Keuzes in toezicht

De effecten van toezicht op compliance in beeld?

Concept masterscriptie

Master Bestuurskunde

Erasmus Universiteit Rotterdam

Door: Marijke Azier

Studentnummer 509431

Datum: 10 december 2020

Aantal woorden: 19.015

Scriptiebegeleider: Henk Klaassen

Tweede lezer: Vincent Homburg

Inhoud

Lijst met tabellen en figuren 2
Voorwoord 3
Samenvatting 4
1. Inleiding 6
1.1. Inleiding 6
1.2. Aanleiding 7
1.3. Probleemstelling van het onderzoek 9
1.4. Relevantie van het onderzoek 10
1.5. Leeswijzer 11
2. Compliance en de effectiviteit van toezicht 12
2.1. Compliance 12
2.2. Toezicht en toezichtinstrumenten 15
2.3. Effectiviteit van toezicht op compliance en hoe breng je dit in beeld 17
2.4. Conceptueel model 22
3. Onderzoeksmethodiek 24
3.1. Operationalisatie 24
3.2. Onderzoekstrategie 25
3.3. Betrouwbaarheid en validiteit 28
4. Onderzoek 30
4.1. Analyse (beleids)documenten 30
4.2. Analyse interviews respondenten 35
4.2.1. Doelgroepen 35
4.2.2. Interventies en de gehanteerde handhavingsmix 36
4.2.3. Effecten 45
4.3. Evaluatie 48
4.4. Samenvatting onderzoek 53
5. Conclusie, discussie en aanbevelingen 56
5.1. Conclusies 56
5.2. Discussie 57
5.3. Aanbevelingen 59
Literatuur- en bronnenlijst 60
Bijlage 1. Vragenlijst semi gestructureerde interviews 62

Lijst met tabellen en figuren

Tabel 1: Operationalisatie kernconcepten ……24/25
Tabel 2: Operationalisatie overige concepten ……..25
Tabel 3: Overzicht kengetallen compliancefactoren 2016-2019 ……..….32
Tabel 4: Overzicht interventies en ordening ……..37
Tabel 5. Overzicht beantwoording werkhypotheses door respondenten ……49
Figuur 1. Beslistabel effectmeting ……..55

Voorwoord

Beste lezer,

Deze scriptie is geschreven in het kader van de afronding van mijn studie Bestuurskunde aan de Erasmus Universiteit Rotterdam. Als medewerker bij de Belastingdienst heb ik de keuzes in toezicht als onderwerp voor deze scriptie gekozen vanuit grote persoonlijke interesse in en betrokkenheid bij het toezicht wat door de Belastingdienst MKB uitgevoerd wordt. De wijze waarop de Belastingdienst met een veelheid aan taken en doelstellingen zijn werk in een steeds complexere omgeving moet doen en keuzes moet maken is intrigerend. De ontwikkelingen in de automatisering geven steeds meer mogelijkheden om data te analyseren. De toenemende aandacht voor privacy vanwege de AVG (Algemene verordening gegevensbescherming) brengt echter ook kaders aan om decentraal over data te beschikken. Het is een uitdaging om theoretische inzichten en centrale keuzes in verbinding te brengen met de dagelijkse uitvoering en ervaringen op de werkvloer, waar steeds minder ruimte komt voor maatwerk.

In de periode waarin ik deze scriptie heb geschreven en het (voor)onderzoek heb gedaan (zomer 2019 – december 2020) zijn er alweer de nodige ontwikkelingen geweest, met name in de aanpassing van de begrotingsdoelstellingen. Dit leest u verderop in deze scriptie. Het continu keuzes maken om met de beschikbare middelen en instrumenten een optimale bijdrage te leveren aan de maatschappelijke opdracht is en blijft een uitdaging. Zo ook het schrijven van deze scriptie.

Voor de uitvoering van dit onderzoek ben ik vele collega's dank verschuldigd. De collega's die met mij gespard hebben over de haalbaarheid en formulering van de onderzoeksvraag en aanpak, mij van informatie hebben voorzien en mee hebben gewerkt aan de interviews en (delen van) de scriptie hebben meegelezen, zijn van groot belang geweest om de scriptie tot een goed einde te brengen. Ook de scriptiekring, en met name mijn scriptiebegeleider Henk Klaassen, ben ik grote dank verschuldigd, mede omdat zij mij iedere keer wezen op gebruik van interne terminologie die voor hen soms onnavolgbaar was. Daarnaast heeft Henk mij op zeer betrokken wijze bij ieder contact weer bruikbare feedback gegeven.

Tot slot ben ik dank verschuldigd aan mijn gezin, die mij ruimte heeft gegeven om deze studie te volgen. Ik was lang niet altijd beschikbaar om hen te helpen bij hun schoolwerk of alle huishoudelijke taken op de verwachte momenten op mij te nemen.

Ik hoop dat deze scriptie ook weer zal leiden tot een kleine stap in de richting van waar wij als Belastingdienst heen willen qua ontwikkeling in toezicht.

Ik wens u veel leesplezier!

Marijke Azier, Rhoon, 10 december 2020

Samenvatting

Deze scriptie beschrijft een onderzoek naar de effectiviteit van (toezicht)interventies door de Belastingdienst MKB op compliance (naleving van wet- en regelgeving) bij burgers en bedrijven en de wijze waarop dit effect in beeld is gebracht. Naar aanleiding van het onderzoek worden aanbevelingen gedaan om de effectiviteit van de (mix van) toezichtinstrumenten te vergroten.

Uit het onderzoek is gebleken dat beïnvloeding van het nalevingsgedrag niet altijd de reden blijkt te zijn geweest om voor een bepaald toezichtinstrument of andere interventie te kiezen. In een aantal situaties was het doel van toezicht het uitvoeren van een proces om deze goed te laten verlopen en bepaalde 'aantallen' te halen (output). Als gedragsverandering ten aanzien van compliance wel expliciet was beoogd, dan was er in de meeste gevallen sprake van gemeten resultaat op korte termijn (impact) of in een enkel geval van effect op langere termijn (effectiviteit). Het is belangrijk gebleken om op voorhand bewust te zijn van het doel wat wordt beoogd met de inzet van interventies en welke (effect)meting hierbij hoort. In het beslisschema op pagina 55 is dit inzichtelijk gemaakt en blijkt ook aan welke voorwaarden moet worden voldaan om effectmeting te kunnen realiseren. Los van het feit dat effectmeting zeer lastig is, is effectmeting niet noodzakelijk als met de inzet van een toezichtinstrument of andere interventie geen bewust effect op compliance wordt beoogd. De gekozen aanpak qua inzet van interventies zal in dat geval waarschijnlijk niet aangepast worden.

Daar waar effectmeting heeft plaatsgevonden was er in alle gevallen sprake van een *bewuste* aanpak om gedragsverandering te realiseren, was er *vooraf* nagedacht over het beoogde effect en hoe dit te gemeten zou kunnen worden en er werd gebruik gemaakt van *extracomptabele* vastleggingen naast data uit bestaande databases. Daarnaast was er sprake van inzet van een ruime mix aan interventies met duidelijk accent op *preventie en een proactieve houding*: voorlichting, communicatie in de vorm van bellen, brieven sturen en samenwerking met externe partijen zoals fiscaal dienstverleners en brancheorganisaties. Naast de samenwerking met externe partijen werd in die gevallen ook gebruik gemaakt van de ervaringen van belastingplichten en/of belastingschuldigen om interne beelden aan te vullen en de verdere aanpak qua keuzes voor interventies op te baseren.

De conclusie van dit onderzoek is dat er weinig ruimte zit bij de verplichte inzet van toezichtinstrumenten die vanuit de uitvoeringsopdracht in de aangifteregeling of het invorderingsproces worden ingezet, maar dat de toegevoegde waarde en het vergroten

van de effectiviteit van de mix aan toezichtinstrumenten gezocht moet worden in de (aanvullende) instrumenten 'aan de voorkant' van het toezichtinstrumentarium. Daar waar een bewuste inzet van (een mix van) interventies heeft plaatsgevonden, vaak vanuit een integrale aanpak, heeft dit geleid tot meetbare effecten ten aanzien van het nalevingsgedrag.

Het onderzoek heeft geleid tot de volgende aanbevelingen om de effectiviteit van interventies beter in beeld te kunnen brengen en op basis daarvan de keuzes voor de inzet van de (toezicht)instrumenten te kunnen verbeteren en zo het effect op compliance te kunnen vergroten:

- Richt digitale systemen zodanig in dat deze geschikt zijn om als basis te dienen voor de koppeling van interventies aan het effect op compliance, zonder dat extracomptabele vastlegging nodig is;
- Maak de keuze om een afdeling onderzoek zodanig te positioneren en vorm te geven dat deze de kwaliteit van de effectmeting, monitoring en evaluaties optimaal kan ondersteunen. Hieronder valt ook een traject van vooraf bewuste keuzes maken;
- Zet meer in op klanttevredenheidsonderzoek om waardevolle feedback op te halen bij belastingplichtigen en/of belastingschuldigen en hiermee de uitvoering van het toezicht beter vorm te kunnen geven ten aanzien van de juiste inzet van de interventies en de menselijke maat hierin.

1. Inleiding

1.1. Inleiding

In 2013 bracht de Wetenschappelijke Raad voor het Regeringsbeleid (WRR) een rapport uit met de naam 'Toezien op publieke belangen'. In dit rapport wordt onder andere ingegaan op de vraag hoe de maatschappelijke meerwaarde van het toezicht door de overheid kan worden vergroot. Het rapport beschrijft de paradox tussen de wens van de maatschappij om publiek geld efficiënt te besteden en tegelijk de behoefte aan méér en strenger toezicht, zeker nadat zich incidenten voor hebben gedaan. De WRR (WRR, 2013) roept toezichthouders op om, vanuit hun unieke kennispositie, een bijdrage te leveren aan het onderkennen van kansen en risico's en deze terug te koppelen naar het bestuur, de politiek en de maatschappij om daarmee bij te dragen aan beleidsontwikkeling. Hierbij wordt vooral opgeroepen om in een breder perspectief naar dit toezicht te kijken. Zo is niet alleen het toezicht door de overheid aan de orde, maar neemt ook het maatschappelijk middenveld en de burger zelf een aandeel in dit toezicht. Hiermee nemen ook de verschijningsvormen van toezicht toe. Wat gelijk blijft is het maatschappelijk doel, het publieke belang, wat de uitvoering van toezicht geacht wordt te dienen. De WRR (WRR, 2013) pleit dan ook voor een verruimd perspectief op het rijks toezicht. Hiermee wordt bedoeld dat niet alleen naar de kosten en opbrengsten van het toezicht moet worden gekeken, maar vooral ook hoe het toezicht bijdraagt aan het publieke belang en wat derhalve de maatschappelijke opbrengst is. Hierbij past een actieve houding ten aanzien van actuele ontwikkelingen en oog voor het krachtenveld waarbinnen het toezicht plaats vindt, de samenwerking met andere toezichthouders en gedragsmechanismen die het gevolg zijn van regelgeving en het uitgevoerde toezicht hierop, zo stelt de WRR (WRR, 2013).

Eén van de belangrijkste toezichthouders in Nederland is de Belastingdienst. In de missie van de Belastingdienst staat dat hij bijdraagt aan een financieel gezond Nederland. Het publieke belang zit in de financiering van een veelheid aan overheidstaken. Nu, met de Coronacrisis in 2020, blijkt maar weer hoe belangrijk deze rol van de overheid is. Dit geldt zowel voor een evenwichtige bijdrage aan de schatkist als een keuze in de uitgaven, ook op Europees niveau.

De taken van de Belastingdienst bestaan uit het heffen (opleggen van aanslagen) en innen van belastingen, premies en accijnzen en het uitkeren van toeslagen. Dit gebeurt op basis van beschikbare data en ingediende aanslagen en aanvragen (Ministerie van Financiën, 2019). Op basis van diezelfde data vindt ook intensiever toezicht op wet- en regelgeving plaats door diverse vormen van controle, naast bijvoorbeeld het geven van

voorlichting of het geven van zekerheid door het voeren van vooroverleg met bedrijven. Bij de uitvoering van zijn taken zet de Belastingdienst, naast de administratieve processen, in op het actief uitvoeren van toezicht. De vraag is welke verschijningsvormen het toezicht door de Belastingdienst heeft en hoe de maatschappelijke meerwaarde van dit toezicht in beeld wordt gebracht. Daarnaast is er de vraag hoe deze meerwaarde kan worden vergroot.

1.2. Aanleiding

De maatschappij (burgers en bedrijven) moet erop kunnen vertrouwen dat de ter beschikking gestelde middelen en menskracht op doeltreffende wijze door de Belastingdienst worden ingezet om (het toezicht op) de uitvoering van de belastingwetten vorm te geven. Deze uitvoering bestaat uit administratieve (registrerende) taken, toezichtstaken en opsporingstaken. Dit onderzoek richt zich op de toezichtstaak van de Belastingdienst, onderdeel Midden- en Kleinbedrijf (MKB).
Voor de onderzoekomgeving van het MKB is gekozen omdat hier, vanwege de omvang en met name de diversiteit van dit segment, veel keuzes kunnen en moeten worden gemaakt ten aanzien van de inzet van een breed assortiment van toezichtinstrumenten en het verwachte rendement hiervan.

Met ruim twee miljoen middelgrote en kleine bedrijven moet er een evenwicht worden gevonden in de aard en de omvang van dit toezicht. De complexiteit hiervan maakt een onderzoek interessant. Bij de inrichting van het toezicht wordt ook de samenwerking gezocht met de fiscaal dienstverleners, brancheorganisaties, dienstverlenende partijen (bv. softwareleveranciers) en andere toezichthouders binnen de overheid. Dit sluit aan bij de oproep van de WRR (WRR, 2013) om ook het krachtenveld rondom de regelgeving, het effect van toezicht én de samenwerking met andere toezichthouders invulling te geven. Met het onderzoeken en in beeld brengen van de wijze waarop keuzes in het toezicht door de Belastingdienst worden gemaakt, en welke effecten hiermee worden beoogd, kan een antwoord worden gevonden op de vraag over de maatschappelijke bijdrage van dit toezicht en hoe dit het beste in beeld kan worden gebracht. Dit kan een verdere basis vormen voor toekomstige beleidsontwikkeling binnen de Belastingdienst en bijdragen aan een advies ten aanzien van politieke keuzes qua (sub)doelstellingen.

De taak van de Belastingdienst MKB is het toezien op een eerlijke en zorgvuldige heffing en inning van belastingen van middelgrote en kleine bedrijven en de daaraan verbonden natuurlijke personen. De Belastingdienst MKB heeft voor zijn taak een veelheid aan toezichtinstrumenten tot zijn beschikking. Deze variëren van voorlichting en vooroverleg aan de ‘voorkant’ van het toezicht tot fraudebestrijding achteraf. De complexiteit van het

toezicht zit in de keuze voor het inzetten van (een mix van) díe instrumenten bij de diverse doelgroepen, zodanig dat deze optimaal bijdragen aan de doelstellingen van de dienst: bijdragen aan een financieel gezond Nederland, zoals dit in de missie in het jaarplan 2020 van de Belastingdienst is weergegeven (Ministerie van Financiën, 2019).

De strategie op basis waarvan de Belastingdienst zijn doelstellingen wil bereiken, het toezicht inricht en keuzes maakt, is gebaseerd op de compliance-doelstellingen[1]. Onder compliance wordt het gedrag verstaan waarbij burgers en bedrijven zonder (dwingende en kostbare) acties van Belastingdienst structureel uit zichzelf (fiscale) regels naleven. De keuze van de inzet van de diverse toezichtinstrumenten is bedoeld om dit gedrag te stimuleren en compliance-risico's tegen te gaan. Het belang is hierin gelegen dat bij een hoge compliance het mogelijk is om met een kleinere hoeveelheid mensen en middelen het toezicht in te richten en een 'slankere' Belastingdienst in te richten. Het is daarom ook belangrijk om het toezicht zodanig vorm te geven dat er optimaal gebruik wordt gemaakt van de 'goedkopere' instrumenten zoals voorlichting, ondersteunende tools op de Belastingdienstsite en telefonisch contact en zo min mogelijk gebruik hoeft te worden gemaakt van de 'duurdere' toezichtinstrumenten zoals een boekenonderzoek of een (langdurige) juridische procedure.

De vraag voor dit onderzoek is hoe de keuze en de verhouding van inzet op deze toezichtinstrumenten, de zogenaamde toezichtmix, tot stand komt en hoe er bij deze keuze rekening gehouden wordt met de impact die dit heeft op de beoogde compliance. De onderliggende vraag is of er voldoende kennis is over het effect van de inzet van de diverse toezichtinstrumenten op compliance of dat de keuze op basis van andere criteria plaatsvindt.

Zowel de commissie Stevens (Stevens, Pheiffer, Van den Broek, Keijzer, & Van der Hel-van Dijk, 2012) als de WRR (WRR, 2013) zijn er niet van overtuigd dat de keuzes die de Belastingdienst ten aanzien van zijn inzet maakt onderbouwd kunnen worden vanuit de effectiviteit van dit toezicht op de beoogde compliance. De commissie Stevens geeft specifiek ten aanzien van het zogenaamde horizontale toezicht[2] aan dat er nog "onvoldoende ondersteuning is door informatiesystemen die het mogelijk maken het juiste toezichtinstrument te kiezen en de effecten van de nalevingsprocessen te meten". Stevens (Stevens et.al, 2012) geeft hiermee aan dat er onvoldoende registratie

[1] Gebaseerd op de OESO-factoren zijn door de Belastingdienst de volgende vier aspecten van compliance benoemd: 1. Registratie van belastingplicht, 2. Tijdigheid van aangifte, 3. Juistheid en volledigheid van aangiften en 4. Tijdigheid en volledigheid van betalingen.

[2] Onder horizontaal toezicht wordt de samenwerking met andere partijen verstaan, in casu de fiscale dienstverleners, waarmee ook convenanten worden afgesloten over de wijze van aangifte doen en het aangepaste toezicht wat de Belastingdienst daar tegenover stelt.

plaatsvindt van de toegepaste toezichtinstrumenten op klantniveau, zodat er geen kwantitatieve analyse mogelijk is om het effect van de individuele toezichtinstrumenten op compliance in beeld te brengen en zodoende de meest effectieve keuze hierin te kunnen maken. Vanuit de WRR (WRR, 2013) wordt dit belang onderschreven om zo de vraag te kunnen beantwoorden of de inzet van horizontaal toezicht een efficiënte keuze is en daarmee het publiek belang dient.

De vraag is vervolgens welke informatie beschikbaar zou moeten zijn om een onderbouwde keuze te kunnen maken voor het 'juiste' toezichtinstrument. Door de uitvoering van dit onderzoek wordt in beeld gebracht wat er in de wetenschappelijke literatuur bekend is over criteria die van belang zijn voor effectmeting, in het bijzonder ten aanzien van compliance, en welke criteria binnen de Belastingdienst nu onderkend en gehanteerd worden bij de huidige werkwijze en keuzes.

1.3. Probleemstelling van het onderzoek

De Belastingdienst MKB wil sturen op een doeltreffende inzet van de ter beschikking gestelde middelen en menskracht door te kiezen voor de juiste mix van in te zetten toezichtinstrumenten bij het toezicht op burgers en bedrijven. Hiervoor is het nodig inzicht te hebben in de criteria die bijdragen aan compliance en deze te verbinden met de keuze voor (de mix van) de diverse toezichtinstrumenten.

Doelstelling

Het doel van dit onderzoek is het in beeld brengen van de effectiviteit van de diverse toezichtinstrumenten op de compliance van burgers en bedrijven bij de Belastingdienst MKB en het doen van aanbevelingen om de keuze voor de toezichtmix te verbeteren *door* gebruik te maken van wetenschappelijke theorieën over de criteria van effectief toezicht en deze te vergelijken met de uitvoeringspraktijk.

Vraagstelling

<u>Centrale vraag</u>
Welke toezichtinstrumenten zet de Belastingdienst MKB in ter bevordering van de compliance van burgers en bedrijven, hoe wordt deze inzet gemotiveerd en wat zijn de mogelijkheden om de effectiviteit van de mix van toezichtinstrumenten te vergroten?

Deelvragen

Deelvraag 1. Welke theoretische inzichten bestaan er over de relatie tussen toezicht en compliance?

Deelvraag 2. Hoe is het toezichtinstrumentarium van de Belastingdienst MKB vormgegeven?

Deelvraag 3: Welke criteria worden gehanteerd bij de inzet van de diverse toezichtinstrumenten?

Deelvraag 4. Welke huidige inzichten bestaan er bij de Belastingdienst inzake de effectiviteit van de ingezette toezichtinstrumenten?

Deelvraag 5. Welke mogelijkheden zijn er om de effectiviteit van (de mix van) het toezichtinstrumentarium te vergroten?

1.4. Relevantie van het onderzoek

Wetenschappelijke relevantie

Veel publieke organen worstelen met de vraag hoe effectief hun dienstverlening en/of toezicht is en hoe hier maatschappelijke verantwoording over kan worden afgelegd. Het vinden van de juiste constellatie van factoren die deze effectiviteit beïnvloeden is lastig. De wijze waarop hun toegevoegde waarde in beeld kan worden gebracht is zo mogelijk nog moeilijker. Hier is nog beperkt wetenschappelijk onderzoek naar gedaan. Dit onderzoek heeft als doel om het proces van de (bewuste) inzet van toezichtinstrumenten te koppelen aan de (beoogde) effecten en aanbevelingen te doen hoe de effectiviteit (beter) in beeld kan worden gebracht. De uitkomsten en aanbevelingen van dit onderzoek kunnen een bijdrage leveren aan de theoretische inzichten hierover.

Maatschappelijke relevantie

Het is belangrijk dat de Belastingdienst met de beschikbare mensen en middelen zijn toezicht zo doeltreffend mogelijk inricht. Het belang van de effectiviteit van het toezicht (compliance) is groot, omdat de inzet van de juiste combinatie van toezichtinstrumenten kan leiden tot een efficiëntere, en dus kleinere en goedkopere, organisatie om de financiële middelen van de overheid te borgen. Met het geven van een goed inzicht van deze effectiviteit kan de Belastingdienst de maatschappelijke meerwaarde van zijn toezicht ook beter onderbouwen. Daarnaast is het van belang om het toezicht dusdanig flexibel in te richten dat er snel en actief ingespeeld kan worden op wijzigingen in het krachtenveld binnen de samenleving, zoals de WRR (WRR, 2013) ook voorstaat. Hierbij past een organisatie die gericht is op compliance, waardoor een financieel gezond Nederland met een zo effectief en efficiënt mogelijke uitvoerings- en toezichtorganisatie

bereikt en behouden kan worden. Op deze wijze kan de maatschappelijke opbrengst van de Belastingdienst worden geborgd.

1.5. Leeswijzer

In hoofdstuk twee worden de kernbegrippen 'compliance' en 'toezicht' en hun onderlinge samenhang beschreven en de wijze waarop dit in beeld kan worden gebracht. Het hoofdstuk wordt afgesloten met het conceptueel model waarop dit onderzoek gebaseerd is. In hoofdstuk drie wordt aangegeven hoe het onderzoek is vormgegeven en hoe de kernbegrippen worden geoperationaliseerd. Het feitelijke onderzoek betreft een combinatie van interviews en documentonderzoek. In hoofdstuk vier worden de bevindingen van het onderzoek gepresenteerd. Deze bestaan, naast de bevindingen van het documentonderzoek, uit de beschrijving van de diverse ingezette toezichtinstrumenten, de motivatie van het gebruik hiervan binnen de diverse projecten en ketens én de tot nu toe verwachte en in beeld gebrachte effectiviteit op de compliance hiervan. In hoofdstuk vijf vindt de evaluatie van het onderzoek plaats en in hoofdstuk zes worden de conclusies, de discussie en de aanbevelingen beschreven die uit het onderzoek naar voren zijn gekomen.

2. Compliance en de effectiviteit van toezicht

In dit hoofdstuk wordt nader ingegaan op de relevante theoretische inzichten ten aanzien van de benoemde concepten, in casu compliance, toezicht en de effectiviteit van toezicht op compliance. Ook wordt aandacht besteed aan de wijze waarop de effectiviteit in beeld kan worden gebracht. De verwachtingen van het onderzoek worden als werkhypotheses weergegeven.

2.1. Compliance

Inleiding

Met de term compliance wordt in de wetenschappelijke literatuur gedrag bedoeld dat conform regels of normen is (Goslinga, Siglé, & Van der Hel, 2016). Goslinga et al. (2016, voetnoot 11 p. 10) geeft daarbij aan dat het hierbij niet belangrijk is of dit gedrag uit eigen beweging plaats vindt of een doelbewuste keuze betreft, maar wel gaat om de *bereidheid* tot het nakomen van verplichtingen. De feitelijke mate van naleving van regels of normen is afhankelijk van een veelheid van redenen, die ook onderling op elkaar inwerken (Etienne, 2011). Etienne (2011) baseert zijn compliance theorie op de goal framing theorie (GFT) van Siegwart Lindenberg en hanteert een onderscheid tussen materiële, emotionele en normatieve doelen, die ook onderling interacteren. Ditzelfde verband ziet Etienne (2011) ook bij compliance, waarbij verschillende doelen, voorkeuren en persoonlijke percepties met elkaar strijden om voorrang. De keuzes die gemaakt worden hebben zowel te maken met verwachte "opbrengst" als bijvoorbeeld sociale goedkeuring en vinden plaats binnen een omgeving van beperkte rationaliteit en beperkte cognitieve mogelijkheden. Aanvullend op de doelen theorie benoemt Etienne (2011) factoren voor non-compliance zoals onbekwaamheid, onwetendheid en het verkeerd begrijpen van de regels als oorzaak.

Ten aanzien van motivatie om bepaalde dingen te doen (of te laten) is door Deci en Ryan de self-determination theory (SDT) ontwikkeld. Hierin wordt beschreven dat motivatie de behoefte aan competentie, autonomie en verbondenheid betreft (Deci & Ryan, 2000). Zij geven aan dat de intrinsieke motivatie mede samenhangt met de bevrediging van deze behoeften. Deze intrinsieke motivatie kan gevoed worden door geïnterneerde extrinsieke motivatie (gevoed door eerdere ervaringen). Bij extrinsieke motivatie gaat het om een stimulans van iets of iemand anders, waarbij je kunt denken aan een beloning of straf. Ten aanzien van competentie is het, voor de motivatie om een bepaald doel te bereiken, ook belangrijk dat het te realiseren doel haalbaar is, dat er geloof is in de eigen bekwaamheid om een bepaalde taak uit te voeren en met een positief resultaat af te ronden. Deze factor wordt in de literatuur ook wel aangeduid met self-efficacy,

zelfeffectiviteit (Locke & Latham, 2002). Een belangrijk aspect van bekwaamheid heeft betrekking op de kennis waarover beschikt kan worden. Dit aspect wordt in de volgende paragraaf uitgewerkt.

Kennis

Uit de, voor dit onderzoek, bestudeerde literatuur is gebleken dat kennis over wet- en regelgeving een belangrijke factor is die bijdraagt aan compliance (Goel, Ravindra, Singh, & Sharma, 2014), (Sellens, et al., 2018). De betreffende onderzoeken lopen uiteen van naleving van het rookverbod in India, aanbevelingen van het Q-koorts vaccin, tot het naleven van hygiëne-instructies in de zorg. Wat deze onderzoeken verbindt is de conclusie dat het verstrekken van informatie, het geven van educatie en communicatie-activiteiten zichtbaar bijdragen aan kennis en bewustwording en daarmee bijdragen aan de compliance ten aanzien van de betreffende regelgeving. Op zich is dit geen verrassende uitkomst als we ervan uitgaan dat de gemiddelde burger/medewerker zich wil conformeren aan geldende wet- en regelgeving. Voorwaarde hiervoor is dat je de wet- en regelgeving dan wel moet kennen en op de juiste manier moet kunnen interpreteren. Dit sluit ook aan op de factoren van non-compliance (onwetendheid en verkeerd begrijpen van regels) die Etienne (2011) benoemt. Door het geven van gerichte informatie aan een beoogde doelgroep wordt het kennisniveau van deze groep, en de bewustheid hierover, verhoogd. Het gaat hierbij om algemene kennis over een specifiek onderwerp waarvan gebleken is dat de compliance in naleving hierop bij deze groep achterblijft. Dit betekent dat verwacht mag worden dat gerichte informatievoorziening bijdraagt aan de compliance van de beoogde doelgroep.

Dit leidt tot de eerste werkhypothese voor dit onderzoek:

WH1: Naarmate de voorlichting specifieker op de doelgroep wordt gericht, neemt de compliance toe.

Feedback als onderdeel van bewustwording

Een onderzoek van Reich, et al. over hygiëne op de ic-afdeling van een ziekenhuis laat zien dat educatie en het helpen herinneren aan de hygiëneregels pas tot een daadwerkelijke gedragsverandering leiden als dat wordt gecombineerd met feedback over het gedrag (Reich, et al., 2015). In dit onderzoek werd de feedback over het percentage compliance negen maanden lang iedere maand breed gecommuniceerd. De bereikte compliance bleek ook na afronding van het onderzoek een blijvend karakter te hebben gekregen. Communicatie in de vorm van feedback over compliance droeg dus bij

aan de beoogde hygiënedoelstelling van het ziekenhuis. Feedback kan in die zin gezien worden als een onderdeel van bewustwording zoals hiervoor is beschreven.
Feedback kan op veel verschillende manieren gegeven worden. In een onderzoek van Feys (Feys, Libbrecht, Anseel, & Lievens, 2008) werd onderzocht wat het effect van de kwaliteit van de relatie tussen beoordelaar en beoordeelde (interactie) én de ervaren procedurele rechtvaardigheid was op de gegeven feedback. De geformuleerde werkwerkhypotheses over de veronderstelde positieve relatie tussen het interactie-effect en de rechtvaardigheidspercepties ten aanzien van feedback konden bevestigend beantwoord worden. Indien zowel de relatiekwaliteit als de rechtvaardigheid van de inhoudelijke feedback positief beoordeeld werden had dit het grootste positieve effect op het gewenste gedrag. Dit onderzoek vond plaats ten aanzien van prestatiebeoordeling van een medewerker door diens leidinggevende. De vraag is in hoeverre deze variabelen ook een factor kunnen zijn in de relatie tussen een (overheids)organisatie en de maatschappij. Tenslotte gaat het om gedrag, in dit geval niet van een medewerker, maar van burgers en bedrijven. De (fysieke) afstand is weliswaar groter en de context is anders, echter zoals aangegeven gaat het in beide gevallen om gedrag van mensen en de wijze waarop dit beïnvloed kan worden.

Dit leidt tot een tweede werkhypothese in dit onderzoek:

WH2: Communicatie in de vorm van gerechtvaardigde feedback draagt bij aan een hogere compliance.

(Non) compliance – individuele afweging
In een publicatie over ontduiking van inkomstenbelasting (Allingham & Sandmo, 1972) worden een aantal factoren benoemd die bijdragen aan de keuze voor non-compliance. In hun theoretische analyse over belastingontduiking benoemen Allingham en Sandmo als factoren de inschatting van de kans op controle (pakkans) en de mate waarin de belastingbetaler bereid is om dit risico te lopen en de bijbehorende sanctie te ondergaan (alsnog belasting betalen plus een boete of celstraf).
Ook Goslinga et al. (2016) geven aan dat de pakkans een belangrijk aspect is bij (non)compliance. Zij onderkennen bij de beïnvloeding van compliance drie achterliggende oorzaken: economische (financieel), sociaalpsychologische (rechtvaardigheid) en economisch-psychologisch (onzekerheid en risicobereidheid). Onder economische oorzaken worden door Goslinga et al (2016) de voordelen benoemd die worden beoogd bij non-compliance ten opzichte van de risico's van de pakkans met eventuele sancties. De gepercipieerde pakkans (de kosten van correctie en boete) is daarmee een factor, die individueel ingeschat en op waarde gewogen wordt. Het individu maakt daarmee een bewuste keuze voor (non)compliant gedrag op basis van het

economisch nadeel wat hij daarmee verwacht in relatie tot het risico wat hij bereid is te lopen, het economisch-psychologische aspect. De onzekerheid en risicobereidheid worden mede ingegeven door eigen ideeën en veronderstellingen.
Tot slot wordt nog verwezen naar Voermans (Voermans, 2005). Als één van de aspecten van een effectief controlesysteem (gebaseerd op de checklist van De Moor-van Vugt, als onderdeel van een preadvies voor de Vereniging voor milieurecht uit 2003) benoemt hij de intensiteit van het toezicht. Deze moet volgens Voermans voldoende hoog zijn om er een preventieve werking van uit te laten gaan.

De gepercipieerde pakkans is daarmee een factor die medebepalend is voor de effectiviteit van de verschillende interventies die vanuit het toezicht kunnen worden ingezet. De eigen inschatting van deze pakkans kan te maken hebben met de vorm en intensiteit van het toezicht (een vragenbrief of diepgaand onderzoek), maar in ieder geval met de inschatting van de kans om met enige vorm van het toezicht te maken te krijgen. Dit betreft een kwantitatief aspect van het toezicht. Voor dit onderzoek wordt intensiteit vertaald met het percentage onder toezicht gestelden dat 'geraakt' worden met enige vorm van toezicht, het zogenaamde rakingspercentage. Op basis van bovenstaande onderzoeken kan gesteld wordt dat de intensiteit van het toezicht invloed heeft op de wijze waarop de gevolgen van (non)compliance worden ingeschat.

Dit leidt tot de volgende werkhypothese:

WH3: Een hoge intensiteit van toezicht (rakingspercentage) van onder toezicht gestelden door toezichtinstrumenten verhoogt de compliance.

2.2. Toezicht en toezichtinstrumenten

Toezicht
In de, voor dit onderzoek, beoordeelde literatuur is er een behoorlijke overeenstemming over de betekenis van het begrip toezicht. Ten aanzien van toezicht wordt over het algemeen de volgende driedeling gehanteerd: het verzamelen van data/gegevens, het geven van een oordeel en het zo nodig interveniëren/handhaven, waaronder het opleggen van sancties (Winter & De Ridder, 2010) (Samadi, 2016) (Voermans, 2005). Winter en de Ridder (2010) voegen hier nog twee specifieke kenmerken aan toe. Volgens hen is er bij toezicht sprake van de (wettelijke) bevoegdheid van een actor om in te grijpen én sprake van de afwezigheid van hiërarchie. Zij beschrijven dat het bij de afwezigheid van hiërarchie gaat om relaties waarbij er sprake is van onder toezicht gestelden die (primair) zelf verantwoordelijk zijn voor hun keuzes en handelingen. Bij de bevoegdheid van een actor gaat het om de legitimiteit om op te treden. Zo heeft de

politie bepaalde bevoegdheden, zoals het geweldsmonopolie, die maakt dat zij en alléén zij op een bepaalde manier mogen optreden. Denk hierbij aan het binnentreden van een woning of het met geweld aanhouden van een verdachte. De burger is zelf verantwoordelijk voor zijn gedrag, maar kan gecorrigeerd worden door een toezichthouder waarmee geen hiërarchische verhouding bestaat.

Bovenstaande bevoegdheid is essentieel als het gaat om het opleggen van sancties om zo de gewenste gedragsverandering af te dwingen. Dit mandaat wordt ook expliciet door Voermans (2005) benoemd. Daarnaast onderscheidt Voermans (2005, p. 71 voetnoot 10) drie vormen van toezicht: nalevingstoezicht, bestuurlijk toezicht en mededingingstoezicht. Alhoewel het onderzoek van Voermans zich richt op de naleving en handhaving van Europese wetgeving, zijn ook hier elementen uit te benoemen die breder toepasbaar zijn. In het kader van dit onderzoek is het nalevingstoezicht (naleving van wet- en regelgeving) aan de orde. Onder nalevingstoezicht wordt door Voermans (2005) het toezicht verstaan dat zich richt op de zelfstandige en rechtstreekse controle op de naleving van de regels, de zogenaamde eerstelijnscontrole. Winter en de Ridder (2010) geven hierbij overigens nog aan dat het maatschappelijk effect (vergroten belastinginkomsten) en het nalevingseffect (normconformiteit) bij belastingwetgeving samenvallen.

Toezichtinstrumenten en doelgroepen

In hun rapport geeft de WRR (WRR, 2013) aan dat veel toezichthouders hun toezichtinstrumenten koppelen aan de soort onder toezicht gestelden. Hierbij worden drie groepen onderkend. Als eerste groep worden de 'koplopers' of normgetrouwen benoemd, waarbij het risico op niet-naleving laag wordt inschat. Deze groep leent zich voor het maken van afspraken in de vorm van convenanten (het zogenaamde horizontale toezicht) of het maken van afspraken over het gebruik van gecertificeerde kwaliteitssystemen. Als tweede groep wordt de welwillende midden categorie benoemd, waarbij de naleving kan worden verhoogd door een combinatie van dienstverlening (advies en voorlichting) en handhaving (correctie). De derde groep benoemt het WRR (WRR, 2013) als de groep waarbij de naleving van de norm achterblijft en die vooral via handhaving (correctie en sancties) aangepakt moet worden.

De WRR (2013:30) geeft als doel van toezicht, met verwijzing naar Mertens (2006:266), aan "het wettelijk gelegitimeerd beïnvloeden van bepaalde aspecten van het gedrag van een rechtspersoon ten einde een maatschappelijk gewenst handelingspatroon beter te bereiken". Hiermee komen we dus bij de relatie tussen de *bewuste* inzet van toezicht en het *beoogde* effect, wat in de volgende paragraaf behandeld wordt.

De aard van het toezicht en de toezichtinstrumenten veranderen overigens mee met de ontwikkeling in de maatschappij. Zo valt in verschillende rapporten te lezen dat de ontwikkeling van 'een oordeel over' is doorontwikkeld naar een aanpak van 'met elkaar in gesprek gaan'. Een bredere integrale aanpak, meer maatwerk en meer samenwerking worden genoemd als instrument om tijdig bij te sturen richting het gewenst effect. In een artikel over het toezicht op sociale verzekeringen lezen we zelfs dat "De belangrijkste toezichtinstrumenten, namelijk overleg en overreding, zijn niet in de wet vastgelegd" (Sureveen & Rijsbergen, 1997). Als meest vergaande wettelijke toezichtinstrumenten door het College van Toezicht Sociale Verzekeringsbanken worden de repressieve instrumenten van het geven van een aanwijzing en het openbaar maken van toezichtbevindingen genoemd. Bovenstaande betekent dat toezichthouders goed moeten (blijven) monitoren wat in een bepaalde situatie de beste wijze van toezicht houden is en dat dit door meerdere aspecten beïnvloed wordt.

WH4: Hoe specifieker de keuze van (de mix van) toezichtinstrumenten op de diverse doelgroepen is afgestemd, hoe groter de effectiviteit hiervan op (het behoud van) compliance bij de betreffende doelgroepen.

2.3. Effectiviteit van toezicht op compliance en hoe breng je dit in beeld

Randvoorwaarden

In 2010 hebben Winter en de Ridder (Winter & De Ridder, 2010) een literatuurstudie gedaan naar de beschikbare kennis over de relatie tussen toezicht en de effecten daarvan. Zij omschrijven 'effectiviteit van toezicht' als de mate waarin de uitvoering van het toezicht bijdraagt aan de verwezenlijking van de beoogde doeleinden van dat toezicht. Dit kan gaan om een beoogde gedragsverandering, maar ook over de continuering van gewenst gedrag. In het kader van dit onderzoek gaat het om het verhogen van de compliance. Toezicht wordt daarbij gezien als middel en niet als zelfstandig doel. Winter & de Ridder (2010) constateren dat de keuzes voor bepaalde toezichtactiviteiten weliswaar bewust gemaakt kunnen worden, maar dat de uitdrukkelijke vastlegging van wat men wil bereiken volgens de Algemene Rekenkamer ontbreekt. Hierdoor is er te weinig kennis over het effect van de ingezette toezichtactiviteiten en daarmee op het effect hiervan op de naleving van de achterliggende doelstelling. Het gaat hierbij dan uitdrukkelijk om het leggen van het *verband* tussen het optreden van de toezichthouder en het nalevingsniveau. Het kan daarbij gaan om het verband tussen één interventie en het nalevingsniveau, maar ook om een specifieke combinatie van interventies door de toezichthouder. Winter & de Ridder (2010) geven daarbij aan dat het voor het aantonen van een dergelijk verband

helpt als er gebruik gemaakt kan worden van bestaande (betrouwbare) databases. Het éénduidig vastleggen van interventies is daarbij een belangrijke randvoorwaarde om een verondersteld verband met naleving te kunnen toetsen.

In de inleiding is aangegeven welk belang de WRR (2013) hecht aan de bijdrage van de toezichthouder aan de beleidsontwikkeling. Het is daarom van belang dat de toezichthouder zelf ook goed zicht heeft op de uitvoering van zijn toezicht en met onderbouwde beelden aan kan geven welke bijdrage vernieuwingen leveren aan de toename van de maatschappelijke waarde van dit toezicht. Goslinga et al. (2016) beschrijven in dat kader een aantal voorbeelden van de door de Belastingdienst ingezette toezichtinstrumenten die de compliance door de burger zouden verhogen. Zo geeft hij aan dat aan het geven van voorlichting bij de jaarlijkse aangiftecampagne de aanname ten grondslag ligt dat dit het kennisniveau en het tijdig doen van aangifte verhoogt. Het toetsen van deze aanname vindt volgens Goslinga et al (2016) echter niet structureel plaats. Een ander voorbeeld wat hij noemt is het dienstverlenend bezoek van de Belastingdienst aan een recent gestarte ondernemer. De veronderstelling was dat deze bezoeken zouden bijdragen aan het beter en sneller op de rails hebben van belastingzaken. Alhoewel de meerderheid van de bezochte ondernemers het bezoek waardeerden en nuttig vonden, bleek dit volgens Goslinga et al (2016) geen waarneembaar effect te hebben op de diverse aspecten van compliance. Wel bleken deze ondernemers volgens hem meer rekening te houden met (onverwachte) toekomstige (belasting)uitgaven dan ondernemers in de controlegroep en financieel dus voorzichtiger te zijn voor het geval zij een financiële tegenslag te verwerken kregen.

Uit bovenstaande voorbeelden blijkt dat aannames van ingezette toezichtinstrumenten in de praktijk niet altijd blijken te werken. De hiervoor genoemd randvoorwaarde dat er een structurele vastlegging nodig is om het (gebrek aan) effect doorlopend te kunnen monitoren en bij te kunnen sturen blijkt wederom van belang. Aan het einde van deze paragraaf wordt hierop teruggekomen.

Gedragsverandering door interventie

In het rapport van de WRR (2013: 62) wordt aangegeven dat gedragswetenschappers steeds meer onderzoek doen naar de factoren die een rol spelen bij nalevingsgedrag. Zij baseren zich op Huisman en Beukelman (2007) die intrinsieke motivatie, sociale normen en de gepercipieerde pakkans als elementen benoemen. Ook de relatie tot het toezicht wordt benoemd. Vooral bij extrinsieke motivatie is bijvoorbeeld het effect van een sanctie groter dan die bij de welwillende midden categorie. Bij deze categorie kan een sanctie zelfs averechts werken. Dit betekent dat er in het toezicht ruimte moet zijn

om de interventie af te kunnen stemmen op de doelgroep om optimaal effect van het toezicht te mogen verwachten.

WH5: Hoe beter handhaving (correctie van gedrag) is afgestemd op een bepaalde doelgroep, hoe groter het effect op compliance.

Een onderdeel van de handhaving is het corrigeren van het gedrag en het (eventueel) opleggen van een sanctie. Voermans benoemt daarbij het belang van de beoordeling of die middel geschikt en relevant is in de betreffende context (Voermans, 2005). Is dit niet het geval, dan kan gesteld worden dat dit niet een doeltreffende reactie op de overtreding is. Daarbij is ook het aspect van proportionaliteit van belang. Hiermee wordt de relatie tussen de ernst van het feit en de zwaarte van de sanctie bedoeld.
De individueel ervaren rechtvaardigheid wordt daarmee een factor in de uiteindelijke keuze voor wel of geen compliant gedrag. Zo zal een redelijke boete kunnen bijdrage aan compliant gedrag, terwijl een onredelijk boete juist weerstand kan oproepen (Goslinga, Siglé, & Van der Hel, 2016).

WH6: Hoe specifieker een sanctie wordt afgestemd op de betreffende overtreding van wet- en regelgeving, hoe beter het effect op compliance.

Output en outcome (impact en effect)
Bij het meten van effectiviteit gaat het niet altijd om de output (producten) van het toezicht, maar meestal om de outcome, het gegenereerde effect van het toezicht (Bordewijk & Klaassen, 2011). Van belang hierbij is dus of de output zelf het doel is of dat de output het middel is om een ander doel (outcome) te bereiken. In de planning- en control cyclus worden deze output en outcome beschreven en vertaald in indicatoren of kengetallen (Bordewijk & Klaassen, 2011). Kengetallen worden hierbij beschreven als absolute getallen of een verhoudingsgetal. De eenheid waarin wordt gemeten is afhankelijk van het beleidsterrein dat in beeld moet worden gebracht. Om een dergelijk kengetal betekenis te geven moet dit worden gerelateerd aan een vooraf bepaald referentiepunt. In dit onderzoek gaat het om de eerder benoemde compliance factoren van tijdige aangifte, tijdige betaling en het juist en volledig doen van aangifte.

Desmidt en Heene (Desmidt & Heene, 2018, p. 426) maken bij de outcome onderscheid tussen impact (korte termijn) en effect (lange termijn). Zij baseren zich daarbij op Bouckaert et al. (2003) en Van Dooren én Bouckaert en Halligan (2010), die stellen dat impact het effect op korte termijn is. Bij het effect op lange termijn is sprake van geïnterneerd gedrag, waardoor minder/ander toezicht aan de orde kan zijn.

Een mooi voorbeeld hiervan is een toezichtproject in een intensive care unit (ICU) in de USA ten aanzien van hand hygiëne. Het inregelen van een interventie door tijdelijk extra schriftelijk feedback te geven op het gedrag van de ICU-collega's verhoogde de naleving van hygiëneregels binnen een half jaar van 65% naar bijna 92% (impact van het toezicht). In de twee jaar daarna bleef deze score rond de 90% liggen, zonder dat daar extra toezicht voor nodig was (Reich, et al., 2015). Hier was dus sprake van een blijvend effect op lange termijn doordat de regels geïnterneerd waren. De extrinsieke motivatie is daarmee omgezet in intrinsieke motivatie om zich aan de regels te conformeren.
Om dit over langere termijn te kunnen monitoren is het dus wederom van belang om structurele vastleggingen te hebben.

In hoofdstuk 4 wordt bij de documentanalyse verder ingegaan op de wijze waarop benoemde tussendoelen en compliance met elkaar samenhangen en welke ontwikkelingen dit beleidsmatig heeft doorgemaakt ten aanzien van sturing op output, impact en effect.

Doeltreffendheid (effectiviteit)
Bij het onderzoeken van de effectiviteit van het toezicht is het belangrijk om nog even stil te staan bij de betekenis van het begrip effectiviteit. Hiervoor zijn verschillende invalshoeken te onderkennen vanuit technisch en/of economisch perspectief (Klaassen, Maks, & Anderson, 2010). Voor dit onderzoek gaat het erom een zo groot mogelijke bijdrage aan compliance te bewerkstelligen. Dit kan betekenen dat een relatief 'duur' toezichtinstrument ingezet kan worden als dit een grotere bijdrage levert aan de beoogde compliance dan een goedkoper toezichtinstrument. Aan de andere kant kan ook gekozen worden voor een relatief 'goedkoop' toezichtinstrument wanneer een minder groot effect op compliance verwacht wordt, maar het toezicht om andere redenen wel uitgevoerd moet worden. De termen 'duur' en 'goedkoop' hebben betrekking op de hoeveelheid in te zetten tijd en middelen. Met dezelfde mensen en middelen kan bij een 'goedkoop' toezichtinstrument een groter deel van de onder toezicht gestelden bereikt worden. Dit eerder genoemde 'rakingspercentage' kan leiden tot een groter effect, in de zin van een breder bereik. Ook is er de keuze om bij een 'goedkoop' instrument hetzelfde aantal onder toezicht gestelden te bereiken met de inzet van minder mensen en middelen.
De gepercipieerde pakkans kan onderdeel uitmaken van de afweging welke toezichtinstrument in welke omvang ingezet kan worden. Zo kan zichtbaarheid bij een groter deel van de onder toezicht gestelden wellicht evenveel of meer impact hebben dan diepgaand onderzoek bij enkelen. Aan de andere kant kan één diepgaand onderzoek wat in de publiciteit komt wellicht ook zorgen voor een groot uitstralingseffect. Juiste deze

afweging betreft de essentie van de onderbouwing van de gekozen mix aan toezichtinstrumenten in een bepaalde setting vanuit het oogpunt van doeltreffendheid. Bij effectiviteit gaat het om de relatie tussen outcome en doelen. Hoe groter de bijdrage van de outcome aan de doelstelling, hoe effectiever dit proces verlopen is.
Om de effectiviteit van een proces te meten is het nodig om hier van tevoren normen voor vast te stellen en deze gewenste normen te vergelijken met de feitelijke effectiviteit van het proces.

Het begrip doeltreffendheid is ook gedefinieerd ten aanzien van de rijksbegroting. We vinden daar terug dat onder doeltreffendheid wordt verstaan: de mate waarin de (beleids)doelstelling dankzij de inzet van de onderzochte (beleids)instrumenten wordt gerealiseerd (Rijksoverheid, 2020). Voor het aantonen van doeltreffendheid is het dus noodzakelijk dat er een verband aanwezig is (causaliteit) tussen de prestatie (output) en het effect (outcome). Ook moet vaststaan dat het effect niet door andere oorzaken is veroorzaakt. Het meten van dit effect is eigenlijk pas goed mogelijk als de situatie kan worden gemeten mét en zonder de onderzochte inspanningen.

Op basis van bovenstaande uiteenzetting kan geconcludeerd worden dat er een bepaalde afweging moet zijn om te komen tot een optimale verhouding tussen de aard en omvang van het aantal ingezette toezichtmomenten.

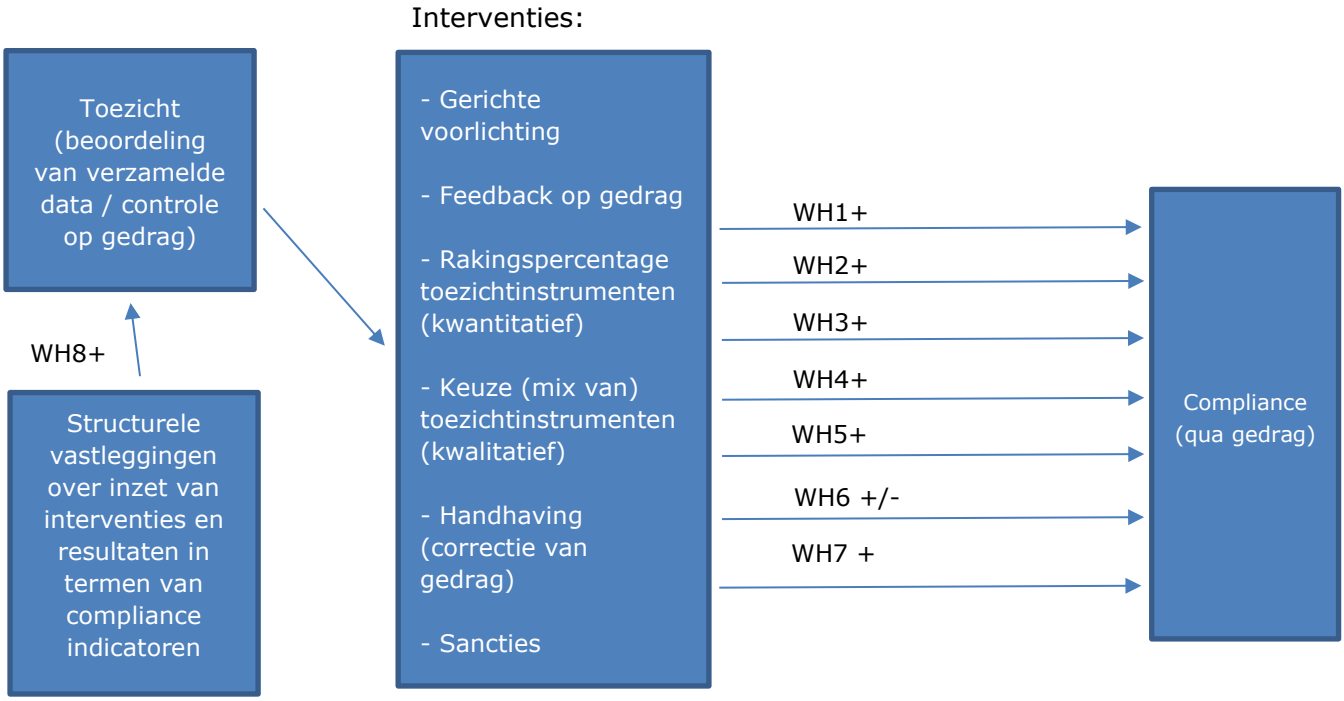
WH7: Hoe beter de afweging tussen de aard en omvang van de inzet van toezichtinstrumenten hoe groter de impact op compliance is.

Structurele vastlegging van data
Om het effect van toezicht in beeld te brengen is eerder als randvoorwaarde aangegeven dat het nodig is om structureel data te verzamelen en vast te leggen op een manier waarbij de impact/het effect beoordeeld kan worden. Winter en de Ridder benadrukken, naar aanleiding van een onderzoek naar verkeersveiligheid, het belang van goed geregistreerde cijfers uit *bestaande* databases (Winter & De Ridder, 2010). Dit verhoogt de betrouwbaarheid van de effectmeting. Ook Klaassen et al. (2010) bevestigen het belang van structurele vastlegging van kengetallen, in dit geval ook losgekoppeld van prestatiemeting, om te voorkomen dat strategisch of tactisch gedrag de uitkomsten zou kunnen beïnvloeden en er geen zuivere meting van effecten (outcome) plaats kan vinden.

WH8: Structurele vastlegging van kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten op het moment dat de gebeurtenis zich voordoet verhoogt de betrouwbaarheid van de effectmeting op compliance.

2.4. Conceptueel model

Op basis van het theoretisch kader is onderstaand conceptueel model opgesteld.



Toelichting op het conceptueel model

In de theorie wordt 'toezicht' in brede zin verdeeld over een drietal componenten: voorlichting, beoordeling en handhaving. Handhaving kan alleen plaats vinden als er voorafgaand een beoordeling heeft plaatsgevonden die deze handhaving rechtvaardigt. Naast het actief inzetten van een gekozen mix aan instrumenten van toezicht en handhaving kunnen ook voorlichting (preventief) en het geven van feedback (reactief) ingezet worden om vooraf of tijdens het toezichtproces een rol te spelen in het stimuleren van de gewenste compliance. Er wordt een positief verband verondersteld tussen de inzet van voorlichting, feedback op gedrag, rakingspercentage, keuze van de inzet van toezichtinstrumenten en handhaving (WH1, 2, 3, 4 en 5) op compliance.

Daarnaast kan het inzetten van sancties, mits gepast en proportioneel, een extra stimulans zijn om compliance te bevorderen (WH6+). Is aan deze beide voorwaarden niet voldaan, dan kan een sanctie een negatief effect op compliance hebben (WH6-).

Met WH7 wordt een positief verband verondersteld tussen de afweging van het soort in te zetten toezichtinstrument en de omvang van die inzet ten aanzien van de impact op compliance.

Om de effecten te kunnen beoordelen is het van belang zowel de input van mensen en middelen, de ingezette toezichtinstrumenten, de handhavingsactiviteiten én de compliance indicatoren structureel vast te leggen (WH8). Indien dit kwalitatief goed en gestructureerd gebeurt is dit bovendien een goed uitgangspunt voor beoordeling vooraf én een mogelijk latere kwantitatieve analyse achteraf.

Werkhypothesen:
WH1: Naarmate de voorlichting specifieker op de doelgroep wordt gericht, neemt de compliance toe.

WH2: Communicatie in de vorm van gerechtvaardigde feedback draagt bij aan een hogere compliance.

WH3: Een hoge intensiteit van toezicht (rakingspercentage) van onder toezicht gestelden door toezichtinstrumenten verhoogt de compliance.

WH4: Hoe specifieker de keuze van (de mix van) toezichtinstrumenten op de diverse doelgroepen is afgestemd, hoe groter de effectiviteit hiervan op (het behoud van) compliance bij de betreffende doelgroepen.

WH5: Hoe beter handhaving (correctie van gedrag) is afgestemd op een bepaalde doelgroep, hoe groter het effect op compliance.

WH6: Hoe specifieker een sanctie wordt afgestemd op de betreffende overtreding van wet- en regelgeving, hoe beter het effect op compliance.

WH7: Hoe beter de afweging tussen de aard en omvang van de inzet van toezichtinstrumenten hoe groter de impact op compliance is.

WH8: Structurele vastlegging van kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten op het moment dat de gebeurtenis zich voordoet verhoogt de betrouwbaarheid van de effectmeting op compliance.

3. Onderzoeksmethodiek

In dit hoofdstuk worden de kernbegrippen gedefinieerd en geoperationaliseerd. Daarnaast wordt aangegeven hoe het onderzoek en de analyse van de verkregen data is vormgegeven en de werkhypothesen getoetst worden.

3.1. Operationalisatie

Definities
Onder compliance wordt voor dit onderzoek verstaan: de bereidheid van burgers en bedrijven om uit zichzelf (fiscale) regels na te leven (gedrag), zonder dwingende en kostbare acties van de Belastingdienst (Ministerie van Financiën, 2019). Onder dwingende en kostbare acties kun je denken aan een juridisch traject (dwingend) van correctie van de aangifte, eventueel gevolgd door bezwaar en beroep, maar ook aan het instellen van een boekenonderzoek of fraudeonderzoek (kostbaar) door de Belastingdienst. Ook kan er sprake zijn van het opleggen van boetes en/of het inzetten van dwanginvordering. Het gedrag is dan niet meer vrijwillig, maar afgedwongen door derden in casu de Belastingdienst of de rechter. Compliance wordt gemeten aan de hand van de door de Belastingdienst benoemde compliance factoren, die bij de prestatie-indicatoren ten aanzien van de algemene doelstellingen in het jaarplan 2020 zijn benoemd (Ministerie van Financiën, 2019):

1. Tijdigheid van aangifte;
2. Juistheid en volledigheid van aangiften;
3. Tijdigheid en volledigheid van betalingen.

Onder toezicht wordt voor dit onderzoek verstaan het geheel aan acties op initiatief van de Belastingdienst gericht op dataverzameling, oordeelsvorming en interventie, met als doel het verhogen van de naleving van wet- en regelgeving.

Onder effectmeting wordt voor dit onderzoek verstaan de mate waarin het effect van het toezichtinstrument op compliance in beeld kan worden gebracht.

Operationalisatie van de concepten:

Kernconcept	**Dimensie**	**Subdimensie**	**Meetpunten**
Compliance (naleving van wet- en regelgeving)	Gedrag	Naleving regelgeving	Nalevingspercentage tijdig doen van aangifte
			Nalevingspercentage juist en volledig doen van aangifte
			Nalevingspercentage tijdige en volledige betaling

Toezicht	Toezichtinstrumenten	Reden/doel van inzetten	- Preventief en proactief (voorlichten/informeren) - Feedback geven - Toetsing (beoordeling, output gericht) - Toetsing (beoordeling, handhaving/correctie) - Effectgericht
Effectiviteit	Behoud van gewenst gedrag en beïnvloeding van ongewenst gedrag	Bewuste en beoogde impact door toezichtinstrumenten	Toename van het nalevingspercentage van de drie benoemde compliance factoren

Tabel 1. Operationalisatie van de kernconcepten 'compliance', 'toezicht' en 'effectiviteit'.

Overige begrippen	**Definitie**	**Meetpunten**
Doelgroep	Een afgebakende groep onder toezicht gestelden	Objectieve criteria: branche, omzet, rechtsvorm, huidige naleving compliance etc.
Rakingspercentage	De mate waarin de onder toezicht gestelden te maken hebben gehad met beoordeling en/of handhaving	Aantal unieke toezichtmomenten, uitgedrukt in percentage van het totaal aantal onder toezicht gestelden
Onder toezicht gestelden	Alle burgers en bedrijven die onder het toezicht van de Belastingdienst MKB staan	Aantallen geregistreerde belastingplichtigen
Toezichtinstrumenten	Alle instrumenten die belastingplichten ervaren als toezicht	Aantallen per soort instrument, bij voorkeur gekoppeld aan een individuele belastingplichtige/belastingschuldige

Tabel 2. Operationalisatie van de overige begrippen uit de werkwerkhypotheses.

3.2. Onderzoekstrategie

Onderzoeksopzet

Het doel van deze onderzoeksopzet is de aanpak van het onderzoek te beschrijven en daarmee de haalbaarheid van het onderzoek aan te geven.

Soort onderzoek

Dit onderzoek betreft een kwalitatief onderzoek. Met dit onderzoek worden de begrippen 'toezicht', 'compliance' en het effect van toezicht op compliance in beeld gebracht en vergeleken met de invulling die de Belastingdienst MKB hieraan gegeven heeft. Er wordt onderzocht welke toezichtinstrumenten voor het verhogen van de compliance ingezet worden en hoe de effectiviteit daarvan in beeld wordt gebracht. Naast het onderzoeken van de hoofdvraag zal dit onderzoek ook inzicht geven in het proces rondom de huidige effectmeting. Het onderzoek zal uitgevoerd worden door middel van een combinatie van deskresearch en semigestructureerde interviews.

Deskresearch

Deskresearch heeft plaatsvinden op basis van het jaarplan 2020 (Ministerie van Financiën, 2019), de handhavingsbrief 2018 (Ministerie van Financiën, 2017), de Beleidsdoorlichting 2017 (Ministerie van Financiën, 2017) en de begrotingen paragraaf IX van 2016 tot en met 2021. Deze documenten zijn beoordeeld op aanwezigheid van (aanvullend) beschreven meetpunten ten aanzien van de effectmeting, beschrijving van gehanteerde handmatige en/of digitale vastleggingen en de uitkomsten van de effectmetingen.

Respondenten

Op basis van de meetpunten vanuit de literatuur zijn semigestructureerde interviews gehouden met 12 medewerkers van de Belastingdienst. Het betreft (landelijke en regionale) handhavingsadviseurs, procesregisseurs, (landelijke en regionale) projectleiders en een wetenschappelijk medewerker.

De keuze voor deze medewerkers is gebaseerd op de aanname dat zij het overzicht hebben op de redenen waarom er voor de inzet van bepaalde toezichtinstrumenten is gekozen en aan kunnen geven welk effect hiermee is beoogd. Tevens zouden zij aan moeten kunnen geven op welke wijze data wordt verzameld om zich te kunnen verantwoorden over de inzet van mensen en middelen. Tot slot kunnen zij aangeven of dit (in)zicht voldoende is om effectmeting te kunnen uitvoeren of wat hier aanvullend voor nodig zou zijn.

Semi-gestructureerde interviews

Door middel van semi-gestructureerde interviews wordt onderzocht hoe de (bewuste) keuze wordt gemaakt ten aanzien van de ingezette toezichtinstrumenten en welk effect hiermee is beoogd én verkregen. Ook wordt onderzocht of en hoe het effect in beeld wordt gebracht en in hoeverre er ondersteuning is vanuit structurele vastleggingen om het beoogde effect in beeld te kúnnen brengen. Daarnaast wordt onderzocht welke behoefte er is aan ondersteuning om de effectmeting te verbeteren. Tot slot wordt in beeld gebracht welke relevante beoordelingspunten de respondenten aanvullend kunnen benoemen vanuit de praktijk. Het doel is te onderzoeken in hoeverre de bijdrage van het toezicht aan de compliance factoren, en de keuzes voor de diverse toezichtinstrumenten hierbij, in beeld gebracht wordt en welke aanbevelingen kunnen worden gedaan om dit te verbeteren.

Ten behoeve van de semi-gestructureerde interviews is de volgende vragenlijst samengesteld. Deze is ook apart in bijlage 1 opgenomen.

Vraag 1a. Welke toezichtinstrumenten zijn ingezet bij…? (afhankelijk van de respondent)
Vraag 1b. Is er ook gebruik gemaakt van communicatie in de vorm van het geven van gerechtvaardigde feedback?

Vraag 2a. Wat was de reden om voor deze toezichtinstrumenten te kiezen?
Vraag 2b. Is hierbij ook specifiek onderscheid gemaakt naar doelgroep? En zo ja, waarom?
Vraag 2c. Wat was de reden om voor deze mix van toezichtinstrumenten te kiezen?
Vraag 2d. Is de intensiteit van het toezicht (rakingspercentage) onderdeel geweest van de afweging van deze mix?
Vraag 2e. Is er specifiek aandacht geweest voor het inzetten van een sanctie t.a.v. overtreding van wet- en regelgeving?

Vraag 3. Wat was het verwachte/beoogde effect op de compliance?

Vraag 4. Op basis waarvan werd dit effect verwacht?

Vraag 5. Is dit effect in beeld gebracht? En zo ja, hoe?
Vraag 5a. Worden de kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten structureel vastgelegd?
Vraag 5b. Indien niet: is dat wel mogelijk? En zo ja, hoe?
Vraag 5c. Is het in beeld gebrachte effect betrouwbaar? Zo ja, waarom? Ze nee, waarom niet?

Vraag 6. Indien dit effect niet in beeld is gebracht, welke mogelijkheden ziet u om dit in beeld te brengen?
Vraag 6a. Zijn de kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten structureel vast te leggen?
Vraag 6b. Zo nee, waarom niet? Zo ja, hoe wel?

Vraag 7a. Was het beoogde effect ook te verwachten geweest bij de inzet van andere instrumenten?
Vraag 7b. Kunt u dit toelichten?

Vraag 8a. Zijn er ook effecten opgetreden die niet vooraf verwacht werden?
Vraag 8b.En zo ja, kunt u dit toelichten?

Analysemethode
Bij de analyse zullen de benoemde meetpunten uit de literatuur en de praktijk vergeleken en beoordeeld worden op hun relevantie voor de effectmeting bij de Belastingdienst MKB. Voor een deel betreft dit een vergelijking met reeds gevonden literatuur, voor een deel zal wellicht aanvullende theorie gezocht moeten worden om aangereikte meetpunten te valideren of af te wijzen als relevant voor effectmeting bij de Belastingdienst MKB.

Bij de codering van de interviews is gebruik gemaakt van een 26-tal begrippen. Voor de verwerking van de data ten behoeve van de analyse is gebruik gemaakt van Excel.

De werkhypothesen
Bij iedere respondent is op basis van de verstrekte informatie geanalyseerd welke toezichtinstrumenten ingezet zijn, welke compliance factor hiermee beoogd is te beïnvloeden en welke informatie en/of data beschikbaar waren om dit effect in beeld te brengen. Als het gerealiseerde effect op compliance op enige wijze met data onderbouwd kan worden zal de werkhypothese bevestigd beantwoord worden. Afhankelijk van de omvang en mate van bevestiging is een onderscheid gemaakt in de bewoording van deze bevestiging. Hierbij zijn de volgende gradaties gehanteerd: aannemelijk, getoetst door pilot, getoetst door breed onderzoek, bewezen met gebruik making van gestructureerd vastgelegde data. De keuze voor deze gradaties zal per werkhypothese nader toegelicht en onderbouwd worden. Indien meerdere respondenten een gerealiseerd effect kunnen onderbouwen, zal dit optellen tot een steviger uitspraak.
Bovenstaande aanpak sluit aan bij de gekozen definitie van effectmeting: mate waarin het effect van het toezichtinstrument op compliance in beeld kan worden gebracht.

Om het onderzoek robuuster te maken zijn de werkhypothesen aan het eind van het interview expliciet voorgelegd aan de respondenten. De beantwoording door de respondenten is vertaald naar 6 opties: Volledig mee eens, mee eens, gedeeltelijk mee eens, gedeeltelijk mee oneens, mee oneens, geen uitspraak.

3.3. Betrouwbaarheid en validiteit

Interne validiteit
Voor dit onderzoek zijn interne documenten gebruikt en deskundigen geïnterviewd, die kennis hebben van de handhavingsregie ten aanzien van de beoogde gedragsbeïnvloeding en de voor de Belastingdienst MKB beschreven beleidsstrategie. Vanuit de wetenschappelijke theorie zijn voor dit onderzoek afhankelijke variabelen benoemd, die van invloed zijn op de compliance van burgers en bedrijven. Door zowel te

kiezen voor de invalshoek van beschreven effectmeting in documenten, interviews met deskundigen als voor een check van de werkhypothesen bij de geïnterviewden, is beoogd de data onderling te checken en de interne validiteit te borgen. Ook de onderbouwing van geclaimde uitkomsten van effectmeting is, voor zover beschreven en vastgelegd, gecheckt aan de hand van opgevraagde documenten.
De combinatie van bovenstaande activiteiten vergroot de betrouwbaarheid van het onderzoek.

Externe validiteit
Dit onderzoek vindt plaats in de hele specifieke omgeving van de Belastingdienst MKB. Mogelijk zijn (onderdelen van) de uitkomsten en aanbevelingen van toepassing op andere onderdelen van de Belastingdienst of andere toezichthouders.
De naleving van wet- en regelgeving, en de daarbij behorende variabelen, gelden breder dan alleen voor de Belastingdienst. Uitspraken over specifieke toezichtinstrumenten die de Belastingdienst hanteert zullen wellicht lastig te vertalen zijn naar andere onderdelen of toezichthouders. Bevindingen over de wijze van het proces van vastleggen en effectmeting zullen buiten de omgeving van dit onderzoek bruikbaar zijn.

4. Onderzoek

In dit hoofdstuk wordt allereerst de documentanalyse weergegeven, gevolgd door de uitkomsten van de interviews. In de derde paragraaf wordt de evaluatie van de werkhypotheses besproken en in de laatste paragraaf wordt een samenvatting van het onderzoek gegeven.

4.1. Analyse (beleids)documenten

In verschillende beleidspublicaties is aandacht besteed aan de compliance factoren, het nalevingsgedrag en de gehanteerde handhavingsstrategie om deze te vergroten. In deze paragraaf worden de relevante passages uit diverse publicaties voor dit onderzoek vermeld om zo de context te schetsen van het huidige toezicht en de wijze van in beeld brengen van de resultaten daarvan.

Handhavingsbrief Belastingdienst 2018 (Dienstverlening, toezicht en opsporing)
Als aanleiding voor deze handhavingsbrief (Ministerie van Financiën, 2017) wordt het onderzoek van de Algemene Rekenkamer uit 2016 naar het handhavingsbeleid van de Belastingdienst genoemd. Hierin is te lezen dat de Algemene Rekenkamer van mening is dat "de Belastingdienst een goed doordacht beleid heeft dat systematisch en gestructureerd inspeelt op de oorzaken van het nalevingsgedrag van belastingplicht", maar ook dat "de Belastingdienst nog informatie mist voor de verdere onderbouwing van de keuzes in het handhavingsbeleid" (Algemene Rekenkamer, 2016). In de handhavingsbrief wordt aangegeven dat "de Belastingdienst hiermee inzicht wil geven in de afwegingen die zij maakt over de inzet van mensen en middelen naar de verschillende doelgroepen en activiteiten". Daarnaast wordt vermeld dat de Belastingdienst zich hierbij meer wil richten op beïnvloeding van gedrag en zich daarvoor ook meer wil focussen op de oorzaken van niet-naleving in plaats van het alleen achteraf controleren van aangiften. Aangegeven wordt dat dit ook zijn oorzaak heeft in capaciteitsgebrek om zich volledig op de traditionele toezichtinstrumenten zoals controle van aangiften en boekenonderzoeken te kunnen richten.

Als instrumenten die de Belastingdienst gebruikt om goede naleving van wet- en regelgeving te bevorderen worden genoemd (Ministerie van Financiën, 2017, p. 3):
- actieve en duidelijke communicatie over wet- en regelgeving zodat burgers en bedrijven kunnen weten wat van hen wordt verwacht;
- goede dienstverlening zodat burgers en bedrijven gemakkelijk aan hun verplichtingen kunnen voldoen;

- risicogericht en zichtbaar toezicht om burgers en bedrijven bewust te maken van de consequenties van niet naleven van hun verplichtingen, waaronder corrigerend optreden.

In de paragraaf in de Handhavingsbrief 2018 over MKB (Ministerie van Financiën, 2017, pp. 25-26) wordt aangegeven dat de samenwerking met fiscaal dienstverleners, al dan niet in het kader van afspraken onder horizontaal toezicht, een belangrijk onderdeel is in de strategie voor het borgen van de juistheid van de aangiften. Daarnaast wordt het standpunt ingenomen dat fouten in de aangifte, met name in het Kleinbedrijf, meer met onkunde en slordigheid te maken hebben dan met onwil. De doelgroep van Kleinbedrijf vertegenwoordigt met 1,5 miljoen ondernemers de grootste groep binnen het MKB-segment met in totaal 1,9 miljoen ondernemers. De strategie voor Kleinbedrijf is er volgens de Handhavingsbrief 2018 dan ook op gericht om, via een goed ingerichte administratie (via gecertificeerde online software) en een vorm van vooraf ingevulde aangiften, het doen van aangifte voor zzp'ers veel gemakkelijker te maken en daarmee de kwaliteit van de aangifte (juist en volledig) toe te laten nemen. Ook wordt aangegeven dat er sterk wordt ingezet op begeleiding van de startende ondernemer. Voor het Middenbedrijf worden verder voorbeelden van vooroverleg en samenwerking met andere overheidspartners genoemd als onderdeel van de aanpak.

Ten aanzien van de mix van toezichtinstrumenten wordt aangegeven dat dit een keuzeproces is op basis van de strategie die per groep wordt bepaald. Hieruit volgt de inzet op capaciteit en competenties. Dit betreft de invulling van de handhavingsregie op tactisch niveau (Ministerie van Financiën, 2017, p. 12).

Om deze cyclus ten aanzien van de handhavingsregie compleet te maken is het volgens de Algemene Rekenkamer van belang om de aannames op basis waarvan de keuze werd gemaakt te toetsen op het beoogde effect. Deze effectmeting zou de mogelijkheid geven om de kennis te vergroten en de strategie waar nodig bij te stellen (Ministerie van Financiën, 2017, p. 13).

Toezichtplan Arbeidsrelaties (oktober 2019 – december 2020)

In het toezichtplan Arbeidsrelaties (Belastingdienst, 2020) is goed terug te zien hoe toezicht vormgegeven kan worden door een bewuste combinatie van inzet van interventies. In onderstaand kader, als voorbeeld, een deel uit dit toezichtplan om inzichtelijk te maken hoe de uitwerking van het handhavingsbeleid eruit kan zien. Aanleiding voor dit toezichtplan was het besef dat er in een behoorlijk aantal gevallen sprake was van gebrek aan kennis van wet- en regelgeving en/of van in meer of mindere mate onjuist handelen. Daarnaast vonden de opdrachtgevers de Wet deregulering

beoordeling arbeidsrelaties (DBA) complex, was handhaving voor de Belastingdienst niet eenvoudig vanwege de zware bewijslast en leidde contact met de opdrachtgevers meer dan eens tot aanpassing van de werkwijze en dus tot verbetering. Het toezichtplan bouwde voort op het toezichtplan uit 2018.

Het toezichtplan wordt in lijn met de strategie ingevuld met een mix aan instrumenten:
- algemene communicatie, waaronder informatie via website, voorlichting en publicatie van modelovereenkomsten en via handhavingscommunicatie;
- beoordelen van voorgelegde modelovereenkomsten;
- vooroverleg conform Besluit Fiscaal Bestuursrecht en beschikking verzekeringsplicht;
- bedrijfsbezoeken en boekenonderzoeken bij individuele opdrachtgevers;
- sectoraanpak als problematiek breed in de sector speelt.

De Belastingdienst maakt zo gericht mogelijk gebruik van bedrijfs- en boekenonderzoeken door:
- vervolg te geven aan de eerdere bedrijfsbezoeken uit het Toezichtplan 2018;
- een selectie te maken uit de goedgekeurde, niet goedgekeurde en ingetrokken modelovereenkomsten;
- op basis van risicodetectie met behulp van query's uit systemen; en
- op basis van andere signalen, zoals personeelsadvertenties etc.

Bron: Toezichtplan Arbeidsrelaties (2019-2020)

Rijksbegroting IX (2016 tm 2021)

In de begroting van het ministerie van Financiën wordt in paragraaf IX (Financiën en Nationale Schuld) aangegeven wat de gerealiseerde en te realiseren doelstellingen zijn. Onderstaand eerst een overzicht van de beschikbare kengetallen ten aanzien van de compliance doelstellingen over de jaren 2013 tot en met 2019 vanuit het Rijksjaarverslag 2016 (Ministerie van Financiën, 2017) en de Rijksbegrotingen 2017 tot en met 2021 (Ministerie van Financiën, 2016) (Ministerie van Financiën, 2017) (Ministerie van Financiën, 2018) (Ministerie van Financiën, 2019) (Ministerie van Financiën, 2020).

Kengetallen in % (gerealiseerd)	2013	2014	2015	2016	2017	2018	2019
Percentage aangiften omzetbelasting tijdig ontvangen	95,3	95,1	95,3	94,9	95,3	95,6	95,6
Percentage aangiften loonheffingen tijdig ontvangen	99,1	99,2	99,2	99,2	99,3	99,1	99,1
Percentage aangiften inkomensheffingen en vennootschapsbelasting tijdig ontvangen	nb	nb	94,5	94,5	nvt	94,7	91,9
Juist en volledig aangifte doen; percentage nalevingstekort op basis van de steekproef Midden- en kleinbedrijf (uitgedrukt als percentage van de totale belastingopbrengst voor dit segment)[1]	4,4	nvt	6,0	5,2	nvt	5,7	nvt
Tijdige betaling van belastingen en premies	nb	98,4	98,4	nvt	98,4	98,5	98,5
Oninbare belastingen en premies	0,6	0,6	0,6	0,3	0,3	0,3	nb

Tabel 3. Overzicht kengetallen compliance factoren (Bron: Rijksjaarverslag 2016 IX, paragraaf 4.1. en Rijksbegrotingen IX artikel 1 2017 - 2021)

[1] De uitkomsten van de steekproef Midden- en kleinbedrijf die in 2016 zijn vastgesteld hebben betrekking op de aangiftejaren 2013. De uitkomsten mogen niet worden verabsoluteerd aangezien het om een schatting gaat.

Verdere toelichting:
- Daar waar 'nvt' vermeld staat betekent dit dat de prestatie-indicator niet of anders is gemeten of gerapporteerd in de begroting, waardoor vergelijking over de jaren niet mogelijk is. Daar waar 'nb' vermeld staat waren (nog) geen cijfers over dit aspect in de begroting opgenomen.
- Aanvullend op bovenstaande cijfers blijkt uit de Beleidsdoorlichting toezicht en opsporing en massale processen Belastingdienst van oktober 2017 dat het nalevingstekort van ongeveer 5% van de totale verschuldigde belasting een resultante is van ingediende aangiften en aangebrachte correcties. Bij ongeveer 60% van de aangiften van MKB-ondernemingen wordt geen enkele correctie aangebracht. Ter vergelijking: bij Particulieren blijkt uit steekproeven dat ongeveer 90% van de aangiften (inkomstenbelasting) zonder fouten is (Ministerie van Financiën, 2017).

Uit bovenstaand overzicht, en de toelichting in de begrotingen, blijkt dat de kengetallen van het tijdig doen van de aangiften van de diverse belastingmiddelen en de tijdige en volledige betaling in 'harde' cijfers zijn weer te geven. Het kengetal over juiste en volledige aangifte komt tot stand op basis van een steekproef. Het feit dat over de jaren 2013 tot en met 2019 drie keer 'n.v.t.' staat vermeld geeft aan dat deze compliance factor het meest lastig te meten is en blijkbaar gezocht wordt naar een juiste meting en weergave van dit kengetal, waardoor vergelijking over de jaren niet mogelijk is.

Voor 2021 zijn twee nieuwe kengetallen toegevoegd: belastingmoraal en vertrouwen in de Belastingdienst. Met belastingmoraal wordt bedoeld de intrinsieke motivatie van burgers en bedrijven om belastingverplichtingen na te komen. Aangegeven wordt dat deze de bepalende, en mogelijk zelfs de meest bepalende, factor is voor bereidwillige naleving van belastingverplichtingen door belastingplichtigen. Dit inzicht wordt van strategisch belang geacht. Het kengetal zal ieder jaar gemeten worden met behulp van een monitor. Het vertrouwen in de Belastingdienst wordt als een belangrijke determinant van compliance en belastingmoraal beschouwd. Met dit kengetal wil de Belastingdienst drie dimensies meten: zien burgers en bedrijven de Belastingdienst als capabel, in staat om zijn taken goed uit te voeren (competentie), zien zij de Belastingdienst als betrokken met oog voor hun belangen (welwillendheid) en zien zij de Belastingdienst als eerlijk, rechtvaardig, consistent en voorspelbaar (integriteit). De dimensie 'welwillendheid' kunnen we, gezien de huidige tijdgeest, zien als een vertaling van de 'menselijke maat', zoals deze inmiddels in ons huidige spraakgebruik is ingeburgerd. Met de prominente aandacht voor het vertrouwen van burgers en bedrijven in de Belastingdienst wordt ook een hernieuwde invulling gegeven aan het maatschappelijk belang hiervan.

Prestatie-indicatoren

In uitwerking van de kengetallen over compliance worden prestatie-indicatoren benoemd ten aanzien van de resultaten. Tot 2016 waren onder andere het *aantal* behandelde aangiften IH/Vpb en boekenonderzoeken een graadmeter voor het behaalde *resultaat*. Van 2017 tot en met 2020 is dit gewijzigd in een bruto correctie opbrengst voor de

aangiften en boekenonderzoeken. In diezelfde periode veranderde de doelstelling voor MKB-ondernemingen onder een fiscaal dienstverlenersconvenant ook van aantallen naar, in dit geval, een percentage kasontvangsten. Daarnaast bleef over de jaren 2016 tot en met 2020 een percentage ingediende bezwaren naar aanleiding van een correctie een meetpunt.

Voor 2021 zien we hierin een drastische wijziging naar prestatie-indicatoren (in de begroting wordt gesproken van een 'herziene set') ten aanzien van *adequate behandeling* (juist, tijdig en proportioneel). Hierbij zijn al indicatoren benoemd voor een beoogd percentage tijdige afhandeling van bezwaren en klachten en indicatoren voor het tijdig terugbellen en tijdig behandelen van verzoeken tot registratie. Ook is een indicator voor kwaliteit toegevoegd. Een aparte indicator voor hoe burgers en bedrijven hun behandeling als adequaat hebben ervaren (correct en passend) is nog niet bekend. De genoemde indicatoren hebben allen betrekking op het onderdeel 'toezicht', maar niet allemaal met de keuze en effectiviteit van een interventie, wat de focus is van dit onderzoek. Als beleidsinstrumenten naast toezicht wordt ook 'effectief informeren' benoemd als onderdeel van dienstverlening. Onder dit zogenaamde 'tussendoel' is een veelheid aan prestatie-indicatoren benoemd voor klant(on)tevredenheid over bereikbaarheid en kwaliteit van de belastingtelefoon, website en balie. Als derde tussendoel wordt 'gemak' benoemd. Omdat dit geen relatie heeft met het inzetten van interventies, de focus van dit onderzoek, wordt dit aspect niet verder toegelicht.
Tot slot wordt het tussendoel 'corrigerend optreden' benoemd. Hier worden concrete indicatoren benoemd voor fraudeaanpak en invordering. Daarnaast worden nog indicatoren benoemd voor het percentage correctie van de belastingmiddelen bij nalevingstekort. Deze indicator is bedoeld ter vervanging van de bruto opbrengst en beoogd de effectiviteit van de handhaving te meten. De meetwaarde van deze indicator is nog niet bekend.

In de begroting van 2021 zien we voor het eerst een vollediger doorvertaling van de handhavingsstrategie op de aspecten van dienstverlening, toezicht en fraudebestrijding naar prestatie-indicatoren. De focus op aangiften en boekenonderzoeken wordt hiermee verbreed met interventies, waaronder toezichtinstrumenten, die niet eerder als indicator benoemd waren.

Jaarplan 2020 Belastingdienst

In het jaarplan 2020 (Ministerie van Financiën, 2019) staat de volgende strategie vermeldt: "Onze meerjarige strategie laat zich in één zin samenvatten: we streven er naar dat burgers en bedrijven bereid zijn uit zichzelf (fiscale) regels na te leven, zonder

dwingende en kostbare acties van de Belastingdienst". De strategie wordt inmiddels beschreven als een uitvoerings- en toezichtstrategie (eerder handhavingsstrategie), waarmee expliciet een verbreding naar reguliere uitvoering en ook interactie wordt gemaakt, die beter aansluit op de praktijk. Er wordt verwezen naar drie beleidsinstrumenten: voorkomen van fouten, dienstverlening, toezicht en fraudeaanpak. Bij de inzet van de deze instrumenten wordt een optimale mix beoogd met een optimaal effect op naleving.

Bovenstaande documenten vormen het kader waarbinnen de analyse van de interviews begrepen kan worden.

4.2. Analyse interviews respondenten

In deze paragraaf worden de uitkomsten van de interviews besproken. Allereerst wordt de aandacht voor de verschillenden doelgroepen benoemd. Daarna worden de benoemde interventies vermeld en wordt aangegeven wat de reden was om deze (toezicht)instrumenten in te zetten. Ook wordt inzicht gegeven in de handhavingsmix die is gehanteerd. Vervolgens wordt besproken wat het verwachte effect was van de individuele interventies, of de mix daarvan, en of dit in beeld gebracht is.

4.2.1. Doelgroepen

Acht respondenten gaven aan te werken met een bepaalde onderverdeling in doelgroepen of deze te herkennen in de aanpak waarbij zij betrokken waren. Bij vier respondenten was dit niet aan de orde. Eén respondent zag een doelgroepindeling als wens voor de toekomst en één respondent vond het juist ongewenst om vooraf een indeling te maken. In dit laatste geval was er dan ook geen sprake van een behandeling van een bepaalde doelgroep met bijvoorbeeld een specifiek probleem, maar juist van een generieke geïntegreerde aanpak. Benoemde doelgroepen waren:
- (sub)branches (5 keer, waarvan één keer als aanvullend criterium), vanwege specifieke problematiek
- gebied (fysiek),
- daar waar (fiscale) noodzaak vanwege bepaalde problematiek werd gezien over de branches heen (2 keer)
- organisatorische segmenten (klein, midden, groot)
- indeling van compliant gedrag (volstrekt compliant, hulp nodig, grens verkennend en malafide)
- wel/niet gecertificeerd (als aanvullend criterium)
- onderscheid naar belang en risico

- al dan niet aangesloten bij fiscaal dienstverlener (als aanvullend criterium)
- startende ondernemers
- verbindende factor (brancheorganisatie, gedrag, georganiseerdheid, fiscaal dienstverlener, geografisch, gemeenschappelijk probleem)

Deelconclusie doelgroepen

Daar waar de WRR aangeeft dat veel toezichthouders hun toezichtinstrumenten koppelen aan de mate waarin onder toezicht gestelden compliant gedrag vertonen (WRR, 2013, p. 30), blijkt uit dit onderzoek dat de doelgroepen, als ordeningscriterium, in de praktijk op veel diverse manieren tot stand komen en ook om verschillende redenen. Het gaat dan om organisatorische indelingen, indelingen op basis van fiscaal risico of fiscaal belang (waaronder branches met specifieke problematiek) of aansluiting bij andere organisaties als ingang voor contact. In alle gevallen betreft het keuzes die bewust en onderbouwd zijn gemaakt en ook een vervolg kregen in de behandelwijze en een gerichte inzet van interventies. De keuze voor de diverse doelgroepen is ook breder dan staat beschreven in de Handhavingsbrief 2018, waar met name een onderscheid gemaakt wordt tussen Kleinbedrijf, waaronder starters, en Middenbedrijf (Ministerie van Financiën, 2017). Alleen de aandacht voor starters sluit aan bij de strategische keuze voor deze doelgroep. Daar waar in de Handhavingsstrategie gesproken wordt over samenwerking met andere overheidspartners zien we in de praktijk een groter accent op samenwerking met brancheorganisaties en fiscaal dienstverleners.

4.2.2. Interventies en de gehanteerde handhavingsmix

In deze paragraaf worden achtereenvolgens de verschillende interventies vermeld die door de respondenten zijn benoemd. Aangeven wordt wat volgens hen de reden was om deze interventies in te zetten. Hieruit blijkt dat niet iedere interventie als toezichtinstrument gebruikt wordt. Zoals in het theoretisch kader is aangegeven wordt met toezicht in het kader van dit onderzoek díe interventie bedoeld die een reactie zijn op een beoordeling. Gekozen is om de interventies te ordenen naar proactief, preventief en reactief. Uit het onderzoek is echter gebleken dat deze ordening niet één op één aangebracht kan worden omdat interventies om meerdere redenen ingezet kunnen worden. Om vooraf structuur aan te brengen wordt allereerst een overzicht gegeven van de benoemde redenen van het inzetten van die interventies in combinatie met de ordering naar proactief, preventief en reactief. Daarna zal per interventie en/of toezichtinstrument een analyse gegeven worden van de uitkomsten van de interviews.

Interventies	Proactief	Preventief	Reactief	
			Toezicht	Handhaving
	informeren, wijzen op mogelijkheden / ondersteunende tools	wijzen op mogelijk gebrek wat nog voorkomen of hersteld kan worden (algemeen geconstateerde fouten, fiscaaltechnisch of branche specifiek)	bespreken van vragen / informatie ophalen n.a.v. beoordeling (beoordeling aangifte, feedback n.a.v. verzuim, betalingsachterstand)	(Voorgenomen) correctie/ handeling, eventueel gevolgd door een sanctie (boete)
Voorlichting (individueel, groepsgewijs, via website of media)	x			
Actualiteitenbezoek	x	x		
Vooroverleg*			X	
Behandelen bezwaar*			X	x
Bellen	x	x	X	x
Brieven	x	x	X	x
Samenwerking horizontaal toezicht /	x	x	X	x
Beoordelen aangifte			X	x
Boekenonderzoek			X	x
Vaststellingsovereenkomst				x

Tabel 4. Overzicht van interventies en ordening

* initiatief ligt bij belastingplichtige

<u>Nadruk op proactief en preventief</u>

Voorlichting

Zes respondenten gaven aan voorlichting in te zetten of met dit instrument bekend te zijn. Naast het informeren en kennis laten vergaren werd ook door één respondent specifiek aangegeven dat de reden van inzetten van dit instrument bewustwording was. Diverse vormen van voorlichting werden benoemd: startersvoorlichting 'De goede start' en intermediairdagen, maar ook berichten op social media, gebruik van Facebook en twitter.

Aangegeven werd dat met voorlichting wordt beoogd een grote groep te bereiken (hoge raking) en zowel ziet op het voeren van een goede administratie als op fiscale onderwerpen, om zo een juiste en volledige aangifte te kunnen doen.

Website

Door zes respondenten werd de inzet van de website als instrument benoemd. Hierbij ging het met name om de website van de Belastingdienst. Eén respondent gaf de inzet van de website van fiscaal dienstverleners of bestaande platforms aan als instrument om in te zetten voor verspreiding van handhavingscommunicatie. Bij de website van de Belastingdienst ging het om inhoudelijke informatieverstrekking, de ondernemer 'vertellen' wat hij in een bepaalde situatie moest doen (denk hierbij ook bijvoorbeeld aan uitstel van betaling aanvragen en dergelijke) of de mogelijkheid om ondersteuning te krijgen van tooling (hulpprogramma's) die daar te vinden zijn en overige handhavingscommunicatie. Wat door één respondent gemist werd was de chatfunctie op de site. De behoefte hieraan was tijdens interviews met ondernemers door hen aangegeven.

Publiciteit

Vier respondenten gaven aan gebruik te maken van de media om informatie te verstrekken die een bepaalde doelgroep raken, vooraankondigingen te doen of toelichting te geven over de inzet van een bepaalde actie. Als medium werd, naast de reguliere media, ook twee keer interviews in vakbladen benoemd als toezichtinstrument.

Actualiteitenbezoek (bedrijfsbezoek of waarneming ter plaatse)

Zes respondenten benoemden het actualiteitenbezoek als toezichtinstrument. Op twee na werd dit niet apart beargumenteerd, maar als onderdeel van de totale aanpak (mix van instrumenten) benoemd. Eén respondent gaf aan dit instrument in te zetten om informatie op te halen en het bedrijf te leren kennen. Reden was het verder verfijnen van branchegerichte acties. Een andere respondent gaf aan onderscheid te maken naar een informerend of een corrigerend gesprek. In beide gevallen was het gesprek bedoeld om de juistheid en tijdigheid van de aangifte te verhogen.

Feedback

Op basis van theoretische inzichten is de respondenten gevraagd of er gebruik wordt gemaakt van gerechtvaardigde feedback als toezichtinstrument. Bijna alle respondenten vonden dit een lastige vraag omdat dit geen gebruikelijke term is binnen de Belastingdienst.

Elf respondenten konden, na een toelichting over het reageren op gedrag, voorbeelden benoemen van wat hieronder zou kunnen vallen. Zes respondenten gingen uit van een setting waarin fysiek contact aan de orde was, in drie gevallen werd telefonische feedback benoemd en in twee gevallen een combinatie daarvan. Voorbeelden van feedback waren: feedback aan belastingplichtige over de administratie, gedrag of

aangifte tijdens een boekenonderzoek (4 respondenten), feedback tijdens gesprekken met de fiscaal dienstverlener (2 respondenten), feedback aan externe partners of feedback aan een groep (informerend of naar aanleiding van nalevingstekort).

Daarnaast werden voorbeelden gegeven van individueel telefonisch contact (bellen bij aangifteverzuim, waarschuwend over aankomende betaaltermijn) of interactieve feedback bij het gebruik van een digitale tool (adviserend of belonend).
Ook werden er voorbeelden gegeven van ontvangen feedback vanuit de belastingplichtige/belastingschuldige op het toezicht door de Belastingdienst bij een contact of de onbekendheid met de wijze waarop informatie wel of niet toegankelijk was.

Toezicht op verzoek van belastingplichtige

Door de respondenten is vooroverleg drie keer benoemd als toezichtinstrument, waarbij één respondent aangaf dat dit een te behandelen signaal is. Van bezwaarbehandeling werd drie keer melding gemaakt, waarbij één respondent aangaf dat dit toezicht op initiatief van belastingplichtige betreft. Een andere respondent gaf aan dat de Belastingdienst dit als een 'soort' toezichtinstrument ziet. In alle gevallen zat er een vraag achter ten aanzien van het kwalificeren van vooroverleg of bezwaar als toezichtinstrument of als uit te voeren proces. Dit leidt tot de conclusie dat, alhoewel er in beide gevallen duidelijk sprake is van toezicht (immers beoordeling), beide werkstromen niet echt als (in te zetten) instrument worden herkend en derhalve ook geen onderdeel uit (kunnen) maken van een bewust gekozen (mix van) instrument(en) om in te zetten om de compliance te verhogen. Uiteraard kan de (kwalitatieve) wijze waarop deze werkstromen uitgevoerd worden een effect hebben op compliance.

Toezichtinstrumenten die zowel proactief, preventie als reactief ingezet worden

Bellen

Zeven respondenten noemden bellen als toezichtinstrument. De redenen voor het inzetten van dit instrument waren divers. Drie respondenten gaven aan dat dit instrument ingezet werd bij aangifteverzuim. Het gaat dan om een dienstverlenend gesprek met als doel alsnog een (juiste en volledige) aangifte te ontvangen (feedback op gedrag). Twee respondenten gaven aan dit instrument ingezet te hebben in opvolging van handhavingscommunicatie (brief). In één geval was dit om aan te geven dat deze belastingplichtigen geen verder toezicht zouden krijgen nadat zij in eerste instantie betrokken waren in een actie, in het andere geval was dit om na te vragen wat het effect van een attentiebrief op hen was geweest. Dit betrof dus een vorm van informatie geven en informatie vragen. Twee respondenten gaven aan te bellen bij een aankomende vervaldatum van betalen of bij niet betalen te vragen of uitstel van betaling of verlenging

van uitstel nodig was. Dit laatste had ook te maken met het feit dat de belastingschuldige mogelijk niet op de hoogte was van de toen geldende aangepaste regelgeving vanwege Corona. Door vier respondenten werd aangegeven dat de (aanvullende) reden om dit instrument in te zetten was dat het een goedkoop instrument is danwel een hoge raking geeft, waarbij één keer werd aangegeven dat de hoge raking juist een *gevolg* was van de bewuste keuze om te bellen. Eén respondent gaf aan dat het bellen bij een startende ondernemer vooral ook bedoeld was om proactief en laagdrempelig te helpen bij vragen over administratieve verplichtingen en belastingplicht. Ook werd door één respondent aangegeven dat het bellen zowel voor het stellen van vragen, het geven van advies of een corrigerend gesprek kon zijn. Tot slot werd nog één keer aangegeven dat gebeld werd bij een voorgenomen correctie bij aangiftebehandeling. Gezien de gegeven voorbeelden kan geconcludeerd worden dat het instrument bellen breed wordt ingezet en zowel gebruikt wordt om te informeren, feedback te geven op betaal- en/of aangiftegedrag om dit gedrag te verbeteren (beoogd effect) als het afgeven van een waarschuwing of bespreking van een voorgenomen correctie nadat hiertoe vanuit toezicht (beoordelen) noodzaak toe werd gezien.

Handhavingscommunicatie (brieven)

Tien respondenten benoemden het sturen van brieven als toezichtinstrument. Door acht respondenten werd aangegeven dat de reden om brieven te sturen was dat er fouten waren geconstateerd in de aangifte of dat er vragen hierover waren (reactief). In twee gevallen werd aangegeven dat dit een nalevingstekort in een hele groep betrof en in één geval werd melding gemaakt van individuele brieven bij nalevingstekort. In beide gevallen was de reden het informeren van belastingplichtige/belastingschuldige en daarmee te waarschuwen zodat het gedrag kon worden aangepast. Door vijf respondenten werden brieven verstuurd om informatie te verstrekken (preventief). In één geval werd naast het verstrekken van informatie ook een vervolgactie aangekondigd. Tot slot werd door één respondent aangegeven dat er standaardbrieven naar aanleiding van stagnatie in de voortgang van het betalingsproces verstuurd werden (waarschuwen). Ook ten aanzien van het versturen van brieven is een brede diversiteit aan redenen genoemd. Naast informeren lag daar in een groot deel van de gegeven voorbeelden een bevinding vanuit toezicht aan ten grondslag.

Samenwerking en horizontaal toezicht

Alhoewel samenwerking geen specifiek onderdeel van de vragenlijst was waarop doorgevraagd werd, gaven acht respondenten uit zichzelf aan de samenwerking met andere (overheid)partners een (belangrijk) instrument te vinden om toezicht vorm te geven. Het betrof hier dan bijvoorbeeld de afstemming van de individuele toezichtaanpak en gebruik maken van elkaars expertise om vanuit de eigen aanpak een

gemeenschappelijk doel te bereiken. Ook het samen opzetten van acties met andere overheidspartners valt hieronder. Zeven keer werd het belang van samenwerking met de fiscaal dienstverleners benoemd, drie keer de samenwerking met brancheorganisaties, twee keer met overige platforms en twee keer met andere overheidspartners. Daarnaast werd ook twee keer het belang van samenwerking met de ondernemers genoemd.

De samenwerking met de fiscaal dienstverleners werd belangrijk gevonden als schakel voor de kwaliteit van de administratie, als ingang om branche specifieke issues te bespreken en daar afspraken over te maken, als ingang om publicaties verder te verspreiden en last but nog least als partner om horizontaal toezicht vorm te geven. Binnen het horizontaal toezicht werden nog twee aanvullende instrumenten benoemd: het kennis nemen van kwaliteitssystemen van fiscaal dienstverlener als onderdeel van het traject om deel te nemen aan het horizontaal toezicht en de evaluatie achteraf met fiscaal dienstverlener op basis van informatie over de fiscaal dienstverlener en zijn klanten.

De samenwerking met de brancheorganisaties werd gezocht om de kwaliteit te verbeteren. Enerzijds omdat soms gewerkt wordt met certificering en brancheorganisaties meekijken naar kwaliteit en ook omdat de respondenten verwachtten dat de organisaties de beoogde effecten van de Belastingdienst kunnen helpen uitdragen. De samenwerking met platforms werd gezocht om informatie met de achterban te kunnen delen en zo het kennisniveau te verhogen en daarmee het nakomen van fiscale verplichtingen te verbeteren. De samenwerking met andere overheidspartners kwam aan de orde om het toezicht op een bepaalde doelgroep met elkaar vorm te kunnen geven en aansluiting te houden in regelgeving en toezicht (barrièremodel).
Tot slot werd de samenwerking met ondernemers genoemd. Enerzijds om hun perceptie op de handhavingsacties te horen (ophalen mening/behoefte), anderzijds om een relatie op te bouwen en daarmee de compliance te vergroten. Daarmee heeft samenwerking in een aantal gevallen de vorm van een bewust gekozen interventie/toezichtinstrument, ook al is het contact met belastingplichtige indirect.

<u>Toezichtinstrumenten (reactief)</u>

Beoordelen aangifte
Acht respondenten gaven aan gebruik te maken van heffing als toezichtinstrument. Bij één respondent was dit instrument niet relevant voor het werkpakket en bij drie respondenten was dit geen expliciet onderdeel van hun aanpak. Door drie respondenten werd aangegeven dat dit een traditioneel en repressief instrument betreft wat tot doel heeft uitvoering te geven aan de wet (verplichting om de aangifte vast te stellen) en

daarnaast een instrument is met een opbrengstdoelstelling, zoals eerder bleek uit de begroting over de jaren 2017 tot en met 2020. Ook werd aangegeven dat de beoordeling van aangiften een bijdrage moet leveren aan de gepercipieerde pakkans door een bepaald volume (raking) te halen. Twee respondenten gaven aan dat het beoordelen van aangiften een (bewust) onderdeel van een gekozen handhavingsmix uitmaakte bij een bepaalde aanpak, terwijl één respondent aangaf dat dit juist alleen vanuit uitworp selectie of een steekproef zou moeten gebeuren. Eén respondent gaf aanvullend nog aan dat door het regelen van aangiften ook informatie verzameld wordt over wat er bij een bedrijf aan de hand is. Daarnaast werd aangegeven dat met het regelen van aangiften wordt beoogd een juiste en volledige aangifte bij het individu te bewerkstelligen, maar ook de ingeschatte pakkans bij een bredere groep te beïnvloeden en daarmee een juiste(re) en volledig(er) aangifte in de toekomst te borgen. Uit bovenstaande blijkt dat hier sprake is van toezicht (het *beoordelen* van een aangifte), maar dat de reden om dit instrument in te zetten verschillend is: een (bewuste) keuze om heffing vanuit een bepaalde aanpak, danwel het uitvoeren van een vooraf gekozen proces (selectie of steekproef) met een bepaalde omvang (jarenlange doelstelling begroting tot en met 2016).

Boekenonderzoek

Tien respondenten maken melding van het inzetten van het boekenonderzoek als toezichtinstrument. Bij twee respondenten was dit instrument niet relevant voor hun werkpakket. Voor het inzetten van dit instrument worden verschillende redenen gegeven. Drie respondenten gaven aan de het een politieke doelstelling was om boekenonderzoeken te doen, gekoppeld aan het borgen van opbrengst door een bepaalde omvang in aantallen. Dit sluit aan bij de begrotingsdoelstellingen van 2016 tot en met 2020. Vanaf 2021 staat het uitvoeren van boekenonderzoek als doelstelling niet meer in de begroting. Door twee respondenten werd aangegeven dat het uitvoeren van boekenonderzoeken (als klassiek instrument) een bewust onderdeel was van de handhavingsmix, maar dan wel beperkter en gerichter ingezet dan vroeger. Eén respondent gaf ook aan dat dit noodzakelijk toezicht betreft voor grens verkennende ondernemers (branche-aanpak), wat door een andere respondent expliciet werd verwoord dat een (repressief) boekenonderzoek wordt ingezet als de ondernemer niet doet wat de Belastingdienst wil. Eén respondent gaf aan dat een boekenonderzoek niet alleen met een repressief karakter kan worden ingezet, maar ook met het bredere doel om de administratie te beoordelen, te beoordelen of de onderneming ‘op orde’ is en ook gebruikt kan worden voor positieve feedback, waarbij de ondernemer ook in de gelegenheid zou kunnen worden gesteld om zelf zijn gevolgen van eventuele fouten aan te kunnen passen. Tot slot gaven twee respondenten aan dat boekenonderzoeken

worden ingezet vanwege een steekproef om inzicht te krijgen in het nalevingstekort. Overall werd aangegeven dat een boekenonderzoek een 'zwaar' instrument is wat de Belastingdienst veel tijd kost en ook belastend is voor de ondernemer.

Sanctie

Het inzetten van een sanctie werd door tien respondenten herkend als mogelijk gevolg van een correctie bij een boekenonderzoek vanuit de toepassing van het boetebeleid, maar niet als apart en bewust gekozen toezichtinstrument. Bij de andere twee respondenten was dit vanuit hun proces niet aan de orde. Daarnaast worden door de respondenten individueel nog aanvullende maatregelen benoemd die zij als (mogelijke) sanctie zien: verzuimboete bij het niet/te laat indienen van een aangifte, verzuimboete bij te laat betalen van de voorlopige aanslag, overweging bestuursverbod, opzeggen convenant bij horizontaal toezicht, inzetten waarschuwing tot en met strafrechtelijke vervolging als onderdeel van een vooraf gecommuniceerde aanpak. Twee respondenten gaven aanvullend aan dat een correctie al vervelend genoeg is en zetten vraagtekens bij de effectiviteit van een aanvullende sanctie.

Vaststellingsovereenkomst

Drie respondenten gaven aan gebruik te maken van een vaststellingsovereenkomst (VSO) als instrument om verschillen van inzicht uit een interventie (sneller) af te ronden en/of afspraken voor de toekomst te maken. Eén van de respondenten gaf aan hiermee gedragsverandering bij de belastingplichtige en adviesgedrag bij de adviseur te willen bewerkstelligen.

<u>Overige genoemde instrumenten</u>

Zes respondenten noemden individueel nog instrumenten die niet in bovenstaande categorieën passen en hieronder staan opgesomd:

- risicomodellering ten behoeve van selectie, selectiemodule
- uitvoeren steekproef om risicomodel te voeden
- query: hiermee werd aangegeven dat de toepassing van gedigitaliseerde selectiecriteria en de daaruit voortkomende selecties (query's) ook als toezichtinstrument werden beschouwd.
- protocollen: gericht informatie geven over wat wel/niet toelaatbaar is of toegestaan is
- inzetten strafrecht/FIOD (3 keer)
- ten aanzien van de invordering: aanmaning, dwangbevel, beslagopdracht, hernieuwd bevel tot betalen, beslag, verkoop in beslag genomen goederen (instrumenten en volgorde vloeien rechtstreeks voort uit de wet), allemaal ten doel om een tijdige en volledige betaling te stimuleren.

Gehanteerde handhavingsmix

Alle respondenten konden (uitgebreid) aangeven waarom zij voor een bepaalde inzet van een instrument of combinatie van instrumenten (handhavingsmix) hadden gekozen. Afhankelijk van de setting waarin zij werkzaam zijn kon dit een beperkte of zeer uitgebreide mix zijn. Opvallend was de nadruk die zij legden op het belang van aandacht aan de voorkant van het toezichtproces om fouten in de aangifte te voorkomen (preventie door informeren en bewust wording) en de bewuste/selectievere inzet van meer repressieve instrumenten zoals een boekenonderzoek.
Vrijwel alle respondenten gaven aan dat een accent op de voorkant van het toezichtproces ook hielp bij het continue keuzevraagstuk tussen beperkte capaciteit en gewenste beïnvloeding. Juist bij preventie werd een hoge raking belangrijk gevonden en werd geconstateerd dat de hogere raking juist een gevolg was van de keuze om de nadruk aan de voorkant (goedkoper) van het toezicht te leggen waardoor met dezelfde capaciteit meer contacten konden plaatsvinden. Door één respondent werd ook het belang van repressieve raking genoemd als belangrijk voor de gepercipieerde pakkans. Toezicht werd één keer omschreven als toezicht wat de ondernemer als zodanig ervaart, wat een andere betekenis kan hebben dat wat de Belastingdienst als toezicht bedoelt. Zo kan een zichtbare actie in 'jouw' branche of wijk bij andere ondernemers óók als toezicht ervaren worden door de belastingplichtige zelf. Daarnaast werd één keer de verwachting uitgesproken dat aandacht aan de voorkant van het proces naar verwachting een langer durend effect zou opleveren. Alhoewel de focus bij de handhavingsmix in principe op de juiste en volledige aangifte ligt, werd één keer ook aangegeven dat aandacht aan de voorkant belangrijk is om procesverstoringen te voorkomen. Vrijwel alle respondenten gaven het belang aan van een opbouw van 'lichte' naar repressieve instrumenten. Dit werd niet alleen onderbouwd vanuit de handhavingsstrategie en de beleidsopdracht, maar ook werd twee keer aangegeven dat men de ondernemer niet onnodig wil belasten.

<u>Deelconclusie interventies en handhavingsmix</u>

Uit bovenstaande bevindingen blijkt dat er van een breed palet aan interventies en toezichtinstrumenten gebruik wordt gemaakt en dat er een grote diversiteit aan redenen is aangegeven om bepaalde interventies in te zetten. Er wordt in de praktijk op verschillende wijze actief gezocht naar mogelijkheden om een handhavingsmix te ontwikkelen in de lijn van het jaarplan 2020 (fouten voorkomen en dienstverlenend zijn naast toezicht en fraude aanpak). In een aantal gevallen betreft dit een specifieke groep, gebaseerd op een specifieke branche of specifiek (fiscaal) probleem. De ruimte voor deze ontwikkeling blijkt in alle gevallen te liggen in instrumenten die niet vanuit de wet zijn ingeregeld en (aanvullend) worden ingezet zoals bellen, handhavingscommunicatie en

samenwerking met diverse partijen, waaronder de ondernemer zelf. Dit heeft ook te maken met de beschikbare capaciteit.

De gedachte en verwachting achter de inzet van de benoemde (mix van) interventies was dat dit een positief effect op naleving zou hebben. Echter het concreet en expliciet benoemen van het beoogde en verwachte effect op één of meerdere compliance factoren bleek lastiger. In twee gevallen betrof de inzet van instrumenten een verplichte werkwijze waarbij wel effect op compliance verwacht werd maar niet expliciet werd benoemd en gemonitord. In verreweg de meeste gevallen werd benoemd dat de inzet van de instrumenten onderdeel uitmaakte van een handhavingsmix die in zijn totaliteit op basis van de handhavingsstrategie was ontwikkeld. De effecten en de meting hiervan worden in de volgende paragraaf verder uitgewerkt en toegelicht.

4.2.3. Effecten

In deze paragraaf worden de verwachte en beoogde effecten per toezichtinstrument of bundeling van toezichtinstrumenten beschreven en aangegeven of deze getoetst zijn. Ook wordt aangegeven wat er volgens de respondenten nodig is om effectmeting te verbeteren en daarmee de effectiviteit van de inzet van toezichtinstrumenten, en de keuze daarin, te vergroten. Ook wordt apart ingegaan op de behoefte aan vastlegging van kengetallen om effectmeting te ondersteunen.

Verwachte en/of beoogde effecten

Bij het beschrijven van effecten wordt uitgegaan van effecten ten aanzien van de gehanteerde mix van toezichtinstrumenten. Daar waar mogelijk worden verwachte en gemeten effecten ten aanzien van de individuele instrumenten vermeld.

In vier gevallen werd aangegeven dat het beoogde en verwachte effect (met name) een meer juiste en volledige aangifte was. Daar waar dat benoemd kon worden was deze verwachting gebaseerd op theoretische kaders, ervaring en/of pilots. Vier keer werd aangegeven het beoogde en verwachte effect van de aanpak op alle drie de compliance factoren betrekking had, waarbij één respondent aangaf dat verwacht werd dat dit effect groter zou zijn bij de meer compliante ondernemers, één respondent aangaf dat het beoogde effect ook het verminderen van procesverstoringen was en één respondent aangaf dat vooral de lichte instrumenten (voorlichten, bellen, handhavingscommunicatie) in principe betrekking hadden op aangifte- en betalingsgedrag en de meer repressieve instrumenten in principe meer ingezet worden op de juistheid en volledigheid van de aangifte. Eén keer werd aangegeven dat het beoogde en verwachte effect op alle vier de compliance factoren betrekking had, dus ook op registratie van belastingplicht. Door één respondent werd aangegeven dat het beoogde effect een hogere gepercipieerde pakkans was. Of dit ook verwacht mocht worden was onbekend. Hier werd geen mogelijkheid

gezien om dit te meten. Eén keer werd gemeld dat de toezichtinstrumenten met name betrekking hadden op tijdige en volledige betaling terwijl een andere respondent juist aangaf dat dit lastiger te beïnvloeden was omdat dit niet alleen met compliance, maar ook met cash flow te maken had. Eén keer werd aangegeven dat de aanpak met name gericht was op betaal- en inlevergedrag. In dit laatste geval werd wel beoogd fouten in de aangifte te voorkomen maar werd niet verwacht dat het effect van een meer juiste en volledige aangifte gemeten kon worden en als zodanig daarmee ook niet expliciet beoogd.

In beeld gebrachte/gemeten effecten
Door vier respondenten werd aangegeven dat het beoogde en/of verwachte effect niet in beeld gebracht kon worden. Door twee respondenten werd aangegeven dat dit in principe wel mogelijk was maar om praktische redenen niet uitvoerbaar was of vrijwel niet uitvoerbaar was in de zin van feitelijk wetenschappelijk onderzoek op langere termijn. Er werden wel mogelijkheden gezien om meer te doen met effectmeting.

In de volgende zes gevallen was het wel mogelijk om effecten te meten. Dit betroffen niet in alle gevallen wetenschappelijke effectmetingen.
- gedragsverandering bij eerstvolgende aangiften na handhavingscommunicatie ten aanzien van *specifieke* onderdelen van de aangifte door hier gericht aandacht voor te vragen. Dit betrof zowel fiscale items als verhoudingsgetallen.
- wetenschappelijke aanpak met nulmeting, controlegroep en check op de juistheid en volledigheid van de aangifte na de interventie (betrof het effect van een *specifieke* handhavingsbrief).
- bij de verzuimaanpak is gebleken dat het effect van het sturen van brieven en bellen tijdelijk is. Stopt de aanpak, dan loopt het verzuim weer op.
- de kwaliteit van de aangiftes onder horizontaal toezicht bleken significant beter.
- meer vooroverleg (afstemming vooraf over fiscale duiding) onder horizontaal toezicht.
- groepsgewijze benadering bij nalevingstekort met nulmeting ten aanzien van correcties, verzuim en betalingsachterstand met monitoring op verschillende momenten na interventies. Dit betrof het hele palet van voorlichting tot en met boekenonderzoeken. De effecten van de diverse instrumenten gingen op een gegeven moment samenlopen en waren niet meer individueel te meten. Effecten bleken ook uit gesprekken met ondernemers.
- behoorlijk effect bij uitbrengen hernieuwd bevel tot betalen.
- rechtvaardiging om bij laag belang en laag risico lichte instrumenten in te zetten (gerelateerd aan uitkomsten in vergelijking met de hele groep die een gelijke behandeling onderging).

- significante verbetering in betaal- en aangiftegedrag bij integrale beoordeling en *specifiek en bewust* gekozen behandeling op basis van beoordeling, in te zetten instrument en expertise, waarbij begonnen werd met inzet van een licht instrument, eventueel gevolgd door een zwaarder instrument.

Voorwaarden voor verbetering effectmeting
Om effecten beter in beeld te kunnen brengen werden de volgende voorwaarden genoemd:
- inzet van instrumenten bewust kiezen volgens handhavingsregie;
- politieke dekking, commitment van de directie tot teamleider, bewustzijn van belang bij medewerker;
- goede doorvertaling van theorie naar de werkvloer;
- beschikbaar stellen van instrumentarium en mogelijkheden om dit vast te leggen
- wetenschappelijke afdeling om effectmeting te ondersteunen en klantonderzoek te kunnen doen (twee keer);
- één keer werd aangegeven dat effectmeting een onjuiste uitkomst aangaf omdat interpretatie door 'eigen' mensen had plaatsgevonden. Dit pleit voor het loskoppelen van vastlegging en effectmeting.
- meer doen met de uitkomsten van de steekproef ondernemingen;
- technisch mogelijk maken om kengetallen vast te leggen / meer vastleggen (drie keer), waarbij één keer werd aangegeven dat het dan om querybare velden moet gaan, zodat data-analyse mogelijk is. Het gaat dan om bijvoorbeeld om unieke codering, zodat interventies gekoppeld kunnen worden aan individuele belastingplichtigen, zoals nu het geval is bij bijvoorbeeld boekenonderzoeken. Dit laatste aspect werd twee keer benoemd;
- langdurige metingen voor beoordelen blijvend effect (twee keer);
- opbouwen en behoud van expertise (kennis en ervaring);
- tussendoelen benoemen;
- causaal verband vaststellen.

Onverwachte effecten
Tot slot werden er nog onverwachte effecten genoemd zoals positieve feedback op het contact wat er geweest was, positief effect ten aanzien van de samenwerking intern of extern, onverwachte feedback over informatie die moeilijk op de site van de Belastingdienst te vinden was, grotere bewustheid bij de leiding van het bedrijf over wat er in hun onderneming speelde, maar ook voorbeelden van ingrijpend effect op het leven van de betreffende ondernemer als gevolg van ingezette interventies.

Vastleggingen

Een belangrijk aspect bij effectmeting is het kunnen beschikken over relevante data. Tien respondenten geven aan dat er op dit moment onvoldoende structurele vastlegging van kengetallen plaatsvindt om effectmeting goed vorm te kunnen geven. Op niet-repressieve instrumenten kan niet gepland worden en zijn er beperktere vastleggingsmogelijkheden. Eén respondent gaf aan dat de data er wel zijn, maar dat dit geen querybare velden betreft. Zes respondenten geven aan dat de effectmetingen die zij gedaan hebben zijn gebaseerd op extracomptabele vastleggingen. In de literatuur vinden het belang van vastleggingen in bestaande databases / informatiesystemen terug (Winter & De Ridder, 2010) (Stevens, Pheiffer, Van den Broek, Keijzer, & Van der Hel-van Dijk, 2012). Daarnaast is de aandacht voor het vastleggen van gegevens buiten bestaande databases toegenomen en ligt dit vanwege de AVG extra gevoelig. Dit wordt ook expliciet door twee respondenten benoemd.

Deelconclusie effectmeting en vastleggingen

Uit bovenstaande weergave van de bevindingen blijkt dat het mogelijk is om effecten te meten als de inzet van interventies bewust heeft plaatsgevonden met dat oogmerk. In dat geval zijn de voorwaarden geschapen door een nulmeting te doen, een controle groep te benoemen, specifieke items te benoemen waarop effecten werden beoogd en (extracomptabel) resultaten bijgehouden en is soms ook op een later moment gemonitord of dit een blijvend effect tot gevolg heeft gehad. Het betreft echter incidentele effectmetingen, mede ingegeven door de 'drive' van de betrokken zelf. Van strategische effectmeting, zoals beoogd in de Handhavingsbrief (Ministerie van Financiën, 2017, p. 13) op basis van handhavingsregie om de cyclus in de keuze van de meest effectieve inzet van (de mix van) toezichtinstrumenten of andere interventies te verbeteren is (nog) geen sprake. Er wordt dan ook nog niet voldaan aan de genoemde voorwaarden om de effectmeting op strategisch niveau vorm te geven.

4.3. Evaluatie

In dit hoofdstuk wordt de toetsing van de werkwerkhypotheses weergegeven. Ook worden extra observaties vermeld die uit het onderzoek naar voren zijn gekomen.

In onderstaande tabel is de beantwoording door de respondenten van de voorgelegde werkhypotheses aan het eind van de interviews in aantallen opgenomen.

Werk-hypothese	Volledig mee eens	Mee eens	Gedeeltelijk mee eens	Gedeeltelijk mee oneens	Mee oneens	Geen uitspraak
1	1	11	0	0	0	0
2	0	9	2	0	0	1

3	0	6	2	1	3	0
4	3	9	0	0	0	0
5	0	9	1	1	1	0
6	0	6	3	0	2	1
7	2	8	2	0	0	0
8	5	7	0	0	0	0

Tabel 5. Overzicht beantwoording werkhypotheses door respondenten

Bij het voorleggen van de werkhypotheses bleek duidelijk een onderscheid tussen de werkhypotheses waar de respondenten het (absoluut) mee eens waren en werkhypotheses waar door veel respondenten nuanceringen op werden aangebracht. Zo was iedereen positief over werking van doelgerichte voorlichting (WH1) en een goede afweging van de aard en omvang van de inzet van toezichtinstrumenten (WH7) en was men (sterk) overtuigd van de waarde van een bewust gekozen mix aan toezichtinstrumenten (WH4) en het belang en de noodzaak van structurele vastleggingen (WH8). Gevraagd was te antwoorden met ja, nee of een nuancering aan te brengen indien gewenst. In de beantwoording werd voor bovenstaande vier werkhypotheses daarnaast de ruimte genomen om te antwoorden met 'absoluut!' en 'zeker!'. Deze antwoorden zijn in de overzichten vertaald naar 'volledig mee eens'. Daarnaast zijn ja en nee vertaald naar 'mee eens' en 'mee oneens'. Tot slot is de beantwoording met ja met nuancering opgenomen met 'gedeeltelijk mee eens' en is nee met nuancering in de overzichten opgenomen als 'gedeeltelijk mee oneens'. In de beoordeling van de werkhypotheses is de inhoud van die nuanceringen vermeld. Het betrof hier met name de werkhypotheses over het nut van de intensiteit van het toezicht (WH3) en in mindere mate de werkhypotheses over de positieve bijdrage van handhaving/correctie (WH5) en sanctie (WH6) aan compliance. In een enkel geval kon geen antwoord worden gegeven op de betreffende werkhypothese. In dat geval is 'geen uitspraak' gebruikt.

Zoals in hoofdstuk 3 bij methodiek is aangegeven zal op basis van de bevindingen van het onderzoek bij elke werkhypothese aangegeven worden of deze aannemelijk is, getoetst is door een pilot, getoetst is door breed onderzoek of bewezen is met gebruik making van gestructureerd vastgelegde data.

WH1: Naarmate de voorlichting specifieker op de doelgroep wordt gericht, neemt de compliance toe.

Werkhypothese 1 werd door alle respondenten bevestigend beantwoord, ook al had maar de helft van alle respondenten dit instrument zelf ingezet. In dat geval was sprake van

voorlichting gericht op specifieke doelgroepen (starters, fiscaal dienstverleners en een specifieke branche-aanpak) met als doel het informeren en kennis te laten vergaren ten aanzien van de administratie en fiscaliteit. Uit de literatuur is bekend dat kennis een belangrijke factor is die bijdraagt aan compliance (Goel, Ravindra, Singh, & Sharma, 2014) (Sellens, et al., 2018). Op basis van de uitkomsten van de interviews en de antwoorden hiervan kan geconcludeerd worden dat werkhypothese 1 in de praktijk herkend en onderschreven wordt en daarmee aannemelijk is. Er heeft geen specifieke effectmeting plaatsgevonden waardoor de stelling bewezen kan worden geacht.

WH2: Communicatie in de vorm van gerechtvaardigde feedback draagt bij aan een hogere compliance.

Elf van de twaalf respondenten deden een uitspraak over deze werkhypothese. Alhoewel op twee gedeeltelijke bevestigingen na het antwoord op deze werkhypothese bevestigend was, kan geen uitspraak gedaan worden over enige mate van toetsing in de praktijk. Tijdens de interviews konden de respondenten wel voorbeelden geven van wat zij als feedback zagen, bijvoorbeeld bij bellen bij verzuim of het waarschuwen van het verlopen van een betaaltermijn, maar dit betrof in alle gevallen meer een toepassing hiervan tijdens de inzet van andere toezichtinstrumenten en niet het bewust inzetten hiervan als zelfstandig toezichtinstrument. Aanvullend werd door één van de respondenten nog aangegeven wel te geloven in een positieve invloed op compliance indien sprake was van ondersteunende feedback of een kleine 'tik op de vingers' ten behoeve van bewustwording, maar niet van repressieve feedback. Doordat geen bewuste inzet van feedback heeft plaatsgevonden en deze dus ook niet gemeten is, kan deze stelling niet bewezen worden geacht.

WH3: Een hoge intensiteit van toezicht (rakingspercentage) van onder toezicht gestelden door toezichtinstrumenten verhoogt de compliance.

Werkhypothese drie leverde verschillende reacties op, zowel tijdens de interviews als tijdens het beoordelen van de werkhypothesen. Drie respondenten beoordeelden de werkhypothese met 'nee', één respondent was het er gedeeltelijk mee oneens, twee waren het gedeeltelijk eens en zes bevestigden de werkhypothese. Tien van de twaalf respondenten hadden de behoefte hun beoordeling toe te lichten. Dit betrof alle vier de respondenten met een (gedeeltelijk) ontkennende reactie en vier met een (gedeeltelijke) bevestigende reactie. De kern van de toelichtingen betrof de mening dat hogere raking van repressieve instrumenten de compliance niet verhoogt, zeker niet als dat wordt toegepast bij compliante belastingplichtigen. Bij horizontaal toezicht wordt zelfs bewust een lage raking door repressieve instrumenten beoogd, omdat bereidwillige naleving geen hoog rakingspercentage nodig heeft. Door een andere respondent werd dit

verwoord met raking die een ondernemer verdient. Een hoog rakingspercentage zou mogelijk meer het karakter krijgen van afgedwongen compliance dan van compliance door eigen bereidheid hiertoe, ook al verhoogt dit de gepercipieerde pakkans. Twee respondenten gaven aan dat een dergelijk afgedwongen compliance dan ook maar tijdelijk zou zijn en één daarvan legde ook de link met (tijdelijke) extrinsieke motivatie. Voor een positief effect op compliance werd de samenhang van de instrumenten (de handhavingsmix) belangrijker geacht van de intensiteit van het toezicht.
Dit betekent dat werkhypothese in deze vorm verworpen moet worden.

WH4: Hoe specifieker de keuze van (de mix van) toezichtinstrumenten op de diverse doelgroepen is afgestemd, hoe groter de effectiviteit hiervan op (het behoud van) compliance bij de betreffende doelgroepen.

Werkhypothese vier werd door alle twaalf de respondenten bevestigend beantwoord, waarvan door drie met een overtuigend 'Ja'. Uit de uitkomsten van de interviews is ook op te maken dat daar waar een specifieke mix van toezichtinstrumenten doelbewust op een specifiek groep is toegepast, het ook mogelijk was om effecten te meten. Eén respondent maakte daarbij melding van een gerichte groepsgewijze bespreking van een nalevingstekort, waarbij de groep daarna op verschillende manieren benaderd en gemonitord werd. Dit had een positief effect op zowel de juistheid en volledigheid van de aangifte als op het aangifte- en betaalgedrag. De ondernemers gaven ook aan dat zij de Belastingdienst nu serieuzer namen doordat zij zagen dat er meer zicht was op wat er speelde in hun ondernemingen. In een ander voorbeeld werd bewust ingezet op een brede, integrale en opeenvolgende inzet van instrumenten, die gevolgd werd door een effectmeting op aangifte- en betaalgedrag. Ook daar werd een significante verbetering waargenomen. Alhoewel de genoemde effectmetingen in een pilot omgeving hebben plaatsgevonden met behulp van extracomptabele vastleggingen en slechts over relatief korte tijd gemeten kon worden, kan deze werkhypothese bevestigend beantwoord worden.

WH5: Hoe beter handhaving (correctie van gedrag) is afgestemd op een bepaalde doelgroep, hoe groter het effect op compliance.

Tien van de twaalf respondenten hebben bevestigend geantwoord op deze werkhypothese. In vier gevallen werd een nadere toelichting gegeven. Zo werd door één respondent aangegeven dat correcties ook averechts kunnen werken, zeker als dit ook met een boete gepaard gaat. Een andere respondent gaf aan dat het ook negatief zou

kunnen uitwerken als de belastingplichtige het gevoel zou hebben dat zij in beeld was doordat de groep teveel gevolgd werd. Dit houdt ook verband met de reactie van een respondent die het eens was met de werkhypothese als de doelgroep goed gedefinieerd was. Daarnaast gaf één respondent aan dat het positief op de compliance zou kunnen werken als de belastingplichtige zelf zijn fout zou mogen herstellen. Omdat deze werkhypothese in de praktijk niet getoetst is en de invulling van de bepaalde doelgroep teveel ruimte laat voor interpretatie, moet de werkhypothese in deze vorm verworpen worden.

WH6: Hoe specifieker een sanctie wordt afgestemd op de betreffende overtreding van wet- en regelgeving, hoe beter het effect op compliance.

Deze werkhypothese werd door zes van de twaalf respondenten bevestigend beantwoord en gaven drie respondenten aan het hier gedeeltelijk mee eens te zijn. Twee respondenten hebben ontkennend geantwoord en één respondent kon hier geen uitspraak over doen. De respondenten die het niet eens zijn met de werkhypothese onderbouwen dit door aan te geven dat het effect van de sanctie afhangt van hoe de samenwerkende partners hierop inspelen danwel met het standpunt dat sancties sowieso een negatief effect hebben op compliance. In het laatste geval zou het publiceren ervan weer wel een afschrikwekkend effect kunnen hebben op anderen en daar dus een positief effect op compliance kunnen hebben. Het voorbehoud van de respondenten die de werkhypothese bevestigden zag in alle gevallen op het woordje 'specifiek'. In algemene zin werd een boete niet gezien als instrument om compliance te verbeteren, maar juist als incentive voor non-compliance. Door de toevoeging dat rekening werd gehouden met omstandigheden en er maatwerk geleverd kan worden kon de werkhypothese bevestigd worden. Tijdens de interviews werden dan ook diverse voorbeelden van dit maatwerk gegeven. Deze twee kanten van het effect van een sanctie zijn ook in het theoretisch model opgenomen. De uitkomsten van dit onderzoek bevestigen dat. De werkhypothese kan daarmee, met de nodige toelichting, bevestigend beantwoord worden.

WH7: Hoe beter de afweging tussen de aard en omvang van de inzet van toezichtinstrumenten hoe doeltreffender de beoogde impact op compliance is.

Alle respondenten hebben bevestigend op deze werkhypothese gereageerd. Twee respondenten waren daarvan sterk overtuigd, twee anderen waren het gedeeltelijk met de werkhypothese eens. Dit sluit aan met de uitgebreide beantwoording van de bijbehorende vraag uit het interview. Vrijwel alle correspondenten hebben beschreven hoe zij, onderbouwd, bewust keuzes hebben gemaakt ten aanzien van de mix aan

toezichtinstrumenten en ook nog steeds bezig waren deze mix te verbeteren. De effectmetingen vonden in alle gevallen plaats in de zoektocht naar de optimale afweging van deze mix en hadden allemaal betrekking op instrumenten aan de voorkant van het toezicht, op de niet-repressieve instrumenten. Toch hadden ook de repressieve instrumenten een bewuste, maar wel meer bescheiden, plek in een aantal benoemde aanpakken. Hiermee blijkt juist deze werkhypothese de kern van het huidige toezicht en kan deze bevestigend beantwoord worden.

WH8: Structurele vastlegging van kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten op het moment dat de gebeurtenis zich voordoet verhoogt de betrouwbaarheid van de effectmeting op compliance.

Alle respondenten hebben deze werkhypothese bevestigend beantwoord, waarvan vijf met overtuiging. Uit de interviews is gebleken dat structurele vastlegging ook als voorwaarde wordt gezien voor (langdurige) effectmeting. Aangegeven werd dat mede hierdoor effectmeting ook meer betrouwbaar werd, zeker als deze daarnaast door deskundigen zou worden uitgevoerd en/of ondersteund. Door de respondenten zijn twee voorbeelden benoemd waarbij vraagtekens gezet konden worden bij de juiste interpretatie van de resultaten. Ook kon, doordat vaak gebruik gemaakt moest worden van extracomptabele vastlegging, wel de impact gemeten worden, maar niet het langer durende effect. Vastlegging in querybare velden zou beide problemen op kunnen lossen. Tot slot werd nog aangegeven dat de (extra) vastlegging moest dienen ten behoeve van de interne leercirkel en niet als politiek instrument. De werkhypothese kan daarmee bevestigend beantwoord worden.

4.4. Samenvatting onderzoek

In deze paragraaf wordt een samenvatting van de uitkomsten van het onderzoek gegeven en wordt een beslistabel gepresenteerd.

Uit het onderzoek is gebleken dat aan de voorwaarden om effectmeting uit te kunnen voeren niet in alle gevallen wordt voldaan. Het is dan ook de vraag of er daadwerkelijk een effect is gemeten of dat er een resultaat is gemeten (zie bijvoorbeeld het tijdelijke 'effect' bij de verzuimaanpak). Het is daarom belangrijk om op dit moment nog nader in te gaan op het verschil tussen resultaat (output), impact (korte termijn) en effect (impact op lange termijn).

Als we de aangegeven redenen voor inzet van een toezichtinstrument plaatsen in het kader van effectmeting, dan is het nodig om onderscheid te maken tussen beoogde

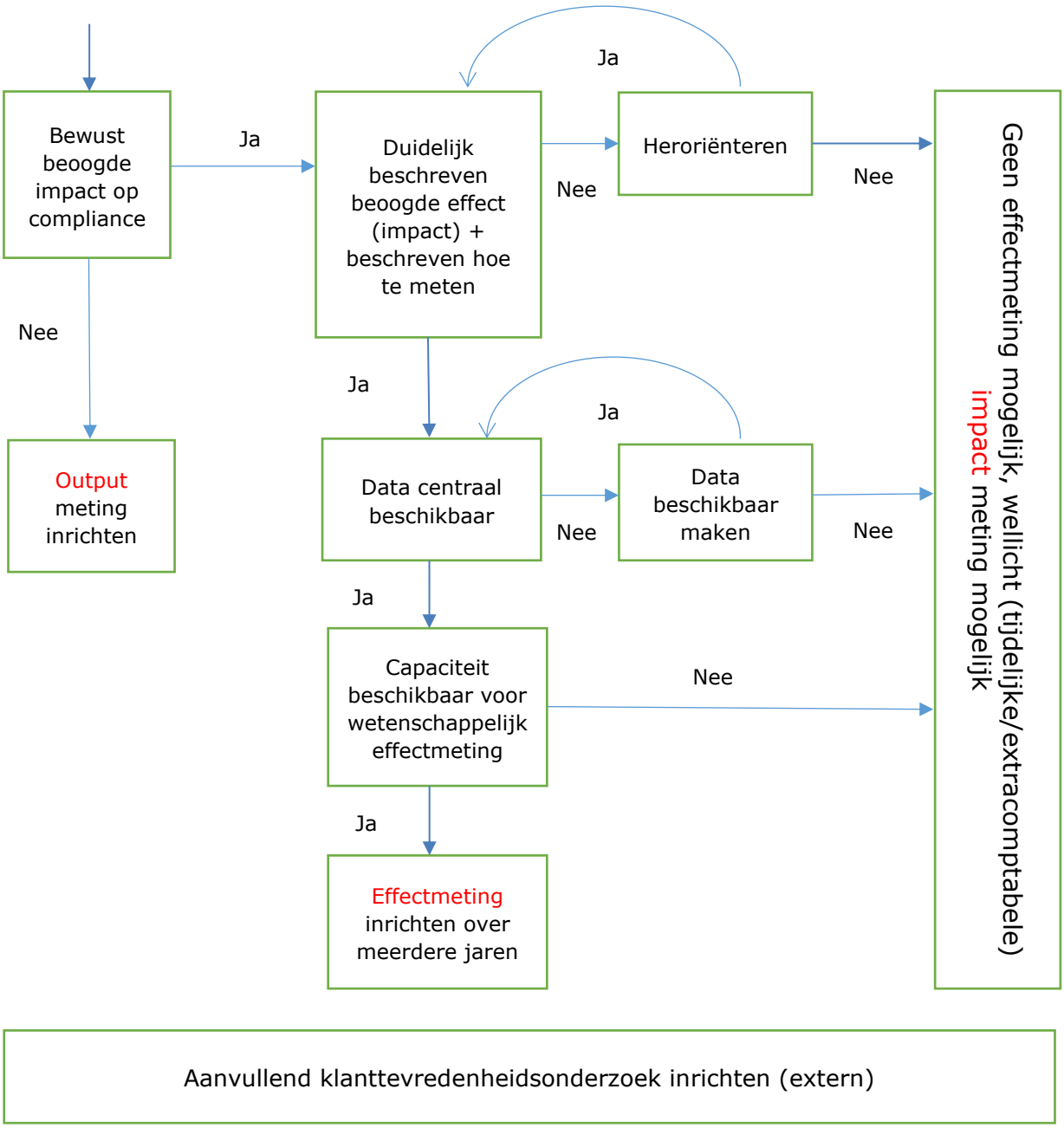
output, impact en effect (Desmidt & Heene, 2018, p. 426). Dit onderzoek richt zich op effectmeting (dus op langere termijn) van de toezichtinstrumenten en de invloed die dit heeft, of zou kunnen hebben, op de keuze van de inzet van (de mix) van toezichtinstrumenten. Uiteraard is het ook mogelijk om effecten te meten als dat niet van invloed is op de keuze van de inzet van de instrumenten, omdat dit een gegeven is vanuit procesbelang of opbrengstdoelstelling, maar dat valt buiten de scope van dit onderzoek.

Het kan zijn dat één toezichtinstrument om verschillende redenen wordt ingezet. Dit is van belang om de zinvolheid van effectmeting te kunnen beoordelen. Indien bijvoorbeeld een boekenonderzoek wordt ingezet als onderdeel van een afgewogen handhavingsmix vanuit een handhavingsstrategie, dan wordt effectmeting als zeer gewenst en relevant beschreven. Als een boekenonderzoek ingezet wordt om correcties te generen of sec aantallen te halen, dan lijkt effectmeting niet relevant voor de inzet van dit instrument, mogelijk ook niet als het effect negatief zou blijken te zijn. Er wordt immers geen vervolg gegeven aan de uitkomsten. Niet het instrument zelf, maar de reden van inzetten hiervan is derhalve van belang voor effectmeting. Als de reden van inzet van instrumenten puur gericht is op het genereren van producten, dan zou outputmeting een logischer keuze zijn, wat in het verleden ook gebeurd is. Uiteraard heeft ook dit toezicht effect op de compliance, maar is daar niet initieel op gebaseerd en is ook niet van invloed op de keuze van de inzet van de toezichtinstrumenten. Het is daarom een voorwaarde voor effectmeting dat deze voorafgegaan wordt door bewust en specifiek keuze, zoals ook beschreven is in de situatie dat effectmeting gerealiseerd is kunnen worden.

Ook is het van belang op dit moment nog de link te leggen tussen extrinsieke motivatie waarmee een resultaat kan worden bereikt en het omzetten van extrinsieke motivatie naar intrinsieke motivatie waarmee een langduriger effect kan worden bereikt (Deci & Ryan, 2000). Deze bewustheid zou ook een plek moeten krijgen in de beoogde aanpak en de keuze van het inzetten van de diverse instrumenten.

Om de link tussen bewuste inzet van toezichtinstrumenten met beoogd effect op compliance en de bijbehorende (effect)meting en monitoring te visualiseren is aan het eind van deze paragraaf een beslistabel opgenomen (tabel 1). Deze beslistabel maakt inzichtelijk in welke situatie investering in effectmeting naar verwachting het meeste rendement oplevert ten aanzien van te maken keuzes in het toezicht. Hierin is ook het belang van structureel beschikbare data en beschikbaarheid van de nodige deskundigheid opgenomen.

Inzet toezichtinstrument



Figuur 1. Beslistabel effectmeting

5. Conclusie, discussie en aanbevelingen

In dit hoofdstuk worden de conclusies van het onderzoek beschreven en wordt de hoofdvraag beantwoord. Daarnaast wordt ook de beperking van het onderzoek aangegeven. Tot slot worden aanbevelingen gedaan die de beoogde doel van dit onderzoek, het vergroten van de effectiviteit van (de mix van) toezichtinstrumenten, dichterbij zou kunnen brengen.

5.1. Conclusies

Aan het begin van dit onderzoek is onderstaande centrale vraag geformuleerd.

"Welke toezichtinstrumenten zet de Belastingdienst MKB in ter bevordering van de compliance van burgers en bedrijven, hoe wordt deze inzet gemotiveerd en wat zijn de mogelijkheden om de effectiviteit van (de mix van) toezichtinstrumenten te vergroten?"

Door de respondenten is tijdens de interviews een overzicht gegeven van de door hen ingezette toezichtinstrumenten c.q. de instrumenten waar zij zicht op hadden. Niet alle instrumenten blijken te zijn ingezet met als oogmerk de compliance te verbeteren. Soms zijn hier andere reden voor gegeven, die onder andere liggen op het gebied van het goed uitvoeren van processen of gewoon als primaire (output) doelstelling. Ook komt het voor dat één instrument voor verschillende doeleinden wordt ingezet. Hier kan compliance er één van zijn. De motivatie voor inzet van de toezichtinstrumenten varieert daarmee. De effectiviteit van de inzet van (de mix van) toezichtinstrumenten kan vergroot worden door meer bewustheid binnen dit keuzeproces te creëren. Dit heeft enerzijds te maken met het bewust(er) benoemen van de beoogde doelstellen en anderzijds met de beperkte mogelijkheid om bij de onderbouwing en beoordeling van de gemaakte keuzes ten aanzien van de ingezette toezichtinstrumenten gebruik te kunnen maken van (effect)meting. Daarnaast kan de effectiviteit worden verhoogd door (nog) meer gebruik te maken van de lichtere instrumenten aan de 'voorkant' van het toezicht, zoals voorlichting, bellen, handhavingscommunicatie en samenwerking met fiscaal dienstverleners, omdat deze instrumenten meer meetbaar effectief zijn gebleken in de casussen die zijn benoemd.

Ten aanzien van de effectmeting kan hier nog een verdieping op worden aangebracht. Uit het onderzoek blijkt dat van structurele effectmeting (nog) geen sprake is. Daar waar (wetenschappelijke) effectmeting wel aan de orde was, betreft het duidelijk afgebakende pilots met beperkte interventies. Dit geeft aan dat effectmeting wel mogelijk is. Als het

doel is om de inzet van interventies te koppelen aan de keuze voor de inzet van toezichtinstrumenten ten behoeve van compliance, dan moet er nog veel gebeuren.

Voorwaarden voor het verder inrichten van effectmeting is in eerste instantie commitment van de hele organisatie, en wellicht van de politiek, om de gekozen strategie (inzet van (een mix van) toezichtinstrumenten) ook daadwerkelijk te laten ondersteunen door effectmeting. Ten tweede is het van belang om meer bewustheid te creëren over beoogde doelen. Het maakt verschil of met de beoogde inzet van een interventie output (bijvoorbeeld beheersing van een proces / producten) of outcome (in casu compliance) wordt beoogd. Tot slot is het van belang of bij de outcome alleen de impact op korte termijn of het effect op lange termijn beoogd is bij de keuze van de interventie. Niet in alle gevallen zal de inzet van toezichtinstrumenten gekoppeld kunnen worden aan een effect op compliance ten aanzien van het nalevingsgedrag. Effectmeting is ook geen doel op zich. Als maatschappelijke verantwoording kan worden afgelegd door een proces goed in te richten en output te meten dan is het goed dit ook zo te benoemen. Als mensen en middelen worden ingezet om de naleving van wet- en regelgeving te verhogen dan heeft effectmeting toegevoegde waarde. Het is dus belangrijk dit onderscheid te maken, gemaakte keuzes goed te onderbouwen en het beoogde doel te meten.

Om dit proces te ondersteunen is een goede samenwerking tussen de diverse afdelingen en stakeholders nodig om samen een gemeenschappelijke taal en aanpak te ontwikkelen. Een afdeling onderzoek om de kwaliteit van dit proces te borgen is een absolute voorwaarde.

Tot slot is het van belang om een goede basis te hebben in de vastleggingen van *alle* interventies om überhaupt materiaal te hebben als basis voor effectmeting en niet afhankelijk te zijn van extracomptabele vastleggingen.

5.2. Discussie

Het doel van dit onderzoek was het in beeld brengen van de effectiviteit van de diverse toezichtinstrumenten op compliance om daarmee aanbevelingen te kunnen doen om de keuze van de toezichtmix te verbeteren. Door de breedte van de opzet van het onderzoek is de diepgang beperkt gebleven en heeft het onderzoek alleen die effectmetingen in beeld gebracht die door de respondenten is aangegeven. Of de kwaliteit van de uitvoering onderdeel uitmaakt van het in beeld gebrachte effect is geen onderdeel van het onderzoek geweest. Het onderzoek heeft geleid tot meer inzicht in de

voorwaarden die van belang zijn om op termijn een vollediger beeld van de effectiviteit van de ingezette (mix van) toezichtinstrumenten te kunnen krijgen. De voorwaarde van de beschikbaarheid van digitale data houdt verband met de uitvoering van kwantitatief onderzoek, met name van de zogenaamde 'niet-traditionele' instrumenten. Een kwantitatief onderzoek, hoe interessant ook, was op dit moment niet mogelijk. Wetenschappelijke effectmeting heeft maar zeer beperkt plaatsgevonden. Alhoewel het aantal bevraagde respondenten klein was, geeft dit wel een beeld van hun realiteit. Een groter aantal respondenten zou uiteraard tot een vollediger beeld en steviger uitspraken hebben geleid. Dit geldt ook voor een kwantitatief onderzoek, wat in het kader van dit onderzoek niet mogelijk was.

Tot slot is dit onderzoek volledig intern gericht geweest. Een waardevol inzicht is dat de omvang en het belang van het vragen van feedback aan belastingplichtigen en belastingschuldigen wel eens een groter zicht kan geven op de effectiviteit van het toezicht dan dat interne effectmeting ooit zal kunnen hebben. Immers, wat één van de respondenten terecht aangaf:

"Effecten zitten altijd in de doelgroep, niet in de organisatie zelf"

Terugkijkend heeft de breedte van de opzet van dit onderzoek ruimte gegeven aan de respondenten om te benoemen wat zij belangrijk vonden ten aanzien van de keuzes in het toezicht en wat hun ervaringen waren ten aanzien van de mogelijkheden en beperkingen van effectmeting. Dit heeft geleid tot een rijk beeld aan praktijkbeschrijvingen op basis waarvan de analyse heeft plaatsgevonden. Deze brede opzet heeft er echter mede toe geleid dat een aantal werkhypotheses achteraf onvoldoende beoordeeld konden worden. Dit hield ook verband met de verschillende interpretaties van begrippen zoals doelgroep, toezicht en toezichtinstrumenten, maar ook met de invulling van de gewenste intensiteit van toezicht (raking). Voor het begrijpen van de respondenten heeft het een voordeel in de organisatie bekend te zijn. Het is echter ook een nadeel in de focus tijdens de interviews en in de voorbereiding van de vragen en de formulering van de werkhypotheses. Het spreken van 'dezelfde' taal bleek tijdens de interviews namelijk een grotere diversiteit te kennen dan verwacht. Een grotere bewustheid van verschil in interpretatie van begrippen tijdens de opzet van het onderzoek had kunnen resulteren in een andere vraagstelling tijdens de interviews en mogelijk tot een betere uitspraak over de werkhypotheses.

5.3. Aanbevelingen

In deze paragraaf zijn de aanbevelingen opgenomen die binnen de Belastingdienst haalbaar zouden moeten zijn om meer zicht op en sturing aan het verhogen van de effectiviteit door inzet van (een mix) van toezichtinstrumenten te kunnen geven.

Aanbeveling 1
Richt digitale systemen zodanig in dat deze geschikt zijn om als basis te dienen voor de koppeling van interventies aan het effect op compliance, zonder dat extracomptabele vastlegging nodig is.

Aanbeveling 2
Maak de keuze om een wetenschappelijke afdeling onderzoek zodanig te positioneren en vorm te geven dat deze in staat is partijen te adviseren over nut en noodzaak van (effect)metingen (bewuste keuzes), in de uitvoering daarvan te ondersteunen en verantwoordelijkheid te nemen voor de kwaliteit en daarmee de beleidsstrategie een nog meer wetenschappelijke invulling te kunnen geven. Deze afdeling kan daarmee als kenniscentrum kunnen fungeren vanuit het overzicht wat zij heeft op bestaande metingen. Deze afdeling zou daarnaast ingezet kunnen worden op systematische evaluaties, zoals aanbevolen in de Beleidsdoorlichting 2017.

Aanbeveling 3
Zet meer in op klanttevredenheidsonderzoek om waardevolle feedback op te halen bij belastingplichtigen en/of belastingschuldigen om hiermee de uitvoering van het toezicht beter vorm te kunnen geven ten aanzien van de juiste inzet van de interventies en de menselijke maat hierin. Het is niet voor niets dat een groot deel van de respondenten verwees naar het belang van samenwerking met externe (overheids)partners, omdat hier onder andere waardevolle feedback vandaan kwam over waar nalevingstekorten veroorzaakt werden door kennisgebrek. De huidige aandacht voor de menselijke maat maakt het daarnaast belangrijk om ook feedback van de belastingplichtige/ belastingschuldige te krijgen om een goede keuze te kunnen maken tussen preventief of meer repressief toezicht. Uit de herijkte indicatoren uit de begroting voor 2021 blijkt ook het belang van het oordeel van de burger over het handelen van de Belastingdienst ten aanzien van adequate behandeling, informatieverstrekking, gemak en corrigerend optreden. Dit zou mede vorm gegeven kunnen worden door een bredere inzet van een vorm van klanttevredenheidsonderzoek.

Literatuur- en bronnenlijst

Algemene Rekenkamer. (2016, november 30). *Handhavingsbeleid Belastingdienst.* Opgehaald van rekenkamer.nl: https://www.rekenkamer.nl/publicaties/rapporten/2016/11/30/handhavingsbeleid-belastingdienst

Allingham, M., & Sandmo, A. (1972). Income tax evasion: a theoretical analyses. *Journal of Public Economics 1(3-4)*, 323-338.

Belastingdienst. (2017). *Handhavingsstrategie Belastingdienst.* Opgehaald van Belastingdienst.nl: https://over-ons.belastingdienst.nl/wp-content/uploads/2017/02/Handhavingsstrategie-Belastingdienst.pdf

Belastingdienst. (2020, januari 23). Opgehaald van belastingdienst.nl: https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/themaoverstijgend/brochures_en_publicaties/toezichtsplan-arbeidsrelaties-2020

Bordewijk, P., & Klaassen, H. (2011). *Begroten met beleid: pleidooi voor planning en control op maat bij de gemeenten.* Den Haag: Sdu.

Deci, E., & Ryan, M. (2000). The "What"en "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behaviour. *Psychological Inquiry 11 (4)*, 227-268.

Desmidt, S., & Heene, A. (2018). *Strategisch Management.* Leuven: LannooCampus.

Etienne, J. (2011). Compliance Theory: A goal framing approach. *Law & Policy 33 (3)*, 305-333.

Feys, M., Libbrecht, N., Anseel, F., & Lievens, F. (2008). Relatiekwaliteit als moderator in de relatie tussen procedurele rechtvaardigheid en reacties op feedback in prestatiebeoordeling. *Gedrag en Organisatie 21 (4)*, 430-450.

Goel, S., Ravindra, K., Singh, R., & Sharma, D. (2014). Effective smoke-free policies in achieving a high level of compliance with smoke-free law: experiences from a district of North India. *Tobacco Control (23)*, 291-294.

Goslinga, S., Siglé, M., & Van der Hel, L. (2016). Het meten van effecten van de handhaving door de Belastingdienst. *Tijdschrift voor Toezicht 3*.

Klaassen, H., Maks, H., & Anderson, R. (2010). *Efficiëntie en effectiviteit van de publieke sector in de weegschaal, hoofdstuk 1.* Studiedienst van de Vlaamse Regering 2010, 2.

Locke, E., & Latham, G. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation. *American Psychologist 57(9)*, 705-717.

Ministerie van Financiën. (2016). *Rijksbegroting 2017*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2017/voorbereiding/begroting

Ministerie van Financiën. (2017, oktober). *Beleidsdoorlichting 2017.* Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/system/files/333/17ix01-fin-rapport-beleidsdoorlichting-toezicht-en-opsporing-en-massale-processen-bdienst.pdf

Ministerie van Financiën. (2017, december 7). *Handhavingsbrief 2018.* Opgehaald van tweedekamer.nl: https://www.tweedekamer.nl/kamerstukken/brieven_regering/detail?id=2017Z17421&did=2017D36175

Ministerie van Financiën. (2017). *Rijksbegroting 2018*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2018/voorbereiding/begroting

Ministerie van Financiën. (2017, mei 17). *Rijksjaarverslag 2016*. Opgehaald van rijksoverheid.nl: hhttps://www.rijksoverheid.nl/documenten/jaarverslagen/2017/05/17/financien-en-nationale-schuld-2016

Ministerie van Financiën. (2018). *Rijksbegroting 2019*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2019/voorbereiding/begroting

Ministerie van Financiën. (2019, november 20). *Jaarplan 2020 Belastingdienst.* Opgehaald van rijksoverheid.nl: https://www.rijksoverheid.nl/documenten/publicaties/2019/11/20/jaarplan-2020-belastingdienst

Ministerie van Financiën. (2019). *Rijksbegroting 2020*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2020/voorbereiding/begroting

Ministerie van Financiën. (2019). *Rijksbegroting 2020*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2020/voorbereiding/begroting

Ministerie van Financiën. (2020). *Rijksbegroting 2021*. Opgehaald van rijksbegroting.nl: https://www.rijksbegroting.nl/2021/voorbereiding/begroting

Reich, J., Goodstein, M., Callahan, S., Callahan, K., Crossley, L., Doron, S., . . . Nasraway Jr, S. (2015). Physician report cards and ranking yield long-lasting hand hygien compliance exceeding 90%. *Critical Care 19*, 292-298.

Rijksoverheid. (2020, april 26). *Rijksbegroting - beleidsevaluaties.* Opgehaald van Rijksbegroting.nl: https://www.rijksbegroting.nl/beleidsevaluaties/evaluaties-en-beleidsdoorlichtingen/handreiking-beleidsdoorlichtingen/het-meten-van-doeltreffendheid

Samadi, M. (2016). Policing the police: het toezicht op de opsporing. *Delikt en Delinkwent 37*, 406-418.

Sellens, E., Norris, J., Dhand, N., Heller, J., Hayes, L., Gidding, H., . . . Bosward, K. (2018). Willingness of veterinarians in Australia to recommend Q fever vaccination in veterinary personnel: Implications for workplace health and sfety compliance. *Plos One 13 (6)*.

Stevens, L., Pheiffer, M., Van den Broek, J., Keijzer, T., & Van der Hel-van Dijk, E. (2012, juni 6). *rijksoverheid.nl/documenten/rapporten.* Opgehaald van rijksoverheid.nl: https://www.rijksoverheid.nl/documenten/rapporten/2012/06/20/rapport-van-de-commissie-stevens-over-horizontaal-toezicht-bij-de-belastingdienst

Sureveen, J., & Rijsbergen, C. (1997). Het toezicht op financiële instellingen. *Maandblad voor accountancy en bedrijfseconomie oktober*, 490-496.

Voermans, W. (2005). De communautisering van toezicht en handhaving. *Data Juridica 0*, 69-87.

Winter, H., & De Ridder, J. (2010). *Toezicht en effectmeting. Het kán!* Groningen: Pro Facto BV.

WRR. (2013, september 9). *wrr.nl/publicaties/rapporten.* Opgehaald van wrr.nl: https://www.wrr.nl/publicaties/rapporten/2013/09/09/toezien-op-publieke-belangen.-naar-een-verruimd-perspectief-op-rijkstoezicht

Bijlage 1. Vragenlijst semi gestructureerde interviews

Vraag 1a. Welke toezichtinstrumenten zijn ingezet bij…? (afhankelijk van de respondent)
Vraag 1b. Is er ook gebruik gemaakt van communicatie in de vorm van het geven van gerechtvaardigde feedback?

Vraag 2a. Wat was de reden om voor deze toezichtinstrumenten te kiezen?
Vraag 2b. Is hierbij ook specifiek onderscheid gemaakt naar doelgroep? En zo ja, waarom?
Vraag 2c. Wat was de reden om voor deze mix van toezichtinstrumenten te kiezen?
Vraag 2d. Is de intensiteit van het toezicht (rakingspercentage) onderdeel geweest van de afweging van deze mix?
Vraag 2e. Is er specifiek aandacht geweest voor het inzetten van een sanctie t.a.v. overtreding van wet- en regelgeving?

Vraag 3. Wat was het verwachte/beoogde effect op de compliance?

Vraag 4. Op basis waarvan werd dit effect verwacht?

Vraag 5. Is dit effect in beeld gebracht? En zo ja, hoe?
Vraag 5a. Worden de kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten structureel vastgelegd?
Vraag 5b. Indien niet: is dat wel mogelijk? En zo ja, hoe?
Vraag 5c. Is het in beeld gebrachte effect betrouwbaar? Zo ja, waarom? Ze nee, waarom niet?

Vraag 6. Indien dit effect niet in beeld is gebracht, welke mogelijkheden ziet u om dit in beeld te brengen?
Vraag 6a. Zijn de kengetallen over de inzet van mensen, middelen, toezichtinstrumenten, handhavingsactiviteiten structureel vast te leggen?
Vraag 6b. Zo nee, waarom niet? Zo ja, hoe wel?

Vraag 7a. Was het beoogde effect ook te verwachten geweest bij de inzet van andere instrumenten?
Vraag 7b. Kunt u dit toelichten?

Vraag 8a. Zijn er ook effecten opgetreden die niet vooraf verwacht werden?
Vraag 8b.En zo ja, kunt u dit toelichten?