillende wijken met elkaar verbindt. Dat zou bijvoorbeeld kunnen door samenwerking in sterke teams. Deze werkwijze zou kunnen voorkomen dat meerdere mensen aan dezelfde case werken zonder het van elkaar te weten.
De vraag is of de nieuwe werkwijze van meer contact en tijd met de cliënt ook meer geld gaat kosten. De consulenten geven aan dat dat misschien op de korte termijn zo is, maar misschien op de lange termijn niet. Het ligt aan de begrenzing van werkzaamheden en aan de samenwerking met andere partijen, die ook bepaalde taken zouden kunnen uitvoeren. De uitvoering hoeft dan niet zozeer meer geld te kosten, maar het onderzoeken en het verdelen daarvan wel. De consulenten van de treatments eigen regie en intensieve begeleiding geven aan dat ze minder geld aan trajecten hebben uitgegeven en meer hebben

teruggeven aan de cliënt om eigen regie te pakken. Deze besparing op het P-budget zou bijvoorbeeld ingezet kunnen worden voor een lagere caseload. Ook is er in deze twee treatments de afgelopen 2 jaar geen enkel bezwaar gemaakt door de klant.

Aanbevelingen

- Alle consulenten zijn het ermee eens dat het voor een goede begeleiding van de cliënten noodzakelijk is om meer tijd beschikbaar te hebben om de gesprekken te kunnen voeren en een relatie te kunnen opbouwen met de cliënt.
- Ook is een aanbeveling om minder rigide naar de wet te kijken en meer naar wat de uitvoeringspraktijk al is. Bij het vormen van nieuw beleid moet er gedacht worden vanuit de juiste uitgangspositie. De verwachting is nu teveel gericht op dat gemeentes moeten handhaven of anders ontheffen. Er zou minder aandacht mogen naar de arbeidsverplichting en meer naar artikel 10, dat iemand begeleiding krijgt als hij/zij het zelf niet kan. En de nadruk op onderzoek wat iemand kan en nodig heeft. Maak het meewerken aan dat proces van onderzoek een verplichting.

3.3.4 Samenvatting

In dit gedeelte worden een aantal belangrijke punten uit het voorgaande verslag kort samengevat per onderwerp.

Effect van de treatments eigen regie en intensieve begeleiding:

- Bij de experimentgroepen eigen regie en intensivering is er meer tijd en aandacht voor cliënt en wordt er niet van consulent gewisseld. Dit draagt bij aan het opbouwen van een vertrouwensband tussen consulent en cliënt.
- Cliënten mogen dingen uitproberen en het mag mislukken zonder dat er financiële consequenties aanhangen. Dat stimuleert eigen regie en vertrouwen. Ook het niet onderleggen van een plan met een contract of beschikking draagt hieraan bij.
- Door de veranderde manier van bejegening, ook op papier, zijn cliënten minder angstig en wantrouwend en meer open in gesprekken, wat de begeleiding ten goede komt.
- In sommige gevallen is het lastiger om een vertrouwensband op te bouwen, bijvoorbeeld met mensen met bepaalde persoonlijkheidsstoornissen of een andere culturele achtergrond.
- Eigen regie werkt vooral als er eerst een intensievere periode van begeleiding aan vooraf gaat waarin men wordt geleerd zelfredzaam te zijn.
- Ook deelnemers die moeilijker in beweging zijn te krijgen of voorheen niet eens zouden zijn begeleid naar werk vanwege bijvoorbeeld een zorgprofiel, boeken resultaat in het experiment.
- Of het toepassen van een maatregel zin heeft, hangt sterk af van het individu die de maatregel opgelegd krijgt. Het geven van een maatregel heeft alleen zin als verwacht wordt dat het effect heeft.
- Deelnemers zijn trots op zichzelf en de stappen die ze zetten vanuit hun eigen plan.
- De klant moet gesprekspartner zijn en blijven. De focus van de begeleiding moet verschuiven van ‘beoordelen’ naar ‘het er samen over hebben’.

Training en uitvoering:

- Training en intervisie zijn belangrijk, maar de opgedane kennis moet wel regelmatig onderhouden worden
- Integraal handelen is belangrijk tijdens de begeleiding, maar is soms lastig omdat de consulent niet altijd deskundig is op elk vakgebied
- Binnen de nieuwe rol van de consulenten bij de groepen eigen regie en intensivering is het lastig in te schatten hoe deze rol moet worden ingevuld bij fraude of ongewenst gedrag.

Leereffecten consulenten:

- Meer vraaggericht werken vanuit wat de cliënt nodig heeft meer effect dan de klant te sturen.
- Aanpassen aan het tempo van de klant.
- De nieuwe manier van begeleiden is fijner voor de consulent omdat ze van mens tot mens kunnen spreken en minder druk ervaren.
- Focus op goede begeleiding in plaats van op uitstroom.
- Nog niet altijd duidelijk waar de rol van de consulent stopt. De begrenzing en het kader is soms onduidelijk.
- Consulenten leren veel van andere consulenten, hebben meer inzicht in wat ze zelf nog moeten leren en meer begrip voor de aanpak van collega's (belang van teamwerk).
- Aspecten van goede begeleiding zijn: tijd, aandacht en breder kijken dan alleen de snelste weg naar werk.
- Het mensbeeld van de consulent of de klant is niet veranderd, maar ze hebben de klant wel beter leren kennen en meer begrip voor hun situatie gekregen.

Aanbevelingen voor de toekomstige werkwijze en rol consulent:

- Een integrale aanpak, expertises samenbrengen.
- Goede samenwerking met accountmanagers van organisaties.
- Beschikbaarheid van plekken waar mensen aan de slag kunnen.
- Met twee of meer mensen aan een caseload werken, meer continuïteit in begeleiding.
- Wijkgericht werken en daarmee aansluiten bij sociale teams.
- Taken van consulenten inventariseren en waar mogelijk verdelen tussen andere partijen waarmee wordt samengewerkt.

4. Bevindingen en conclusies

4.1. Opzet van het experiment en werving

De start

Oss was als gemeente samen met Apeldoorn in 2017 nu ruim 2 jaar geleden een van de voortrekkers in Nederland om een experiment met regelluwe bijstand te starten in de Participatiewet buiten de kaders van de AMvB om. Niet-AMvB gemeenten hebben meer vrijheidsgraden bij het ontwerp van het experiment dan de AMvB gemeenten doordat zij niet gebonden zijn aan de nogal als beperkend ervaren voorwaarden van artikel 83. Anderzijds hebben zij niet de mogelijkheid om de vrijlatingsbepalingen te verruimen. De gemeente wilde een nieuwe aanpak waarin wederzijds vertrouwen tussen cliënt en gemeente alsook eigen regie en maatwerk belangrijke uitgangspunten zijn van de begeleiding. Daarom wordt het experiment ook wel "Zelf aan het stuur en op maat" genoemd. Bij de opzet van het experiment is overigens aangesloten bij de opzet van de experimenten in de AMvB gemeenten. Dat betekent dat ook in Oss en de andere niet-AMvB gemeenten is gekozen voor een 'random control trial' (RCT) design waarbij deelnemers aselect worden toegewezen aan de 'treatments' of behandelingsgroepen (zie onderzoeksvoorstel: Muffels & van der Klink, 2017). De treatment groepen zijn wel verschillend doordat niet-AMvB gemeenten niet mogen experimenteren met een ruimere vrijlating maar dus wel met versoepeling van de verplichtingen (eigen regie) en een andere bejegening en begeleiding van de cliënt (maatwerk en intensieve begeleiding).

Werving en onderzoek: invulling van de vragenlijsten

De werving zou voor de helft uit het zittende bestand en voor de helft uit nieuwe instroom plaatsvinden. Voor de werving van de nieuwe instroom was 6 tot 9 maanden ingeruimd, tussen 1 oktober 2017 en1 juli 2018. Er is veel aan gedaan om de werving tot een succes te maken met voorlichtingsmateriaal, filmpjes, voorlichtingsbijeenkomsten etc. Uiteindelijk leidde dit tot 380 aanmeldingen waarvan 35 aanmeldingen niet voldeden aan de toelatingscriteria. Op 25 juni werd de laatste deelnemer aangemeld. Het streefaantal van 300 werd in de periode tussen 1 oktober 2017 en 1 juli 2018 gehaald. De invulling van de Qualtrics online vragenlijsten door de deelnemers bleek evenwel in de fase daarna geen sinecure. Niettemin is een respons van 62% gehaald, 85% in de nulmeting en 60% in de vervolgmetingen. Door de gunstige arbeidsmarkt nam de uitstroom naar betaald werk sterk toe waardoor het aantal deelnemers in begeleiding snel afnam maar tegelijkertijd ook de respons werd beinvloed. Eenmaal uitgestroomd werd de motivatie minder om de vragenlijst in te vullen. De consulenten vulden tijdens het proces van uitvoering ook een vragenlijst in en deze is voor nagenoeg 100% tijdens het experiment drie keer ingevuld voor alle deelnemers. De gunstige arbeidsmarkt was plezierig voor de deelnemers zelf, een flink groter deel vindt werk, maar daardoor liep het experiment sneller leeg dan verwacht. Het gevolg was dat de kansrijkste deelnemers naar werk uitstromen en de consulenten te maken kregen met deelnemers die moeilijker aan het werk kwamen. Dat maakte de begeleiding c.q. coaching na meting 2 vanuit het perspectief van de consulent een lastiger opgave.

Uitvoering van de treatments

Het experiment beoogt voor deze deelnemers een alternatieve aanpak en bejegening te ontwikkelen uitgaande van vertrouwen, eigen regie, keuzevrijheid, zelfredzaamheid en maatwerk. Niet de uitkomst staat centraal maar het proces van begeleiding en ondersteuning waarin mensen wordt geleerd zelfredzaam te worden, dus zelf aan het stuur en afgestemd op de noden en mogelijkheden van de deelnemer (vraaggericht maatwerk). Deze aanpak is anders en wordt door de consulenten als verfrissend, en vernieuwend ervaren. Het werk als professional werd er boeiender door, relaxter ook, en men ging met meer plezier naar het werk. Het 'echte' gesprek, aandacht op basis van vertrouwen en de eigen regie bij de deelnemer werden als de belangrijkste 'triggers' van de andere aanpak gezien en ervaren. Het werd gezien als een uitdaging. Die uitdaging betreft de treatment van 'eigen regie' maar ook de treatment van intensieve begeleiding waarbij de consulenten een andere aanpak voorstaan en de filosofie erop gericht was om mensen zelfredzaam te maken en de deelnemer te leren de eigen verantwoordelijkheid te nemen en die ook waar te maken. Daartoe hebben de consulenten op basis van de 'skills' methodiek een training gekregen van Roland Blonk waarin vaardigheden zijn geleerd om de zelfredzaamheid van de deelnemers te bevorderen. Eigen regie is dus geen 'loslaten' maar 'leren zelfredzaam te zijn of te worden' door het gesprek aan te gaan, het initiatief bij de deelnemer te laten en handelingsvrijheid te geven om zelf in actie te komen. Tegelijkertijd vond men het onderscheid zeker in het begin tussen eigen regie en intensieve begeleiding moeilijk te duiden en te implementeren. De consulenten in Oss werkten als koppels binnen dezelfde treatment en als team tussen de treatments aan het experiment, hetgeen door henzelf als een van de sleutels tot het gepercipieerde succes werd gezien. Op die manier was continuiteit in de begeleiding maar ook intervisie (elkaar bij de les houden) verzekerd. Door deze werkwijze kreeg men naar eigen zeggen geleidelijk aan meer grip op de treatments en het onderscheidende karakter van beide. Nochtans vinden we in het onderzoek over deze 2 jaar weinig significante verschillen tussen de treatments op basis van de informatie die deelnemers hebben gegeven in de drie vragenlijsten en de uitkomstmaten die we hebben gebruikt (geen verschillen bij subjectieve gezondheid, zelfredzaamheid, keuzemogelijkheden, vertrouwen in de instituties en armoede). Daarentegen vinden we wel duidelijke verschillen bij de consulenten tussen de treatments op basis van de in de procesevaluatie gebruikte uitkomstmaten. Dat uit zich met name in de motivatie voor de eigen treatment bij de deelnemer, de motivatie bij de deelnemer voor het inzetten voor het hoogst bereikbare doel, het verwachtte effect van de treatment voor de deelnemer maar door de consulent, het leereffect bij de deelnemer (zelfredzaamheid in gedrag) en het leereffect bij de consulent. Deze gepecipieerde verschillen tussen treatments uitten zich vooral aan de start van het experiment en later werden deze verschillen met de standaardgroep eerder kleiner dan groter. In de visie van de consulenten zelf zijn de treatments in vergelijking met de standaard treatment heel verschillend maar dat uit zich kennelijk nog niet in de uitkomsten bij de deelnemer.

Verandering tijdens de uitvoering

De wereld staat tijdens het experiment evenwel niet stil. De politieke en beleidscontext verandert maar er zijn ook leerfeffecten bij de coaches waardoor de treatments ook veranderen. Dat maakt het experiment complex want voor een juiste vergelijking zouden de treatments zoveel mogelijk hetzelfde moeten blijven. De verandering in uitvoeringspraktijk en beleidscontext beïnvloedt vermoedelijk ook de standaard treatment waardoor een zuivere vergelijking tussen de reguliere en alternatieve treatments in

het onderzoek werd bemoeilijkt. Het eindverslag gaat daar op verschillende plaatsen nader op in. Uit de proces- en productanalyses wordt duidelijk dat de standaard treatment is meegeevolueerd met de andere treatments. De lijntjes lopen in de tijd bij bijna alle drie treatment groepen en de meeste uitkomstmaten zowel in de product als de procesevaluatie min of meer gelijk op. We vinden significante treatment effecten in de procesevaluatie voor de standaardgroep met name in het interactiemodel met de tijd bij bepaalde uitkomstmaten en we vinden significante treatment effecten in het interactiemodel met de tijd in de productevaluatie bij bepaalde uitkomsten voor de deelnemer. Dit duidt erop dat er over de tijd iets veranderd is in de standaard treatment zelf. Dat is deels ook vanzelfsprekend omdat beleid meeevolueert met de context en de inhoud van de standaard aanpak over de tijd maar deels heeft het ook te maken met een soort "Hawthorne" effect, een effect van het experiment en het onderzoek naar de werkwijze van de consulent zelf. De vergelijking tussen de treatments en de standaardaanpak werd daarmee problematischer.

Uitstroom naar werk en matching

Bij de meting van de uitkomstmaat volledige uitstroom naar werk vonden we slechts kleine verschillen tussen de treatments en de standaardgroep. Deze waren in de orde van grootte van 1 tot 3%. Nochtans waren de uitstroom percentages voor de beide treatments maar dus ook voor de standaardgroep twee keer zo hoog als in de periode daarvoor 30% tegenover 15% in de tweejaarlijkse perioden 2015-2017 en 2016-2018. In de eerste tussenrapportage (Muffels e.a. 2018) zijn deze uitstroomcijfers voor dezelfde doelpopulatie (mensen die voldoen aan de toelatingscriteria voor het experiment) maar dan van niet-deelnemers berekend en gerapporteerd. Het vermoeden bestaat daarom dat de aanpak in de standaardgroep ook is veranderd tijdens het experiment. Voor de 'uitstroom naar werk' uitkomstmaat bleek dat probleem te kunnen worden opgelost door een vergelijking van de uitstroom niet alleen tussen de treatments en de standaardgroep maar met name met de referentiegroep, de groep van bijstandsgerechtigden die niet aan het experiment heeft deelgenomen maar wel tot dezelfde doelpopulatie behoort. Die vergelijking laat zien dat de consulenten een twee keer zo hoge uitstroom hebben gerealiseerd dan gemiddeld genomen het geval is, 30% in plaats van 15% over de experimentperiode van 2 jaar. Omdat de referentiegroep niet is gerandomiseerd weten we niet zeker of dit een causaal effect is. Daarnaast heeft er zelfselectie plaatsgevonden bij de aanmeldingen. Zo bleek dat in de onderzoekpopulatie korte duren waren oververtegenwoordigd en lange duren ondervertegenwoordigd. Deels had dat ook te maken met het feit dat afmeldingen met name plaatsvonden bij mensen uit het zittende bestand met langere duren. Ook naar nationaliteit was de groep niet representatief. Om hiervoor te corrigeren is 'matching' toegepast. Daartoe zijn er twee matching methoden gebruikt: 'nearest neighbour' en 'propensity score' matching. Het idee was om uit de referentiegroep een groep van 345 mensen te vinden die nagenoeg qua achtergrondkenmerken 100% overlap vertonen met de deelnemers. Op die manier kon worden gecorrigeerd voor de selectiviteit die was ontstaan bij de deelnemersgroep ten opzichte van de bijstandspopulatie. Dat bleek heel goed mogelijk, de kwaliteit van de matching was goed, en leidde ertoe dat het gevonden treatment effect naar schatting geen 15% (uitstroom per jaar) bedraagt maar 9 tot 13% na correctie voor selectie, maar nog altijd substantieel.

Niet-werk gerelateerde utikomstmaten

De meeste treatment effecten die we vonden bij de niet-werk gerelateerde uitkomstmaten bleken niet significant in vergelijking tussen de beide treatment groepen en de standaardgroep en we vonden ook geen significante verschillen tussen eigen regie en de standaardgroep. De meeste hypothesen die we formuleerden an het begin worden daarom niet bevestigd met uitzondering van de hypothesen ten aanzien van sociale participatie. Daar vinden we positieve en significante treatment effecten voor eigen regie en intensieve begeleiding vergeleken met de standaardgroep zowel in de procesevaluatie in de metingen bij de consulent als in de productevaluatie in de metingen bij de deelnemers. We vonden geen treatmenteffecten bij welbevinden, gezondheid, zelfredzaamheid, keuzemogelijkheden en vertrouwen. De regressieanalyses zowel van de procesevaluatie als van de productevaluatie gaven naar onze mening duidelijke aanwijzingen voor een verandering van de treatment in de standaardgroep. Bij de uitkomstmaat van welbevinden vonden we een positief significant treatment effect over de tijd bij de standaardgroep. Indien de standaard treatment niet was veranderd zou het vinden van een dergelijk effect onwaarschijnlijk zijn geweest. Met andere woorden, indien de standaard treatment inderdaad is veranderd, is geen zuivere vergelijking met de treatment groepen meer mogelijk. Dan zullen we een vergelijking moeten maken met een vergelijkingsgroep die niet beïnvloed is door het experiment. Op dit moment is zo'n alternatieve vergelijkingsgroep niet beschikbaar.

Effecten over tijd bij de andere uitkomstmaten

We vinden echter wel positieve en significante effecten over tijd in verband met welbevinden maar dan voor alle groepen. Of dit het gevolg is van het experiment, dus een treatment effect, is vooralsnog causaal niet vast te stellen zonder zo'n vergelijkingsgroep. Het is wel opvallend omdat er geen evidentie is dat welbevinden bij bijstandsgerechtigden zou zijn toegenomen in de afgelopen twee jaar, integendeel cijfers van het SCP (Wennekers e.a. 2019) suggereren dat voor mensen met een score op tevredenheid met het leven of geluk van 6 of lager (waaartoe ook de bijstandsgerechtigden behoren) de score de afgelopen jaren – dus na de economische crisisjaren - is verslechterd. Het is dus aannemelijk om te veronderstellen dat we hier te maken hebben met een 'Hawthorne' effect bij de standaardgroep maar zeker weten doen we het niet en vraagt nader onderzoek. Dat de treatment van de standaardgroep is veranderd zou ook afgeleid kunnen worden uit het feit dat we positieve en significante effecten in de tijd vinden bij de standaardgroep in geval van sociale participatie (vrijwilligerswerk, sociaal netwerk). Voor sociale participatie vinden we positieve treatment effecten van de beide treatments ten opzichte van de standaardgroep. Ook in de procesevaluatie vinden we in het interactiemodel met tijd significante treatment effecten voor de standaardgroep erop duidend dat de treatment is veranderd.

Met uitzondering van twee uitkomstmaten (subjectieve gezondheid en vertrouwen in de instituties) worden alle uitkomsten in de tijd gezien beter. We vinden geen causaal verband met de treatments met uitzondering van sociale participatie. Nader onderzoek moet uitwijzen of onderliggend deze stijging niet toch een treatment effect is dat niet zichtbaar wordt omdat de standaard treatment gedurende het experiment is gaan lijken op de andere treatments.

Verschillen tussen de treatments

Opvallend is dat we geen verschillen vinden tussen de treatments. Dit heeft verschillende oorzaken. De eerste is dat in het experiment met opzet is gekozen voor een ontheffings -of 'eigen regie' groep die

mensen leert zelfredzaam te worden. Dat betekent dus geen 'loslaten' zoals stricto senso onder het begrip 'eigen regie' zou kunnen worden verstaan. De training aan de consulenten was gericht op zelfredzaamheid waarbij het onderscheid tussen eigen regie en intensieve begeleiding minder aandacht heeft gekregen. De treatments zijn aan het begin van het experiment nader omschreven en op schrift gesteld. Daarin werd ook aangegeven hoeveel contactmomenten er gemiddeld over de tijd zullen zijn per jaar met de deelnemer. Het was duidelijk dat bij eigen regie het aantal contactmomenten minder zou zijn en dat de regie voor contact bij de deelnemer zou behoren te liggen. De afgesproken 'case load' was voor eigen regie ook hoger vastgesteld waardoor er minder tijd zou zijn dan bij intensieve begeleiding per deeelnemer. De ervaring leerde echter al snel dat de op papier gestelde uitgangspunten in de praktijk anders uitpakten. Reeds bij het begin van het experiment bleek dat ' leren zelfredzaamheid' in het begin juist veel contact vraagt met de deelnemer en dat de meeste deelnemers het juist prettig vonden in het experiment dat ze veel contact hadden met de consulent. De gangbare praktijk was dat mensen die reeds enkele jaren in de bijsjstand zaten, het zogenaamde 'granieten bestand', vaak heel weinig of geen enkel contact meer hadden met de consulent. Vaak werden mensen al ontheven van de arbeids – en/of sollicitatieverplichting. Deelname aan het experiment werd door de deelnemer dan ook soms gezien als een manier om meer ondersteuning te krijgen en meer contact. Door anderen werd eigen regie wel geinterpreteerd als 'loslaten' en handelingsvrijheid. Doordat vanwege de focus op zelfredzaamheid het onderscheid tussen beide treatments niet geheel duidelijk was en ook de contactintensiteit aan het begin niet verschilde was het voor consulenten niet van meet af aan duidelijk wat de verschillen waren tussen de treatments. Geleidelijk aan is het onderscheid tussen eigen regie en intensieve begeleiding in de werkwijze van de consulent helderder geworden en later bleek ook dat er minder contact met de deelnemer nodig was. Dit heeft tijd gevraagd. Het bespreken van dit probleem in het team en in de koppels heeft bijgedragen aan het meer onderscheidend inrichten van de treatments in het tweede jaar van het experiment. Nochtans vinden we dat niet terug in de uitkomsten noch bij de procesevaluatie noch bij de productevaluatie. Mogelijk heeft de snelle (voltijdse) uitstroom naar werk een rol gespeeld. Op dat moment is de 'echte' treatment al op een laag pitje gezet ook al wordt nog contact onderhouden met de deelnemer. Per 1 september 2018 na bijna een jaar was 12% volledig uitgestroomd naar werk maar na 1,5 jaar 23%. Daarnaast is er andere uitsroom vanwege verhuizing, overgang naar een andere regeling, herzien recht op uitkering, arbeidsongeschiktheid en dergelijke waardoor het aantal deelnemers snel afnam. Na 2 jaar was ruim 58% uitgestroomd. Dit heeft mogelijk effect gehad op de gevonden verschillen in uitkomsten tussen de beide treatments. Ook al vinden we geen significante verschillen, de consulenten zelf zijn meer geporteerd van de intensieve begeleiding als treatment. Zij zijn van mening dat de deelnemers hier meer baat bij hebben dan de eigen regie treatment. De pro-actieve rol die consulenten hebben bij intensieve begeleiding past wellicht beter bij de rol die van hen werd en wordt verwacht als professional, 'loslaten' hoorde en hoort daar minder bij. Mogelijk heeft het rolconflict dat consulenten ervoeren bij het uitvoeren van de treatments - zie het verslag van de focusgroepgesprekken- ook een rol gespeeld. Ook al werd van hen gevraagd om de treatments zo zuiver mogelijk uit te voeren en niet te veranderen, vanuit moreel en ethisch oogpunt vonden consulenten het lastig om de deelnemer iets te onthouden wat op dat moment door de consulent als professional nodig wordt geacht. Niet uitgesloten kan worden dat dit een effect heeft gehad op het zuiver uitvoeren van de treatments.

Tot besluit

Enkele van onze verwachtingen worden bevestigd anderen worden weerlegd op basis van deze cijfers. Vooralsnog worden de meeste hypotheses over de verschillen tussen de treatments en de standaardgroep weerlegd. Toch is er ook goed nieuws te melden. Het gevonden positieve treatment effect op sociale participatie is zeker niet onbelangrijk aangezien dit een van de belangrijkste doelen was die de gemeente zich bij aanvang stelde. De scores die consulenten geven aan de deelnemers op de participatieladder lieten een significante en sterke stijging zien voor de meerderheid van de deelnemers waardoor de afstand tot de arbeidsmarkt werd verkleind. Deze sociale stijging was voor de gemeente ook een belangrijk doel. Eerder constateerden we al dat er positieve en significante effecten zijn op volledige uitstroom naar werk maar ook bij deze uitkomstmaat is dus sprake van een duidelijk treatment effect. De lijntjes van de uitkomstmaten laten bijna allemaal een stijging zien ook bij de standaardgroep. Hoewel we geen treatment effecten vonden is dit opmerkelijk. Zo blijkt bij welbevinden en geluk dat de tevredenheid met leven van mensen die een 6 of lager scoren (waaronder bijstandsgerechtigden) afneemt terwijl in het experiment dit stijgt voor alle deelnemers. Ook bij sociaal vertrouwen zien we een stijging bij alle groepen. Nader onderzoek met een andere vergelijkingsgroep is dus nodig om tot definitieve uitspraken te komen over de uitkomsten op deze andere uitkomstmaten. De komende maanden willen we daartoe een poging wagen wanneer we de cijfers gaan koppelen aan de CBS-microdata bestanden.

Beleidsvragen

Niet alle beleidsvragen die aan het begin zijn geformuleerd kunnen worden beantwoord op dit moment. We laten ze hieronder kort de revue passeren:

- *De eerste beleidsvraag had betrekking op het effect van ontheffing*. Doordat we geen verschil vonden tussen de ontheffingsgroep en de intensieve begeleidingsgroep is het moeilijk om het effect van ontheffing en intensieve begeleiding afzonderlijk te meten. Voor mensen die in het zogenaamde 'granieten' bestand zitten is in feite al sprake van ontheffing. Er is nog maar sporadisch contact met de (werk)consulent en velen hebben ook ontheffing van de sollicitatie -en arbeidsplicht. In de ontheffingsgroep meten we dus niet zozeer ontheffing als wel de combinatie van ontheffing en van sturing op zelfredzaamheid. Alle uitkomstmaten wijzen in de goede richting maar zoals eerder vastgesteld is er geen significant effect in vergelijking met de standaardgroep, waarschijnlijk omdat die standaardtreatment zelf is veranderd en meer is gaan lijken op de andere treatments. Voor een deel hangt dat samen met verandering van de beleids -en uitvoeringscontext waarin het gedachtegoed van het experiment geleidelijk aan is geïmplementeerd in de praktijk van de uitvoering (zie verslag focusgroepgesprek). De uitstroom naar werk is de hoogste van alle drie groepen en twee keer zo hoog als bij de referentiegroep die volgens ons de werkelijke vergelijkingsgroep is.
- *De tweede beleidsvraag had betrekking op het effect op de arbeidsinschakeling wanneer mensen op een positieve vraaggerichte manier worden ondersteund en intensiever en op maat worden begeleid (met name intensieve begeleidingsgroep).* Deze onderzoeksvraag is hierboven beantwoord. Er is geen verschil in uitstroomkansen met de andere treatment groep of met de standaardgroep maar wel met de referentiegroep die volgens ons de juiste vergelijkingsgroep is.

De uitstroom is dan twee keer zo hoog en na correctie voor selectiviteit nog steeds zo'n 9 tot 13% hoger in beide treatment groepen.

- *Een derde beleidsvraag voor de gemeente was of mensen die deelnemen aan het experiment meer ruimte ontwikkelen om eigen initiatief te ontplooien, meer in brede zin gaan participeren, en wat het effect hiervan is op de arbeidsinschakeling (eigen regie, intensieve begeleiding).* Ook deze vraag kan bevestigend worden beantwoord. Hoewel ook hier evidentie is dat de standaardgroep ook is veranderd blijft er een significant efffect voor beide treatments op sociale participatie bestaan. De resultaten van de proces -en productevaluatie tonen dat beide aan. Er is opnieuw geen verschil gevonden tussen eigen regie en de intensieve begeleidingsgroep. Op basis van de procesevaluatie denken we dat het werkzame bestanddeel hier lijkt te zijn vertrouwen, aandacht, maatwerk en vraaggerichte ondersteuning.
- *Een belangrijke vraag voor de gemeente was wat de neveneffecten zijn van de ruimere handelingsvrijheid op maatschappelijke participatie, maar ook op zelfredzaamheid, welbevinden en gezondheid (eigen regie, intensieve begeleiding).* Op zelfredzaamheid zijn de effecten van de procesevaluatie positief ten opzichte van de standaardgroep. Bij de productevaluatie vinden we geen effect van de treatments op zelfredzaamheid, noch op gezondheid en welbevinden. We vinden wel positieve en significante treatment effecten op maatschappelijke of sociale participatie, en we vinden significante effecten voor welbevinden over de tijd maar dan voor allle treatment groepn tezamen. We vermoeden dat hier een treatment effect aanwezig is mar daarvoor hebben we een andere vergelijkingsgroep nodig.
- *In het experiment wordt onderzocht wat de uitkomsten over de tijd gezien zijn van meer handelingsvrijheid en een intensievere begeleiding op maat. Dat vraagt een andere manier van ondersteuning en andere kwalificaties van de consulenten. De gemeente wil graag weten hoe dit proces verloopt. Worden de treatments uitgevoerd zoals beoogd en in welke mate zijn ze onderscheidend? Wat werkt in de uitvoering en wat werkt niet? Welke condities, aspecten maken het verschil en wat betekent dit voor de implementatie van de uitkomsten (procesevaluatie).*
  Het bleek lastig om de treatments zoals beoogd uit te voeren. Het onderscheid tussen eigen regie en intensieve begeleiding was bij de start niet duidelijk voor de consulenten. Of dit een 'theory failure' of 'implementation failure' is, is moeilijk te zeggen. Wij denken beide. In de oratie van Blonk (2018) wordt opgemerkt dat er weinig literatuur is over het integratieve gedragsmodel wat betreft handvatten voor de professional. Wellicht heeft dat een rol gespeeld en waren we nog niet zo ver in het uitkristalliseren van dit model in de praktijk. Ook in de uitvoering was er geen methodiek om de treatments inhoud en vorm te geven. Het onderscheidende van de treatments was daardoor minder zeker in het eerste half jaar. We denken dat met name de verandering bij de standaard treatment duidt op een 'Hawthorne' effect waarvoor we niet konden corrgeren. De 'andere bejegening' maakt ook volgens de onderzoekers het verschil hoewel er geen meting daarvan bij de deelnemers heeft plaatsgevonden. De analyses van de vragenlijsten bij de consulenten en de focusgroep gesprekken geeft ons de indruk dat het opbouwen van een persoonlijke relatie, vertrouwen geven en aandacht, meer tijd voor het 'echte' gesprek, sturen op intrinsieke motivatie, een luisterend oor en loslaten sleutelbegrippen zijn voor het succes van de aanpak zoals die door de consulenten is ervaren

Literatuur

Abbring, J.H., G.J. van den Berg, P. Mullenders en J.C. van Ours (1996). Sancties in de ww; een werkend perspectief. In: Economisch Statistische Berichten, jg. 81, p. 750-753.

Abbring, J.H., G.J. van den Berg en J.C. van Ours (2005). The Effect of Unemployment Insurance.

Sanctions on the Transition Rate from Unemployment to Employment. In: Economic Journal, jg. 115, nr. 505, p. 602-630.

Atkinson, A. B. (1996) 'The case for a participation income', Political Quarterly, 67, 1, 67–70.

Bigotta, M., Bonoli, G., Fossati, F., Lalive, R., & Oesch, D. (2018). Helping social assistance recipients find jobs: Evidence from a controlled trial (ESPANet working paper, University of Lausanne, Switzerland)

Bohnet, I., Frey, B. S. and Huck, S. (2001) 'More order with less law: on contract enforcement, trust and crowding', The American Political Science Review, 95, 1,131–44.

Bond, G. R., Drake, R. E., & Luciano, A. (2015). Employment and educational outcomes in early intervention programs for early psychosis: A systematic review. Epidemiology and Psychiatric Sciences, 24(5), 446–457. https://doi.org/10.1017/S2045796014000419

Blonk, R.W.B., M.W. van, Twuijver, H.A. van de Ven en A.M. Hazelzet (2015). Quickscan wetenschappelijke literatuur gemeentelijke uitvoeringspraktijk. Leiden: tno.

Blonk, Roland (2018), We zijn nog maar net begonnen, Openbare rede prof. dr. R. Blonk, Tilburg University op 6 april 2018.

Card, D., Kluve J., en A. Weber (2010), Active labour market policy evaluations: a meta-analysis, *The Economic Journal,* 120 (November), F452–F477.

Card, D., Kluve, J., & Weber, A. (2015). What works? A meta analysis of recent active labor market program evaluations (No. w21431). National Bureau of Economic Research.

CBS (2017), De Geluk meter, CBS Statline, Voorburg/Heerlen.

Cuelenaere, B., E. van den Brink en J. Leenen (2019). Ervaringen gemeenten met Participatiewet. Rapportage derde meting. Tilburg: Centerdata.

Deaton, A., Nancy Cartwright (2017. Understanding and misunderstanding randomized controlled trials, Social Science & Medicine, Vol. 210(2018):2-21(https://doi.org/10.1016/j.socscimed.2017.12.005)

Deci, E. L. and Ryan, R. M. (1985) Intrinsic Motivation and Self-Determination in Human Behavior, New York: Plenum.

Delos, S. en J. Oude Alink (2016), Experiment regelluwe bijstandsverlening, Brief aan portefeuillehouders, Werk en Inkomen, Gemeente Oss.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decade of progress. Psychological Bulletin, 125(2), 276–302. https://doi.org/10.1037/0033-2909.125.2.

Dohmen, T., A. Falk, D. Huffman and U. Sunde (2008), Representative trust and reciprocity: Prevalence and determinants, *Economic Inquiry*, 46 (1), 84–90.

Fehr, E. & Gächter, S. (2000). Fairness and Retaliation: the economics of reciprocity. *The Journal of Economic Perspectives, 14*(3), 159-181.

Forget, E. L. (2011). The town with no poverty: the health effects of a Canadian guaranteed annual income field experiment. *Canadian Public Policy*, *37*(3), 283-305.

Fehr, E. and Schmidt, K. M. (2003) 'Theories of fairness and reciprocity: evidence and economic

applications', Advances in Economics and Econometrics, Econometric Society, Eighth World Congress, 1, 208–57.

Fishbein, M., & Ajzèn, I. (2010). Predicting and changing behavior: The reasoned action approach. New York: Psychology Press.

Frey, B. S. and Jegen, R. (2001) 'Motivation crowding theory', Journal of Economic Surveys, 15, 5, 589–611.

Goodin, R. E., Headey, B., Muffels, R. and Dirven, H.-J. (1999) The Real Worlds of Welfare Capitalism, Cambridge: Cambridge University Press.

Groot, N. de en B. van der Klaauw (2019). The effects of reducing the entitlement period to unemployment insurance benefits. In: Labour Economics, jg. 57, p. 195-208.

Groot, L. and van der Veen, R. (2000) 'Clues and leads in the policy debate on basic income in the Netherlands', in R. van der Veen and L. Groot (eds.), Basic Income on the Agenda: Policy Objectives

Groot, L., Muffels, R. J., & Verlaat, T. (2019). Welfare states' social investment strategies and the emergence of Dutch experiments on a minimum income guarantee, Social Policy & Society (2019) 18:2, 277–287, Cambridge University Press

Headey, B., Muffels, R.J.A., & Wagner, G. (2010). Long-running German panel survey shows that personal and economic choices, not just genes, matter for happiness. Proceedings of the National Academy of Sciences of the United States of America (PNAS), 107(42), 17922-17926

Hemerijck, A. (2013) Changing Welfare States, Oxford: Oxford University Press.

Hemerijck, A. (ed.) (2017) The Uses of Social Investment, Oxford: Oxford University Press.

Hoff Stella Hoff Bart van Hulst Jean Marie Wildeboer Schut Benedikt Goderis (2019), Armoedekaart 2019 (http://digitaal.scp.nl/armoedeinkaart2016), SCP-Den Haag.

Koning, P. (2012a). Beter een stok dan een wortel. In: J. Ultermark, A.J. Gielen en M. Ham (red.), Wat werkt nu werkelijk? Politiek en praktijk van sociale interventies (p. 105-120). Amsterdam: Van Gennep.

Kluve, J., (2010). The effectiveness of European active labor market programs. Labour Econ. 17, 904–918. doi:10.1016/j.labeco.2010.02.004.

Kok, L., D. Hollanders en J.P. Hop (2006). Kosten en baten van reïntegratie. Amsterdam: SEO Economisch Onderzoek.

Koning, J. de, A. Gelderblom, K. Zandvliet en L. van den Boom (2005). Effectiviteit van reïntegratie. De stand van zaken. Literatuuronderzoek. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid.

Koning, J. de, P. de Hek, L. Mallee, F. Rosing en M. Groenewoud (2014). Uitkomsten en ervaringen experimenten netto-effectiviteit re-integratie. Rotterdam/Amsterdam: seor/Regioplan Beleidsonderzoek/Epsilon Research.

Kremer, M., J. van der Meer, en M. Ham (2017). Van werk naar welbevinden – werkt de zachte hand in de bijstand? Sociale Vraagstukken (https://www.socialevraagstukken.nl/van-werk-naarwelbevinden- werkt-de-zachte-hand-in-de-bijstand/)

Mani, A., Mullainathan, S., Shafir, E. and Zhao, J. (2013) 'Poverty Impedes Cognitive Function', Science, August, 976–80.

Marshall, T., Goldberg, R. W., Braude, L., Dougherty, R. H., Daniels, A. S., Sushmita Shoma Ghose, E. D., et al. (2014). Supported employment: Assessing the evidence. Psychiatric Services, 65(1), 16–27. https://doi.org/10.1176/appi.ps.201300262

Morel, N., Palier, B. and Palme, J. (2012) Towards a Social Investment Welfare State? Ideas, Policies and Challenges, Bristol: The Policy Press.

Muffels, R. and B. Headey (2013). Capabilities and choices: Do they make Sense for understanding objective and subjective well-being? An empirical test of Sen's capability framework on German and British panel data." *Social Indicators Research,* **110**(3): 1159-1185.

Muffels, R. & S. Vos (2018), Tussenrapportage No. 1, Vertrouwensexperiment Oss, Tilburg University/Gemeente Oss, p. 24

Muffels, R., K. Stam & S. van Wanrooij (2019), Tweede Voortgangsverslag Vertrouwensexperiment Oss: No. 2, 1 maart 2019, Tilburg University/ReflecT-Tranzo, p. 1-30.

Muffels, R. & J. Van der Klink (2017), Offerte Vertrouwensexperimenten Apeldoorn, Oss, Renkum, Epe en Geldrop-Mierlo: Onderzoekdesign en kostenraming onderzoek, Tilburg University, 10 mei 2017.

Mullainathan, S. and Shafir, E. (2013) Scarcity – The True Cost of Not Having Enough. London: Penguin Books.

Nazar, Amir (2010), Wat werkt? Wat weten we over effectiviteit?, Ministerie van Sociale Zaken en Werkgelegenheid (http:/www.onderzoekwerkeninkomen.nl/rapporten/xbdiusf2).

Ryan, R. M., Kuhl, J. and Deci, E. L. (1997) 'Nature and autonomy: an organizational view of social and neurobiological aspects of self-regulation in behavior and development', Development and Psychopathology, 9, 4, 701–28.

Sen, A. K. (2004) 'Capabilities, lists, and public reason: continuing the conversation', Feminist Economics,10, 3, 77–80.

Sen, A. K. (1999). Development as Freedom, New York, Knopf.

Soentken, M., van Hooren, F. and Rice, D. (2017) 'The impact of social investment reforms on income and activation in the Netherlands', in A. Hemerijck (ed.), The Uses of Social Investment, OUP Oxford, Kindle Edition, 475.

Van Echtelt, Patricia e.a. (2019), Eindevaluatie van de Participatiewet, SCP.

Thaler, R. H. and Sunstein, C. R. (2008) Nudge: ImprovingDecisions about Health,Wealth, and Happiness, New Haven, Conn.: Yale University Press.

Van der Klink JJ, Bultmann U, Burdorf A, Schaufeli WB, Zijlstra FR, Abma FI, Brouwer S, van der Wilt GJ.Sustainable employability--definition, conceptualization, and implications: A perspective based on the capability approach. Scand J Work Environ Health 2016; 42:71-9.

Van Ryn, M., & Vinokur, A. D. (1992). How did it work? An examination of the mechanisms through which an intervention for the unemployed promoted job-search behavior. *American Journal of Community Psychology*, *20*(5), 577-597.

Veenhoven, R. (1996). Developments in satisfaction-research. Social Indicators Research, 37(1), 821–46. https://doi.org/10.1007/BF00300268

Widerquist, K. (forthcoming) The Devil's in the Caveats: A Critical Analysis of Basic Income Experiments for Researchers, Policymakers, and Citizens, Palgrave Macmillan.

Van Hooft E.A.J. (2016). Self-Regulatory Perspectives in the Theory of Planned Job Search Behavior: Deliberate and Automatic Self-Regulation Strategies to Facilitate Job seeking. In:

Wennekers, A. J. Boelhouwer, C. van Campen, J. Kullberg (2019), De sociale staat van Nederland, SCP Den Haag.

Bijlage A. Resultaten regressieanalyses cliëntvragenlijsten

Tabel A1.1. Niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS of logit regressie model (armoede)

Concepten	Eigen regie	Intensieve begeleiding	Meting 2	Meting 3
Welbevinden	0	0	.43***	.29**
Subjectieve gezondheid	0	-2.0***	0	0
Keuzemogelijkheden	0	0	0	0
Perceptie sociale participatie	0	0	0	0
Sociale participatie	.72*	1.49***	.95***	.63***
Zelfredzaamheid	0	0	.32*	-.52**
Vertrouwen instituties	0	0	-.69***	-.72***
Sociaal vertrouwen	0	-1.97***	0	0.73***
Inkomensarmoede	0	0	0	0
Deprivatiearmoede (3 uit 5)	0	0	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 2: Voor de testen van de effecten van eigen regie en intensieve begeleiding was de standaardgroep de referentiecategorie waartegen werd getest. Bij de effecten van meting 2 en 3 was de nulmeting de referentiecategorie.

Noot 3: In dit model waren de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen waren treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom, nummer consulent en een aantal controlevariabelen (leeftijd, geslacht, opleiding, duur uitkering, leefvorm, en huisvesting). Deze uitkomsten zijn de test voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.2.1 t/m 3.2.12. Bij een 0 in de tabel, kon er geen significant effect gevonden worden.

Bron: Deelnemers vragenlijsten in de drie metingen Vertrouwensexperiment Oss, Tilburg University, 2020.

Tabel A1.2: Regressieuitkomsten interactiemodel treatment en meting productevaluatie (significante treatment effecten per meting)

Uitkomstmaten	Eigen regie			Intensieve begeleiding			Standaardgroep		
	Nul meting	M2	M3	Nul-meting	M2	M3	Nul-meting	M2	M3
Welbevinden	0	0	0	0	0	-.88**	0	.68**	.62**
Subjectieve gezondheid	-1.63**	0	0	0	0	0	0	0	0
Keuzemogelijkheden	0	0	0	0	0	0	0	0	0
Perceptie sociale participatie	0	0	0	0	0	0	0	0	0
Sociale participatie	1.39**	0	0	0	0	0	0	0	.82**
Zelfredzaamheid	0	0	0	0	0	0	0	0	0
Vertrouwen instituties	0	0	0	0	0	0	0	0	0
Sociaal vertrouwen	0	0	0	0	0	0	0	0	0
Inkomensarmoede	0	0	0	0	0	0	0	0	0
Deprivatiearmoede (3 uit 5)	0	0	0	0	0	0	0	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 2: voor de effecten van de standaardgroep in meting 2 en 3 is de standaardgroep in de nulmeting de referentiecategorie. Voor de effecten van eigen regie in de nulmeting is de referentie de standaardgroep in de nulmeting en voor metingen 2 en 3 geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van het effect van de eigen regiegroep in de nulmeting waar de standaardgroep de referentie is. Hetzelfde geldt voor het effect van intensieve begeleiding.

Noot 3: In dit model zijn de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen zijn treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en consulentnummer. Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.3.1 t/m 3.3.10. Bij een 0 in de tabel, is er geen significant effect gevonden.

Bijlage B. Resultaten regressieanalyses procesevaluatievragenlijsten

Tabel B1.1: niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS regressiemodel per treatment groep over de tijd en per meting

Concepten	Eigen regie	Intensieve begeleiding	Meting 2	Meting 3
Motivatie eigen treatment	2***	3,5***	0	0
Motivatie hoogst bereikbare doel	1,34**	2,75**	1,74***	2,19***
Realistische eisen	0,93**	0	0,76**	1,29***
Zelf in actie komen	1,02**	0	1,36***	1,87***
Zelf doelen stellen	0	0	0	1,51***
Verwacht effect treatment	2,4***	2,34**	0	0
Sociale participatie	1,75***	0	0,35*	0,91***
Participatieladder	0	0	0,78***	0,82***
Leereffect deelnemer	1,84***	1,42**	0	0
Leereffect consulent	3,92***	2,61***	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 2: Voor het testen van de effecten van eigen regie en intensieve begeleiding is de standaardgroep de referentiecategorie. Bij de effecten van meting 2 en 3 is de standaardgroep in de nulmeting de referentiecategorie.

Noot 3: In dit model zijn de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen zijn treatment en meting. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en consulentnummer. Deze uitkomsten geven de test weer voor de verschillen van de buitenste rij en kolomtotalen van tabellen 3.3.1 t/m 3.3.10. Bij een 0 in de tabel, is er geen significant effect gevonden.

Bron: Vragenlijsten procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.

Tabel B1.2: niet-gestandaardiseerde regressie coëfficiënten (*b*) geschat op basis van OLS interactie model (treatment effecten per meting)

	Standaardgroep		**Eigen regie**			**Intensieve begeleiding**		
Concepten	M2	M 3	Nulmeting	M2	M3	Nulmeting	M2	M3
Motivatie eigen treatment	-1,01**	0	1,58**	1,09*	0	3,08**	1,09*	0
Motivatie hoogst bereikbare doel	1,05*	1,92***	1,06*	0	0	2,45**	1,02*	0
Realistische eisen	0,67*	1,54***	1,10**	0	0	0	0	0
Zelf in actie komen	1,08**	2,04***	1,51**	0	-1,57**	0	0	0
Zelf doelen stellen	0	1,43**	1,15**	0	-1,11*	0	0	0
Verwacht effect treatment	0	0	2,78***	0	-1,15*	2,64**	0	0
Sociale participatie	0	1,24**	1,71***	0,98*	0	0	0,89*	0
Participatieladder	0,66**	1,11***	0	0	0	0	0	0
Leereffect deelnemer	0	0	2,54***	-1,17**	-1,52**	1,97**	0	0
Leereffect consulent	0	0	3,8***	0	0	2,49***	0	0

Noot 1: ***p<0,01; **p<0,05; *p<0,10

Noot 2: voor de effecten van de standaardgroep in meting 2 en 3 is de standaardgroep in de nulmeting de referentiecategorie. Voor de effecten van eigen regie in de nulmeting is de referentie de standaardgroep in de nulmeting en voor de vervolgmetingen geeft de parameter het effect weer van de groep in die meting dat bovenop het effect komt van de eigen regiegroep in de nulmeting waar de standaardgroep de referentie is. Hetzelfde geldt voor het effect van intensieve begeleiding.

Noot 3: In dit model waren de concepten in de eerste kolom de afhankelijke variabelen Y. De belangrijkste onafhankelijke variabelen waren treatment en meting en de interactie tussen deze twee. Verder is gecontroleerd voor zittend bestand/nieuwe instroom en consulentnummer. Deze uitkomsten zijn de test voor de verschillen van de tussenliggende cellen van tabellen 3.3.1 t/m 3.3.10. Bij een 0 in de tabel, kon er geen significant effect gevonden worden.

Bron: Vragenlijsten nulmeting en vervolgmetingen procesevaluatie Vertrouwensexperiment Oss, Tilburg University, 2020.