

Minderheden in de mixed-mode?

Minderheden in de mixed-mode?

Een inventarisatie van voor- en nadelen van mixed-mode dataverzameling onder niet-westerse migranten

R. Feskens
J. Kappelhof
J. Dagevos
I. Stoop

Het Sociaal en Cultureel Planbureau is ingesteld bij Koninklijk Besluit van 30 maart 1973.

Het Bureau heeft tot taak:

- a wetenschappelijke verkenningen te verrichten met het doel te komen tot een samenhangende beschrijving van de situatie van het sociaal en cultureel welzijn hier te lande en van de op dit gebied te verwachten ontwikkelingen;
- b bij te dragen tot een verantwoorde keuze van beleidsdoelen, benevens het aangeven van voor- en nadelen van de verschillende wegen om deze doeleinden te bereiken;
- c informatie te verwerven met betrekking tot de uitvoering van interdepartementaal beleid op het gebied van sociaal en cultureel welzijn, teneinde de evaluatie van deze uitvoering mogelijk te maken.

Het Bureau verricht zijn taak in het bijzonder waar problemen in het geding zijn die het beleid van meer dan één departement raken. De minister van Volksgezondheid, Welzijn en Sport is als coördinerend minister voor het sociaal en cultureel welzijn verantwoordelijk voor het door het Bureau te voeren beleid. Omtrent de hoofdzaken van dit beleid treedt de minister in overleg met de minister van Algemene Zaken, van Justitie, van Binnenlandse Zaken en Koninkrijksrelaties, van Onderwijs, Cultuur en Wetenschap, van Financiën, van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer, van Economische Zaken, van Landbouw, Natuur en Voedselkwaliteit, van Sociale Zaken en Werkgelegenheid.

© Sociaal en Cultureel Planbureau, Den Haag 2010

SCP-publicatie 2010/18

Zet- en binnenwerk: Textcetera, Den Haag

Figuren: Mantext, Moerkapelle

Omslagontwerp: Bureau Stijlzug, Utrecht

ISBN 978-90-377-0517-1

NUR 740

Voorzover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet 1912 dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3060, 2130 KB Hoofddorp, www.repro-recht.nl). Voor het overnemen van (een) gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (art. 16 Auteurswet 1912) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, Postbus 3060, 2130 KB Hoofddorp, www.cedar.nl/pro).

Sociaal en Cultureel Planbureau

Parnassusplein 5

2511 VX Den Haag

Telefoon (070) 340 70 00

Fax (070) 340 70 44

Website: www.scp.nl

E-mail: info@scp.nl

De auteurs van SCP-publicaties zijn per e-mail te benaderen via de website.

Inhoud

1	Naar een nieuw design voor de Survey Integratie Minderheden?	7
2	Modes van data verzamelen	9
2.1	Modes van data verzamelen	9
2.2	Kwaliteit van surveygegevens	10
2.3	Bij wie wordt gemeten: representeren de respondenten de doelpopulatie?	12
2.3.1	Steekproef	12
2.3.2	Dekkingsgraad	12
2.3.3	Non-respons	13
2.4	Meetfouten: worden de concepten goed gemeten?	17
2.5	Kosten	22
2.6	Samenvatting: sterke en zwakke punten van modes	22
3	Mixed-mode onderzoek: voor- en nadelen	24
3.1	Soorten mixed-mode design	24
3.2	Gevolgen van mixed-mode onderzoek	26
3.2.1	Respons en responsamenstelling	26
3.2.2	Meetfouten en vergelijkbaarheid tussen de modes	30
3.2.3	Analyse van mixed-mode gegevens	32
3.2.4	Vragenlijsten: dezelfde of een andere lijst in verschillende modes?	33
3.2.5	Trendgegevens: gevolgen voor de vergelijkbaarheid in de tijd	35
3.2.6	Kosten en doorlooptijd	35
3.3	Samenvatting en conclusies	37
4	Conclusies en varianten	41
4.1	Variant 1: behoud van het huidige design	41
4.2	Variant 2: mixed-mode onderzoek met kwaliteit als criterium	42
4.3	Variant 3: mixed-mode onderzoek met kosten als criterium	42
4.4	Variant 4: experiment met mixed-mode	43
4.5	Conclusies	43
	Noten	45
	Bijlage A Interviewgide	46
	Literatuur	50
	Publicaties van het Sociaal en Cultureel Planbureau	53

1 Naar een nieuw design voor de Survey Integratie Minderheden?

De Survey Integratie Minderheden (SIM) zal in 2010 op verzoek van het ministerie van Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer (VROM) worden uitgevoerd onder regie van het Sociaal en Cultureel Planbureau (SCP). In de SIM2010 worden data verzameld onder de vier grote groepen niet-westerse migranten en hun kinderen (van Turkse, Marokkaanse, Surinaamse en Antilliaanse origine) en autochtone Nederlanders. In eerdere peilingen gebeurde dit door middel van mondelinge dataverzameling: een interviewer komt aan de deur en neemt de vragenlijst af. De opdrachtgever heeft het SCP verzocht te onderzoeken of er – en zo ja, welke – mogelijkheden bestaan om de data te verzamelen met behulp van een mixed-mode design, waarbij ook andere dataverzamelingmethoden (ofwel: modes) worden gebruikt, en welke consequenties dit zou hebben voor de kwaliteit van de data, logistiek en kosten. Dit rapport geeft een overzicht van de literatuur op het gebied van mixed-mode studies en de resultaten van interviews met experts. Het tracht antwoord te geven op de volgende vragen:

- 1 *Wat zijn de voor- en nadelen van de verschillende dataverzamelingmodes?*
- 2 *Wat zijn de gevolgen van het gebruik van meerdere dataverzamelingmodes (mixed-mode) in één onderzoek voor de kwaliteit en kosten van het onderzoek?*
- 3 *Kan er in de SIM gebruikgemaakt worden van een mixed-mode ontwerp? En zo ja, hoe?*

De selectie van de literatuur over mixed-mode onderzoek is tot stand gekomen door zoekopdrachten op mixed-mode survey. Naast deze literatuurstudie zijn er gesprekken gevoerd met experts op het gebied van mixed-mode onderzoek en/of surveys onder migranten. Daarbij is geprobeerd om experts uit de academische wereld, overheidsinstellingen en werkzaam bij veldwerkbureaus te consulteren. Uiteindelijk zijn er gesprekken gevoerd met de volgende personen¹.

naam	instelling	specialisatie
drs. A. Ait Moha	Motivaction	research manager, gespecialiseerd in etnomarketing en onderzoek onder Nieuwe Nederlanders
prof. dr. J. Bethlehem	Centraal Bureau voor de Statistiek en Universiteit van Amsterdam	hoogleraar, deskundig op het gebied van survey en online onderzoek; verricht wetenschappelijk onderzoek gericht op het corrigeren van de uitkomsten van enquêtes voor het optreden van non-respons
dr. E. Cox	GfK/ Innovate	research manager, veldwerkcoördinator van het NELLS (Nederlandse levensloop studie)
dr. L. Heering	NIDI	onderzoeker, coördinator van het TIES (The Integration of the European second generation Study) en verantwoordelijk voor het Nederlandse survey

naam	instelling	specialisatie
prof. dr. E. de Leeuw	Universiteit Utrecht	hoogleraar Methoden en Technieken, met de nadruk op de kwaliteit van de survey; expertise op het gebied van dataverzameling, speciale groepen onderwerpen, non-respons en mixed-mode onderzoek
prof. dr. G. Lensvelt-Mulders	Universiteit voor Humanistiek	hoogleraar Wetenschapstheorie, Methodologie en Onderzoeksleer, onder andere onderzoek op het gebied van mixed-mode onderzoek en dan met name analyse van mixed-mode effecten
dr. A. Scherpenzeel	Centerdata, Universiteit Tilburg	projectleider online dataverzameling van het LISS-panel
drs. Y. Schothorst	Veldkamp	algemeen directeur gespecialiseerd op het gebied van onderzoek onder allochtonen
drs M. Hilhorst	Veldkamp	projectleider veldwerk SIM 2006
dr. I. Stoop	Sociaal en Cultureel Planbureau	hoofd Informatievoorziening en Automatisering, gespecialiseerd in non-respons en dataverzameling
prof. dr. J. Veenman	Erasmus Universiteit	hoogleraar Economische Sociologie, gespecialiseerd op het terrein van sociale ongelijkheid, met bijzondere aandacht voor de positie van allochtonen

2 Modes van data verzamelen

In dit hoofdstuk beschrijven we welke manieren ('modes') van data verzamelen zijn te onderscheiden en welke gevolgen de toepassing van deze modes heeft. Daarbij gaat de aandacht in het bijzonder uit naar de gevolgen voor data verzamelen onder niet-westerse migranten. Dit hoofdstuk richt zich ook op mode-effecten. Gevolgen van het toepassen van mixed-mode onderzoek komen in het volgende hoofdstuk aan de orde.

2.1 Modes van data verzamelen

Van oudsher worden drie hoofdvormen van data verzamelen onderscheiden: face-to-face of mondeling onderzoek, telefonisch en schriftelijk (De Leeuw 1992; Groves en Kahn 1979; Lyberg en Kasprzyk 1991). Face-to-face en telefonisch enquêteren vereisen vrijwel altijd een interviewer (telefonisch enquêteren waarbij een computer de vragen stelt, wordt zelden toegepast). Schriftelijke vragenlijsten vult men zelf in, zonder interventie van een interviewer, al kan een telefonische helpdesk soms ondersteuning bieden. De zelf-invul mode (meestal aangeduid met self-completion) omvat tegenwoordig ook web- of internetsurveys. Al deze vormen kunnen 'computer assisted' (ondersteund) zijn, een woord dat terugkomt in Engelse benamingen voor deze modes: CAPI, CATI, CASI en CAWI, ofwel *Computer-Assisted Personal, Telephone, Self en Web-Interviewing*. Bij vragenlijsten die gebruikmaken van internet bestaat er de mogelijkheid om aan respondenten een (persoonlijke) link te sturen naar een site waar ze de vragenlijst kunnen invullen. Deze vorm van CAWI wordt ook wel web- of online survey genoemd (Bethlehem 2009). Tabel 2.1 geeft een overzicht van de verschillende modes ingedeeld naar de inzet van computer en interviewer.

Tabel 2.1
Dataverzamelmethode(n) (modes)

	computer-assisted		not computer-assisted	
	interviewer	zonder interviewer	interviewer	zonder interviewer
mondeling (face-to-face)	CAPI		PAPI (paper and pencil interviewing)	
telefonisch	CATI	interactive voice response, etc.	telefonisch	
schriftelijk		CASI (e-mail; invullen in aanwezigheid interviewer)		schrifte-lijke vragenlijst (toegestuurd (mail) of achtergelaten door interviewer (drop-off))
internet		CAWI (websurvey)		

Iedere vorm van data verzamelen heeft sterke en zwakke kanten. In dit hoofdstuk wordt geïnventariseerd op welke manier de onderscheiden modes zich verhouden tot de kwaliteit van surveygegevens. De volgende paragraaf beschrijft in het kort welke criteria aangehouden worden om de kwaliteit van surveygegevens te bepalen. De rest van het hoofdstuk concentreert zich op twee aspecten van kwaliteit, namelijk:

- Vormen de respondenten een goede afspiegeling van de populatie (thema's: steekproef, dekking en non-respons)?
- Is sprake van een goede meting (thema's: meetfouten vanwege vragenlijst, interviewer en respondent)?

Deze twee aspecten van kwaliteit worden besproken in relatie tot de onderscheiden modes van onderzoek.

2.2 Kwaliteit van surveygegevens

De kwaliteit van surveygegevens wordt afgemeten aan een aantal criteria (zie Eurostat 2005), namelijk:

- 1 relevantie;
- 2 accuraatheid en betrouwbaarheid;
- 3 tijdigheid;
- 4 coherentie en vergelijkbaarheid;
- 5 toegankelijkheid en begrijpelijkheid.

Andere Eurostat-criteria zijn onder andere het beperken van de responsdruk, het gebruik van goede methoden en effectiviteit van de inzet van kosten. De criteria voor de kwaliteit van surveygegevens worden in deze paragraaf kort besproken.

Het eerste criterium voor de kwaliteit van surveydata is *relevantie*. De relevantie hangt natuurlijk af van het onderzoeksdoel. Als monitoren op een aantal hoofdlijnen voldoende is, is een beperkte gegevensset relevant. Als naar verbanden en verklaringen wordt gezocht, zijn veel meer gegevens nodig. Als de gebruikersgroep groot is, bijvoorbeeld wanneer een bestand door de overheid en door verschillende groepen academische onderzoekers wordt gebruikt, geldt al gauw dat hoe meer onderwerpen aan de orde komen, en hoe meer in detail wordt gevraagd, hoe groter de relevantie is. Dit kan natuurlijk op gespannen voet komen te staan met de responsdruk en kan leiden tot hogere kosten. De maximum tijdsduur voor het beantwoorden van een vragenlijst (en daarmee het aantal onderzoeksvragen dat kan worden beantwoord) is zeer sterk afhankelijk van de mode: een face-to-face survey kan veel langer duren dan een telefonisch of websurvey. De keuze voor een mode kan daarmee beperkingen opleggen aan de relevantie van een survey.

Een tweede criterium is de *accuraatheid en betrouwbaarheid* van de surveydata. De betrouwbaarheid is een afgeleide van de accuraatheid. De accuraatheid is het verschil tussen de geschatte waarde van het survey en de werkelijke waarde van een onderzoeksobject. Afwijkingen kunnen zowel van toevallige als systematische aard zijn. Toevallige afwijkingen hebben invloed op de varianties van schatters: deze zullen groter worden bij

een groot aantal toevallige fouten. Wanneer de fouten systematisch van aard zijn, zal dit tot een vertekening in de schatters leiden. De schatter wordt in dit geval onzuiver genoemd (cf. Bethlehem 2004). Zeker in dit laatste geval kunnen verkeerde conclusies uit de verzamelde data getrokken worden. Er zijn allerlei mogelijke foutenbronnen bij het uitvoeren van een enquêteonderzoek die de accuraatheid van de verzamelde gegevens kunnen bedreigen (Bethlehem, 2009; Biemer en Lyberg 2003; Groves 1989). Hox e.a. (2008) geven aan dat de kwaliteit van enquêtedata bepaald wordt door de kwaliteit van de dekking, steekproef, respons en meting. In het design van een survey is het doel de accuraatheid van de resultaten te maximaliseren binnen beperkingen van budget en tijdsplanning (Biemer en Lyberg 2003).

Kortom, de accuraatheid van de surveydata heeft twee aspecten, namelijk bij wie er gemeten wordt (dekking, steekproef, non-respons) en hoe of wat er gemeten wordt (meetfouten). Op beide punten komen we in het vervolg van dit hoofdstuk uitvoeriger terug.

Tijdigheid, het derde criterium, is de snelheid waarmee gegevens beschikbaar zijn. Deze wordt bepaald door de duur van de voorbereiding, het data verzamelen zelf en de benodigde bewerking van de gegevens. De gemiddelde doorlooptijd van een enquête varieert per afnamevorm. Internet en telefonische enquêtes zijn over het algemeen sneller afgerond dan mondelinge en schriftelijke. Bij een mondelinge enquête moet de interviewer soms herhaaldelijk naar het adres van een beoogde respondent reizen, en is deze voor de duur van het interview bezet. Dat laatste geldt ook voor telefonische interviews, maar herhaaldelijk bellen om contact te leggen met personen verspreid over het hele land kost natuurlijk nauwelijks tijd. Bij een schriftelijke (en ook internet) mode kan aan iedereen tegelijk een vragenlijst worden toegestuurd, maar moet de onderzoeker wachten totdat men een vragenlijst invult en terugstuurt. Met gepaste intervallen kan de onderzoeker een reminder sturen om de respondent aan te moedigen mee te werken aan de enquête, maar de reminder brengt ook weer extra wachttijd mee. Bij een internetenquête is het effect van een reminder eerder bekend dan bij een schriftelijke vragenlijst verstuurd via de post. Een laatste punt dat samenhangt met de doorlooptijd is de snelheid waarmee de gegevens worden verwerkt. De computer ondersteunde varianten hebben als groot voordeel dat de data zo goed als direct beschikbaar zijn nadat het interview is afgenomen. De schriftelijke vragenlijsten nemen ook hier het meeste tijd. De data moet in dat geval nog handmatig worden ingevoerd of gescand. Handmatig invoeren kan overigens een extra foutenbron zijn.

Het vierde criterium is *coherentie en vergelijkbaarheid*. Dit speelt een rol wanneer men gegevens uit verschillende surveys, periodes of landen of van verschillende doelgroepen wil vergelijken. Standaardclassificaties en -definities vergemakkelijken de vergelijkbaarheid (bijvoorbeeld van opleiding en beroep). Vergelijkbaarheid wordt bemoeilijkt wanneer vragen niet op dezelfde wijze zijn gesteld, of verschillende dataverzamelmethodeën zijn gebruikt.

Een vijfde en laatste criterium voor de kwaliteit van surveydata is *toegankelijkheid en begrijpelijkheid*. Als data niet – via bijvoorbeeld een data-archief – toegankelijk zijn, is de waarde zeer beperkt. Als niet bekend is hoe de vragen precies gesteld zijn, hoe de steekproef is getrokken, wat de respons was, enzovoort is de kwaliteit van een survey moeilijk te beoordelen en zijn de gegevens moeilijk te vergelijken.

2.3 Bij wie wordt gemeten: representeren de respondenten de doelpopulatie?

Zoals gezegd staan we bij twee punten met betrekking tot de datakwaliteit uitvoeriger stil. In deze paragraaf gaat het om de vraag of de respondenten een goede afspiegeling vormen van de populatie en wat de gevolgen van de diverse modes hiervoor zijn. Dit raakt aan kwesties als de steekproef, de dekkingsgraad en de non-respons.

2.3.1 Steekproef

Over de steekproef en steekproeftrekking kunnen we betrekkelijk kort zijn. Volgens de steekproeftheorie kan een steekproef een representatief beeld geven van een populatie als alle leden van de populatie een bekende kans groter dan nul hebben om mee te doen. In de Survey Integratie Minderheden (SIM) is dit het geval, en er mag van worden uitgegaan dat ook in de nieuwe SIM een random steekproef getrokken zal worden.

2.3.2 Dekkingsgraad

Voorts is het steekproefkader van belang; dit is de lijst waaruit de personen worden getrokken. Het verschil tussen de daadwerkelijke onderzoekspopulatie en het steekproefkader bepaalt de dekkingsgraad. In de SIM is de gemeentelijke basisadministratie (GBA) het steekproefkader, met een zeer hoge dekkingsgraad.² Serieuze problemen rond de dekkingsgraad spelen wel bij het gebruik van de *telefonische mode*. Het Centraal Bureau voor de statistiek (CBS) schat dat het percentage huishoudens met een (bekende) vaste aansluiting schommelt tussen de 65% en 70%. Onder niet-westerse migranten is dit nog een stuk lager (Schothorst 2002). Ook Banning e.a. (2009) wijzen op onderdekking bij Computer Assisted Telephone Interviewing (CATI) onder niet-westerse migranten. In de recent gehouden SIM³- en SING⁴ is nagegaan of de adresgegevens van te herbenaderen respondenten (personen die vier keer niet thuis bleken te zijn en 'zachte' weigeraars⁵) konden worden uitgebreid met telefoonnummers. Onder migranten bleken zeer weinig personen te beschikken over een vast en bekend nummer. In een onlangs gehouden bijeenkomst over mixed-mode onderzoek stelde CBS-medewerker Schouten dat het hoge aandeel personen zonder geregistreerd telefoonnummer de toepassing van CATI bij de uitvoering van een mixed-mode aanpak serieus bemoeilijkt. Daar komt nog bij dat steeds meer personen aangeven dat hun telefoonnummers niet mogen worden gebruikt om hen te benaderen voor onderzoek. Ook om die reden acht de hierna geciteerde expert het gebruik van de telefoon steeds minder geschikt.

Telefoon is volgens mij een stervend iets. Dat is gewoon een sterfhuisconstructie. Dat mag ik niet zeggen en ik beledig nu een heleboel collega's, maar die telefoon gaat eruit, met het toenemen van mobiele telefoons waardoor je dus helemaal geen random steekproef kunt trekken. Plus het feit dat daar zo'n boycot op komt als www.belmijniet.nl. Daar mag geen random digital dialing, want dat is slecht voor de privacy en zo. Nou, dat is de dood in de pot. Dus als je nu iets nieuws opzet ga je niet inzetten op telefoon.

Dit betekent niet dat de telefoon helemaal niet gebruikt zou kunnen worden. Verschillende van de door ons geïnterviewde experts geven aan dat ze bij de benadering van respondenten het telefoonnummer vragen om een afspraak te maken.

Wat de interviewer nu ook wel doet en dat is iets nieuws, ook met name bij de allochtonen, is een kaartje achterlaten waarbij ze zelf een mobiel nummer kunnen opgeven. Die kunnen ze zelf terugsturen of zelf contact opnemen met de interviewer om een afspraak te maken. Dus dat is iets wat we hebben toegevoegd.

Webenquêtes en dekkingsgraad

Het uitvoeren van webenquêtes onder niet-westerse migranten en hun kinderen heeft ook te kampen met een dekkingsprobleem. Uit een studie van Van Ingen e.a. (2007) blijkt dat er minder computerbezit en internettoegang is onder niet-westerse migranten (in het bijzonder Turkse en Marokkaanse Nederlanders), waardoor een substantieel deel van de doelpopulatie via deze mode niet wordt bereikt.

2.3.3 Non-respons

Naast de dekkingsgraad bepaalt de respons of een steekproef als representatief kan worden beschouwd. De kwaliteit van de respons hangt af van de omvang en samenstelling ervan. Zogenaemde unit-non-respons kan voortkomen uit

- er niet in slagen contact te leggen met de steekproefeenheid;
- de steekproefeenheid is niet in staat tot deelname aan het interview; en
- de steekproefeenheid weigert deelname aan een interview.

Non-respons verkleint de precisie van surveyresultaten, omdat het aantal respondenten kleiner wordt. Non-responsvertekening treedt op als er systematische verschillen zijn tussen de respondenten en de non-respondenten op de kenmerken die je wilt onderzoeken. Hoe meer respondenten en non-respondenten verschillen, hoe groter het effect van dit verschil. Zowel de hoogte van het non-responspercentage als de grootte van het verschil tussen respondenten en non-respondenten heeft dus invloed op de non-responsvertekening.

Op veel manieren kan worden geprobeerd een hoge respons te behalen. Eén manier is vragenlijsten niet al te lang te maken (al is een saaie vragenlijst al gauw lang en wil men over een interessant onderwerp best wat langer praten), schriftelijke vragenlijsten goed te ontwerpen (Dillman 2007) en begrijpelijke en interessante vragen te stellen. Goede introductiebrieven, veel contactpogingen of reminders, cadeautjes bij deelname en kun-

dige, gemotiveerde interviewers kunnen ook helpen. De modes verschillen van elkaar in non-respons. Ruwweg behaalt de duurste mode (face-to-face) de hoogste respons, en de goedkoopste (internet) de laagste. Een verklaring hiervoor is de aan- of afwezigheid van een interviewer. Een goed getrainde interviewer kan weigerachtige respondenten overhalen mee te doen aan het onderzoek (Bethlehem 2009). Ook ongeletterdheid, iets wat bij oudere niet-westerse migranten regelmatig voorkomt, is geen probleem als de interviewer de vragen voorleest. Tweektalige interviewers kunnen ervoor zorgen dat respondenten die de originele taal van de vragenlijst niet beheersen, toch kunnen meedoen. Aan de andere kant wil niet iedereen een interviewer binnenlaten, en geven mensen er soms de voorkeur aan een vragenlijst zelf in te vullen wanneer het hen uitkomt (Bronner en Kuijlen 2006). Dit is mogelijk bij een schriftelijke (en internet)vorm van data verzamelen.

Non-respons is een van de grootste bedreigingen voor de juistheid van uitkomsten van enquêteonderzoek. Zeker in Nederland zijn de responscijfers een reden tot zorg. Een hoog non-responscijfer hoeft op zichzelf nog niet te leiden tot verkeerde conclusies, maar er zijn veel aanwijzingen dat de non-respondenten anders zijn dan de respondenten en er dus sprake is van een selectieve non-respons (Schmeets 2005a en b; Schouten en Cobben 2007; Stoop 2005). In zo'n geval kunnen er verkeerde conclusies uit de data worden getrokken (Groves 1989).

Meer in relatie tot het veldwerk onder niet-westerse respondenten kan het volgende worden gezegd over de respons in verschillende modes.

Schriftelijk

Schriftelijke vragenlijsten leiden over het algemeen tot een lagere respons, en dit geldt vanwege taalproblemen en ongeletterdheid in het bijzonder voor niet-westerse migranten. Een nog steeds substantieel deel van hen beheerst de Nederlandse taal niet (voldoende) of is analfabeet. Dit komt vooral voor bij eerstegeneratie ouderen van Turkse en Marokkaanse origine en huwelijksmigranten. Selectieve respons is het gevolg.

In het verleden hebben we het wel eens eerst schriftelijk gedaan, bijvoorbeeld aan de hand van een vragenlijst. Daarbij merkten we voor het schriftelijke deel met name dat we van de hoger opgeleiden en jongeren meer terugkregen. We hebben het dan over een hele selectieve steekproef onder allochtonen.

Het vertalen van de vragenlijst is natuurlijk een mogelijkheid om non-respons als gevolg van taalproblemen te verminderen. Veel Marokkaanse Nederlanders spreken echter Berbers, dat geen geschreven taal is. Voor hen is het vertalen van een vragenlijst dus geen optie. Een door ons geïnterviewde respondent heeft ook geen goede ervaringen met vertaalde schriftelijke vragenlijsten bij personen van Turkse en Marokkaanse origine.

Toen hebben we ze een Nederlandse en een vertaalvragenlijst toegestuurd. En de vertaalvragenlijst heeft negen exemplaren opgeleverd. En het probleem daarbij is ook nog eens dat je wel wat met de gesloten antwoorden kunt, maar de niet-gesloten antwoorden zijn best ingewikkeld. Wij zijn niet eens begonnen met invoeren. We hebben alleen de gesloten antwoorden ingevoerd. [...] Er zijn er denk ik iets van 500 uitgestuurd. Negen is echt heel erg weinig. Het probleem is dat niemand weet hoe het komt. Die mensen hebben dus een tientje gekregen om dat in te vullen en ze doen het gewoon niet.

Verder zullen zich bij de SIM-doelgroep naar verwachting vaak problemen voordoen met het begrip van de vraag (Van der Zouwen 2006). Het zelf invullen van vragen vraagt meer taalvaardigheden dan het 'alleen' mondeling rapporteren. Het probleem van de lengte van de vragenlijst is bij deze categorie respondenten in het bijzonder aanwezig. Het is niet doenlijk om langere, complexe vragenlijsten af te nemen; dit leidt niet alleen tot een lagere respons, maar ook de kwaliteit van de antwoorden zal afnemen vanwege de langere vragenlijst (Herzog en Bachman 1980; Groves e.a. 2000).

Een ander belangrijk nadeel van een schriftelijke lijst is dat niet valt te achterhalen of de geselecteerde persoon ook degene is die de vragenlijst invult. Een ander familielid kan de vragenlijst invullen voor de respondent of de respondent 'helpen' met het invullen van de vragenlijst. Dit is in algemene zin een nadeel van een schriftelijke lijst, maar zal zich naar verwachting in het bijzonder voordoen bij respondenten die het Nederlands niet machtig zijn.

Daarnaast viel ons de invulling op bij een redelijk aantal vragenlijsten die we terugkregen. Deze zou bijvoorbeeld eigenlijk ingevuld moeten zijn door bijvoorbeeld een vrouw van rond de 65, in Marokko geboren of van de eerste generatie. Dus op een bepaalde leeftijd hier naartoe gekomen, zonder onderwijs of waarvan je aan zou kunnen nemen dat ze heel slecht het Nederlands beheerst. Dat we daarvan een keurig ingevulde vragenlijst terugkregen, die duidelijk door iemand anders was ingevuld. Waarbij niet zeker was dat de persoon die dat had ingevuld een goed intermediair is geweest of dat deze hem eigenlijk naar eigen inzicht had ingevuld. Dus dat soort aspecten kwam daar vooral in terug.

Internet

Hiervoor is al gewezen op het feit dat er minder computerbezit en internettoegang is onder niet-westerse migranten (met name onder Turkse en Marokkaanse Nederlanders), waardoor een substantieel deel van de doelpopulatie niet zal worden bereikt (Van Ingen e.a. 2007). De respons zal ook lager uitvallen dan bij een CATI-mode. Zo zijn geletterdheid, taalbeheersing en ook computerervaring een vereiste bij de webmode, waardoor meer niet-westerse migranten niet in staat zijn om deel te nemen. Ook hier is het vertalen van de vragenlijst natuurlijk een mogelijkheid om non-respons als gevolg van taalproblemen te verminderen (zie eerder). Net als bij schriftelijke vragenlijsten is er een grotere kans op meetfouten, omdat er geen interviewers aanwezig zijn om vragen te verduidelijken wanneer een respondent moeite met de taal heeft. Een geïnterviewde expert die bezig is met de opbouw van een panel zegt hierover het volgende:

Ja, dat zien we nu ook bij de online interviews. Als je sommige niet-westerse allochtonen binnen hebt [als deelnemer van het panel], blijken die soms wel moeite te hebben met onze vragenlijsten. Dat ze het toch niet begrijpen, de vragen zijn soms best moeilijk.

Tot slot moeten de verwachtingen over de respons via het web niet te hoog gespannen zijn (in het volgende hoofdstuk komen we hier nog op terug). Over het algemeen is de respons op internetsurveys niet hoog.

Stel dat je web overweegt, dan moet je toch netjes de steekproeftrekken. Dus dan moet je eerst een brief sturen met een link en de respons daarop is laag. Wij zijn bij het CBS nog niet boven de 30% uitgekomen.

Telefoon

Een telefonische enquête leidt over het algemeen tot een hogere respons dan een schriftelijke vragenlijst. Schothorst wees er in 2002 op dat niet-westerse migranten telefonische interviews als zeer ongebruikelijk ervaren, waardoor er vaker wordt geweigerd (Schothorst 2002). Onduidelijk is of dit in 2010 nog steeds het geval is. Bij een telefonische enquête is sprake van mondelinge antwoorden, zodat ook ongeletterde respondenten kunnen meedoen. Tweektalige interviewers zijn noodzakelijk om respondenten met onvoldoende taalbeheersing te bevragen in de herkomsttaal. De noodzaak voor tweektalige interviewers zal bij CATI groter zijn dan bij CAPI, omdat de mogelijkheid van een (vertaalde) hulpvragenlijst wegvalt. Deze kan immers niet worden getoond. Wel kan een tweektalige interviewer over de telefoon veel interviews afnemen. De vragenlijsten zullen dan wel kort moeten zijn.

Telefonisch kan niet, want die vragenlijst van het SIM is te lang. Dat kan gewoon niet. Dat is echt onmogelijk.

Mondeling

Verschillende auteurs zijn van mening dat een mondelinge mode de beste strategie is om respons onder niet-westerse allochtonen te verkrijgen (Schmeets 2005a; 2005b; Feskens e.a. 2006; 2007). Een mondelinge mode zorgt ervoor dat ook ongeletterde respondenten kunnen meedoen. De kans op taalproblemen blijft in deze mode echter even groot, maar er is wel meer aan te doen. Zo kan men gebruikmaken van tweektalige interviewers om respondenten met onvoldoende taalbeheersing te interviewen in hun eigen taal. De interviewers hebben dezelfde etnische achtergrond als de doelpopulatie en spreken zowel Nederlands als de taal van de doelpopulatie (bijvoorbeeld Turks, Arabisch, Berbers). In de SIM2006 behaalden zij een hogere respons dan autochtone enquêteurs (Kappelhof 2010; Dagevos e.a. 2007). Een bijkomend voordeel is dat deze niet-westerse interviewers bekend zijn met de cultuur en gebruiken en dus een kleinere sociale afstand tot de onderzoeksgroep hebben (Veenman 2002). Er kleven ook nadelen aan het gebruik van tweektalige interviewers, zeker als deze apart geworven moeten worden. Schothorst (2002) stelt dat deze interviewers extra goed geïnstrueerd dienen te worden. Ervaringen met de SIM2006- en SING2009 leren dat het verloop onder deze

enquêteurs groter is dan onder autochtone interviewers. Wanneer er gekozen wordt voor autochtone interviewers is het mogelijk gebruik te maken van een vertaalde (hulp) vragenlijst. Een interviewer kan wanneer de vraag voor de geïnterviewde niet duidelijk is, deze aanwijzen in de vertaalde vragenlijst. Dit zal de kans op meetfouten aanzienlijk verminderen. In dit geval kunnen interviewers dus een positief effect op de kwaliteit van antwoorden hebben.

De hier genoemde voordelen laten onverlet dat de mondelinge, face-to-face afname-methode onder niet-westerse migranten zeker niet zonder problemen is. Onderzoeken als de SIM zijn duur. De inzet van tweetalige enquêteurs levert weliswaar een hogere respons op, maar het heeft ook een vertragende invloed op het tempo en de organisatie van het veldwerk.

2.4 Meetfouten: worden de concepten goed gemeten?

Meetfouten treden op wanneer een antwoord op een vraag niet of niet nauwkeurig het concept meet dat door de onderzoeker is bedoeld. Zowel de vragenlijst, interviewer, respondent als de afnamevorm (mode) kan een bron van meetfouten zijn. De verschillende modes met betrekking tot vragenlijst, respondent en interviewer hebben elk hun eigen repercussies voor meetfouten.

Vragenlijsten

Afnamevormen verschillen in de mate waarin ze geschikt zijn voor bepaalde vragenlijsten (Roberts 2007). De complexiteit van de vragenlijst, de lengte van de vragenlijst en het toestaan om vragen over te slaan (item non-respons) hebben invloed op de optimale keuze voor een mode. Complexe routingschema's en andere methodologische opties zoals randomisering van schaalitems, antwoordcategorieën om antwoordeffecten tegen te gaan en uitgebreide hulpschema's zijn vooral mogelijk bij computer ondersteunde vragenlijsten. De lengte van een vragenlijst is een andere factor in de keuze voor een geschikte mode. Bij internet- en schriftelijke afname kan over het algemeen geen lange vragenlijst worden afgenomen (Roberts 2007). Bij een mondeling interview kan de vragenlijst het langst zijn.

Een andere relevante dimensie bij de vragenlijst is het leiden door een vragenlijst / toestaan van item non-respons. Bij de mondelinge en telefonische afname kan een interviewer de respondent helpen, uitleg geven. De interviewer heeft hierbij de leiding. Bij schriftelijke vragenlijsten kan een respondent ervoor kiezen een vraag over te slaan. Diverse geïnterviewde experts onderschrijven dit:

Het is een stuk minder voor de hand liggend (als er een interviewer is) dat je een blok overslaat als je een vraag krijgt die je saai vindt, niet helemaal begrijpt, of niet wilt. Als er echt een interviewer tegenover je zit die deze dingen vraagt, dan is het natuurlijk veel dwingender om erover na te denken en toch antwoord te geven. Ik denk dat de hoeveelheid blokken die overgeslagen wordt omdat men het niet weet, of iemand die afen toe ergens links of rechts eens een kruisje zet veel groter is bij een schriftelijke vragenlijst. Dat heeft gevolgen voor de kwaliteit.

Bij internetafname moet er goed worden nagedacht of respondenten een vraag mogen overslaan. Het wel toestaan kan leiden tot item non-respons. Het niet-toestaan kan leiden tot satisficing (d.w.z. een antwoord geven om er maar vanaf te zijn, zie verder), onzinnantwoorden of uitval.

Kort gezegd worden meetfouten geminimaliseerd wanneer bij een lange, complexe vragenlijst gebruik wordt gemaakt van CAPI. Is de vragenlijst kort en eenduidig, dan zijn de meetfouten bij een schriftelijke/internet mode minder groot dan bij een lange, complexe lijst.

Interviewer

Bij telefonische en mondelinge interviews kan de interviewer een *respondent helpen* bij het afnemen van het interview door toelichting of uitleg te geven wanneer vragen niet duidelijk zijn. Wanneer het invullen van de vragenlijst zoveel problemen oproept dat er iemand aanwezig moet zijn om de respondent te helpen bij het geven van de gevraagde informatie, is een persoonlijk interview de aangewezen waarnemingsmethode (Van der Zouwen 2006). Overigens zou men ook kunnen concluderen dat een vragenlijst die zoveel problemen oproept geen goede vragenlijst is, of kan leiden tot grote interviewer-effecten.

Een interviewer kan ook een nuttige rol spelen wanneer een respondent de *taal* waarin de vragenlijst is opgesteld slecht beheerst. In dat geval kan de interviewer de respondent – in beperkte mate – helpen bij het afnemen van het interview. Taalproblemen spelen overigens een nadelige rol bij alle dataverzamelingmethoden, zij het dat een (tweetalige) interviewer in sommige gevallen uitkomst kan bieden door vragen te verduidelijken.

Verder kan een interviewer ervoor zorgen dat de *juiste respondent de vragen beantwoordt*, iets wat bij schriftelijke en internetenquêtes niet mogelijk is, en aanvullende gegevens noteren over de situatie waarin het interview plaatsvond. Hierbij valt te denken aan de aanwezigheid en mate van invloed die derden hebben op de antwoorden bij een interview, of de respondent de vraag goed begrepen heeft, de tijd die een respondent voor de vraag neemt, maar ook de aanwezigheid van afleiding, zoals televisie.

Aan de andere kant kan de aanwezigheid van een interviewer tot vermindering van de kwaliteit van de meting leiden. Meerdere studies laten zien dat respondenten eerder geneigd zijn in aanwezigheid van een interviewer *meer sociaal wenselijk* gedrag te rapporteren (De Leeuw 1992; Van der Zouwen en De Leeuw 1990, 1991; De Leeuw 2005; Bronner en Kuijlen 2006). Een geïnterviewde expert heeft de volgende ervaring:

Dat [bij telefonische interviews] de mensen ook wat minder extreem waren in hun antwoorden. Er zat inderdaad ook wat meer sociale wenselijkheid in. Dat zijn van die dingen die je ook wel verwacht. Ik ben geneigd om te denken dat online, zonder interviewer, realistischer antwoorden oplevert. We hebben onlangs ook online onderzoek gedaan naar huiselijk geweld en daar krijg je verrassend open antwoorden over.

Bovendien kan de aanwezigheid van een interviewer ervoor zorgen dat een respondent *minder tijd neemt* om goed over een antwoord na te denken, wat kan resulteren in meer foute antwoorden.

Respondent

Een andere bron van meetfouten zijn respondenten. Er zijn meerdere stappen die een respondent doorloopt bij het beantwoorden van iedere enquêtevraag. Tourangeau e.a. (2000) onderscheiden vier cognitieve stappen die een respondent moet zetten in het responsproces:

- 1 comprehension
- 2 retrieval
- 3 judgment
- 4 report

Ieder stap in dit proces kan een meetfout veroorzaken en de kans hierop varieert per afnamevorm. Het al dan niet aanwezig zijn van een interviewer blijkt een belangrijke variabele. Een aantal van de zojuist genoemde punten bij de interviewer komen hier terug, maar dan vanuit het perspectief van de respondent.

Ad 1 Comprehension

De eerste stap in het responsproces is het begrip van de vraag. Hier kunnen verschillen tussen modes optreden afhankelijk van de aanwezigheid van een interviewer. Het invullen van schriftelijke vragenlijsten vraagt geletterdheid en een relatief hoog niveau van actieve beheersing van de taal (De Leeuw 2008). Dit geldt ook voor het invullen van webenquêtes, dat computervaardigheden vereist en toegang tot het internet. Mondelinge enquêtering kan dus als de meest laagdrempelige manier van data verzamelen gezien worden. Bovendien kunnen interviewers hulp bieden bij het responsproces op het moment dat een respondent een vraag niet goed begrijpt. De verschillen in het begrip van de vraag worden geregistreerd in het gebruik van afzonderlijke modes, maar worden ook gerapporteerd in mixed-mode studies. Aitkin e.a. (2004) geven aan dat de afwezigheid van interviewers tot moeilijkheden bij de interpretatie van vragen kan leiden. Schonlau e.a. (2004) en Nicolaas en Tipping (2004) laten dezelfde moeilijkheden zien voor bepaalde antwoordcategorieën bij de afwezigheid van interviewers (cf. Ariel e.a. 2008).

Ad 2 Retrieval en ad 3 Judgement

In stap 2 van het responsproces probeert de respondent zich de relevante informatie te herinneren door in het geheugen te zoeken (Tourangeau e.a., 2000). In stap 3 beoordeelt de respondent of de herinnerde informatie relevant en compleet is. Tijdens deze twee fasen van het responsproces kan de aanwezigheid van een interviewer een nadelige invloed hebben. De Leeuw (2008) geeft aan dat er in mondelinge conversaties een grotere acceptatie is om een stilte te laten vallen dan in telefonische conversaties. Wanneer de respondent dus het idee heeft snel te moeten antwoorden en als gevolg daarvan zijn of haar geheugen niet compleet heeft afgezocht naar de relevante informatie, kunnen er meetfouten ontstaan. Bij schriftelijke en webenquêtes kan een

respondent alle tijd nemen om zich de informatie te herinneren en te beoordelen of deze relevant en compleet is.

Ad 4 Report

In de laatste fase van het surveyproces geeft de respondent het antwoord op de vraag. Ook in deze fase kan een interviewer een nadelige invloed hebben op de datakwaliteit en tot meer sociaal wenselijke antwoorden leiden (De Leeuw 1992; Van der Zouwen en De Leeuw 1990, 1991; De Leeuw 2005; Bronner en Kuijlen 2006), vooral bij gevoelige vragen en onderwerpen. Wat gevoelige onderwerpen zijn (bv. geld, seks, drankgebruik), verschilt echter per populatie. Het is dus niet altijd vóór de start van het data verzamelen bekend welke onderwerpen onderhevig kunnen zijn aan vertekening door sociale wenselijkheid. Een voorbeeld is afkomstig uit een presentatie van Ahmed Ait Moha (2009), researchmanager bij Motivaction. Uitkomsten van online onderzoek onder 'nieuwe Nederlanders' wijzen op een aanzienlijk hoger drankgebruik dan bij een face-to-face onderzoek. Ook gooit men in online onderzoek het glas minder vaak in de glasbak dan in face-to-face onderzoek. Meerdere auteurs vinden verschillen in de meting van antwoorden op gevoelige vragen (Link en Mokdad 2005; Aquilino 1994; Greenfield e.a. 2000; Jäckle e.a. 2006; cf. Ariel e.a. 2008; Kampen 2006).

Verskillende experts geven aan dat mode-effecten vaak optreden bij sensitieve vragen:

Vragen als hoe oud iemand is en waar iemand is geboren lijken me niet zo aan twijfel onderhevig. Maar bij alles wat met meningen te maken heeft, maakt het natuurlijk veel uit of er wel of niet een interviewer tussen zit.

Ik weet dat bij het woononderzoek momenteel heel sterk mixed-mode gebruikt wordt, omdat je zoveel mogelijk mensen te pakken wilt krijgen. Maar dat zijn volgens mij allemaal heel feitelijke gegevens. Dingen als het hebben van verhuisplannen lijkt me niet zo gevoelig voor welke mode je daarvoor gebruikt. Maar dat is anders bij echt gevoelige dingen.

Maar ook over harde vragen is er wel enige zorg.

Wij bij [...] hebben in ieder geval wel mode-effecten aangetoond bij vrij harde vragen over werkloosheid.

Kortom:

Zodra het over gevoelige onderwerpen gaat, maakt het zo verschrikkelijk veel uit of er een interviewer aanwezig is.

Dat is eigenlijk het enige moment waarop het nadelig is om een interviewer te gebruiken. Over het algemeen is een interviewer natuurlijk positief, want die duwt mensen, trekt ze over de drempel, helpt ze bij het uitleggen van dingen, vragen en zo krijg je een betere kwaliteit.

Het kan iets doen met de kwaliteit van je antwoorden, om de doodeenvoudige reden dat er niemand is om te helpen.

Bij een mondelinge vragenlijst kan een interviewer mensen een beetje helpen als het gaat over hoe je een vragenlijst in moet vullen. Het begint al met het trekken als iemand zegt; nou ik heb eigenlijk niet zo'n zin om mee te doen. De medewerking, dat is één. En in de tweede plaats zijn er ook mensen die zeggen; ja, ik begrijp de vraag niet zo goed. [...] Dan vertaalt hij hem in zijn hoofd mentaal in een vraag die hij wel begrijpt en daar geeft hij dan antwoord op. Maar zo'n interviewer kan nog helpen om de vraag echt uit te leggen en dat leidt tot betere antwoorden op de vragen.

Een andere bron van fouten in de laatste fase van het responsproces is *acquiescence* (Krosnick 1991; 1999), de neiging van respondenten om het met alles eens te zijn, ongeacht de inhoud van de vraag. Het aantal respondenten dat dergelijk gedrag vertoont, is soms wel 15% tot 20% (Krosnick 2009). *Acquiescence* wordt ook wel gezien als een vorm van *satisficing* (Krosnick 1991). *Satisficing* vindt plaats wanneer een respondent aan de stappen 2 en 3 van het responsproces weinig tot geen aandacht schenkt (Krosnick en Alwin 1987). Zodra de vraag begrepen wordt, komt er na een zeer korte 'search' een antwoord. Respondenten lijken meer gemotiveerd consciëntieus te responderen met een interviewer erbij. Bij webinterviews zou dit vaker in het geding zijn:

Wat je dus ook wel gauw ziet bij webvragenlijsten is dat men ze toch wel afmaakt, maar dat er heel veel gekloot wordt. En tatatatat, zo ingevuld en dan weet je al dat het mis is. Iemand had experimenten gedaan en bijgehouden hoe lang het duurde voordat de vragenlijst was ingevuld. En dan bleek een hele lange vragenlijst in vijf minuten te zijn ingevuld. Nou dan weet je het wel. Dus dan nog kan het zijn dat mensen ze wel invullen, maar dat het van een bedroevende kwaliteit is geworden.

Een ander probleem dat in deze fase van het responsproces kan voorkomen is een zogeheten *contexteffect*. Een *contexteffect* kan plaatsvinden als het begrip van een enquêtevraag wordt beïnvloed door andere informatie in de vragenlijst. In een schriftelijk interview kunnen respondenten bijvoorbeeld de hele vragenlijst zien, wat in een mondelinge en telefonische enquête niet het geval is. Bij webenquêtes bestaat die mogelijkheid om alles te zien wanneer er sprake is van een 'one page' ontwerp. In dat geval kunnen de respondenten ook de hele vragenlijst zien. Bij een 'multi page' ontwerp hangt het af van de onderzoeker of de respondent niet-ingevulde pagina's mag overslaan.

Waarnemings- en communicatieproces

Modes verschillen in de manier waarop de vraag wordt gecommuniceerd of overgebracht. Zo leidt het oplezen van antwoordcategorieën bij telefonische interviews tot een grotere voorkeur voor de laatste antwoordcategorie (*recency effect*) en het zien van antwoordcategorieën (bijvoorbeeld via toonkaarten) tot een grotere voorkeur voor de eerste antwoordcategorie (*primacy effect*). Krysan e.a. (1994) vinden empirische ondersteuning voor deze claim: respondenten in de schriftelijke mode rapporteerden twee keer zo vaak de eerste antwoordmogelijkheid en respondenten in de mondelinge mode waren meer geneigd de laatste antwoordmogelijkheid te kiezen. Meerdere experts wijzen op de aanwezigheid van *primacy* en *recency* effecten:

[Als in] een telefonische vragenlijst je een gesloten vraag stelt, dan hebben mensen de neiging om de laatste uit het lijstje te kiezen. En bij schriftelijk en web kiezen ze voor het eerste item op de lijst. Als je exact dezelfde vragenlijst gebruikt in verschillende modes dan kun je heel goed mode-effecten hebben.

2.5 Kosten

De afnamevormen verschillen in de hoogte van de kosten die met het data verzamelen gepaard gaan. Schriftelijk (en ook internet) is het goedkoopst. Deze vormen vallen voornamelijk goedkoper uit omdat er geen interviewers aan te pas komen. De op een na goedkoopste optie is telefonisch. Hier komen wel interviewers aan te pas, maar de kosten van lokaliseren en contacteren van (potentiële) respondenten zijn lager dan bij de duurste variant: mondeling interview. De geïnterviewde experts van de veldwerkbureaus geven indicaties van de kosten van een interview in verschillende modes.

Het web is heel goedkoop. Misschien zou je met het web, om een aantal mensen over de streep te trekken, nog een hogere cadeaubon kunnen geven. Als je ervan uitgaat dat een gesprek door een enquêteur in het veld snel 80 tot 90 euro kost. Met reistijd en zo. Terwijl bij een website, stel dat je 20 euro geeft aan een respondent, de hostingkosten slechts een paar euro per gesprek zijn.

Als je nu al ziet, dat een interview online 5 euro is en een interview face-to-face 100 euro.

De kwaliteit van een survey mag niet los gezien worden van de kosten. Een heel korte vragenlijst en een zeer klein aantal respondenten zullen de kosten reduceren, maar ook de relevantie en de accuraatheid van de uitkomsten verminderen. Een scheve respons (bv. alleen gemotiveerde, zeer geïnteresseerde respondenten) zal weinig bruikbaar zijn in een onderzoek. Daarom zou het beter zijn om over kosteneffectiviteit te spreken (of value for money) dan over kosten sec. Kosteneffectiviteit is echter veel lastiger te kwantificeren dan kosten. In het volgende hoofdstuk komen we hier nog op terug, wanneer we de voor- en nadelen van het gecombineerde gebruik van de modes (mixed-mode) tegen elkaar afwegen.

2.6 Samenvatting: sterke en zwakke punten van modes

Tabel 2.2 geeft een overzicht van de sterke en zwakke punten van de verschillende modes. De mate waarin voor- en nadelen een rol spelen verschilt uiteraard per onderzoekspopulatie en onderzoeksvraag. Niettemin kan de richting per mode wel op hoofdlijnen worden aangegeven.

Tabel 2.2

Overzicht van de sterke en zwakke kanten van de verschillende afnamevormen

kwaliteit	
relevantie	
lengte vragenlijst	Face-to-face-enquêtes kunnen veel meer vragen bevatten dan telefonische, schriftelijke of internetenquêtes. Meer mogelijkheden voor analyse. Meerdere gebruikersgroepen data.
dekking	
uitsluiting groepen	Telefonisch (geen nummer in telefoonboek), schriftelijk (ongeleterd) en internet enquêteren (ongeleterd, geen vaardigheden, geen internettoegang) sluiten specifieke groepen uit.
non-respons	Non-respons is het hoogst bij web- en schriftelijke enquêtes. Bij mondeling gehouden enquêtes is deze het laagst. Getrainde interviewers kunnen weigerachtige respondenten overhalen. Voordeel bij schriftelijke en webenquêtes is er de mogelijkheid om de enquête op ieder gewenst tijdstip in te vullen, wat non-respons kan reduceren.
meetfout	
vragenlijst	Lange, complexe vragenlijsten kunnen het beste via CAPI afgenomen worden; mogelijkheid van (uitgebreidere) hulpfuncties. Bij CATI kunnen complexe vragenlijsten ook gebruikt worden, maar vragenlijst kan minder lang zijn. Vragenlijsten moeten bij schriftelijke en webenquêtes beduidend korter zijn om uitval en non-respons te vermijden. Bij webenquêtes kunnen wel complexe routingschema's worden toegepast. Bij schriftelijke en – soms – bij webenquêtes kunnen contexteffecten optreden.
interviewer	Interviewer kan de kwaliteit van de gegevensverzameling zowel verbeteren (bijvoorbeeld hulp, paradata) als verslechteren (sociale wenselijkheid bij gevoelige vragen).
waarnemings- proces respondent	Afhankelijk van het waarnemingsproces zijn er verschillende fasen in het responsproces die tot een grotere kans op meetfouten leiden (primacy / recency effecten). Verschillen in comprehension, retrieval, judgment en report (hulp, herinneren, uitleg, satisficing, acquiescence).
verwerkingsfout	Bij computerondersteunde varianten is de kans op verwerkingsfouten kleiner door afwezigheid van data-entry, consistentiechecks. Interviewer verkleint kans op invulfouten.
timeliness	
	Bij computerondersteunde afnamevormen zijn de data sneller beschikbaar. De doorlooptijd van telefonische enquêtes is langer dan die bij mondelinge enquêtes. Bij schriftelijke en webenquêtes is het afhankelijk van het aantal reminders en de tijd die een respondent wordt toegestaan. Voordeel is dat iedereen tegelijk kan worden benaderd. Over het algemeen is een webenquête sneller dan een schriftelijke enquête.
kosten	
	Mondelinge enquêtes zijn het duurst, gevolgd door telefonische en schriftelijke enquêtes. De interviewkosten voor webenquêtes zijn het laagst.

3 Mixed-mode onderzoek: voor- en nadelen

In hoofdstuk 2 zijn de voor- en nadelen van de verschillende afnamemethoden besproken. Dat geeft al een beeld van de mogelijk- en moeilijkheden van een survey waarin verschillende methoden worden gecombineerd. Wat de voor- en nadelen zijn van het gecombineerd gebruik van afnamemodes is onderwerp van dit hoofdstuk. We beginnen met een korte toelichting op verschillende vormen van mixed-mode onderzoek.

3.1 Soorten mixed-mode design

Mixed-mode in (her)benadering versus mixed-mode in dataverzameling

In de benadering van respondenten kunnen verschillende modes worden ingezet⁶ (mixed-mode design in de benadering of een mixed-mode system; De Leeuw 2005). Sommige groepen kunnen bijvoorbeeld telefonisch benaderd worden, andere mondeling. Er kan ook gebruikgemaakt worden van meerdere modes gedurende het data verzamelen. Dit kan plaatsvinden op verschillende manieren.⁷ Dit rapport gaat in op het gebruik van meerdere modes van dataverzameling bij groepen van respondenten tijdens één (cross-sectioneel) onderzoek.

Concurrent versus sequentieel mixed-mode onderzoek

In mixed-mode onderzoek kan er gevarieerd worden in het *moment* waarop verschillende modes worden ingezet in het veldwerk. Meerdere modes kunnen gelijktijdig in het veld ingezet worden. Dit wordt wel als een *concurrent* mixed-mode design aangeduid. Er kan ook voor worden gekozen om de dataverzameling te beginnen met de ene mode en na bepaalde tijd non-respondenten uit de eerste mode of een nieuwe groep steekprofeenheden te (her)benaderen met de volgende mode. Zo'n design wordt een *sequentieel* mixed-mode design genoemd. Een sequentiële mixed-mode benadering heeft theoretisch gezien twee voordelen: een hogere respons en een verbeterde responsamenstelling. Het idee daarbij is dat respondenten die je in de ene mode niet kunt 'vangen', wel kunt bereiken in een andere mode. Dat zou dus in theorie de respons moeten verhogen en van de samenstelling van respondenten een betere weergave opleveren van de doelpopulatie.

Toewijzing van respondenten aan een mode

De toewijzing van steekprofeenheden aan een bepaalde mode kan op verschillende manieren plaatsvinden. In een sequentieel design hebben steekprofeenheden geen keuze in welke mode ze zullen 'vallen'. De modes worden één voor één afgewerkt. In een concurrent design kunnen (1) respondenten zelf kiezen in welke mode het interview zal worden afgenomen. Dit geeft mensen de gelegenheid zelf hun voorkeursmode te kiezen, maar tegelijkertijd geeft dit hun een extra gelegenheid om een deelnameverzoek te weigeren (Dillman en Tarnai 1991). Steekprofeenheden kunnen ook aan een mode toegewezen worden (2). Een toewijzing van een mode aan een steekproef-

eenheid door de onderzoeksorganisatie kan ook weer op twee manieren plaatsvinden. Respondenten kunnen (a) aselekt aan een mode toegewezen worden. Dit heeft als voordeel dat respons- en mode effecten beter uit elkaar kunnen worden gehaald (waarover later in dit hoofdstuk meer). Nadeel is dat een toevallige toewijzing van modes inefficiënt kan zijn (bijv. oudere respondenten, die vaker niet over internet beschikken, de websurvey aanbieden, met hogere non-respons tot gevolg). Een andere manier is om steekproefeenheden toe te wijzen aan een mode op basis van het (b) gebruik van achtergrondinformatie. Hierbij wordt getracht de responskansen te maximaliseren door aan te sluiten bij de *ingeschatte* modevoorkeur van steekproefeenheden. In hetzelfde voorbeeld als hiervoor zouden dus oudere steekproefeenheden niet en jongere juist wel aan de webmode worden toegewezen. De gevolgen zijn omgekeerd aan die van een aselekte toewijzing: responskansen worden in dit geval waarschijnlijk geoptimaliseerd, maar responsstelling en meetfouten zijn moeilijker uit elkaar te halen. Bovendien kan een niet-toevallige toewijzing van modes tot een selectieve respons leiden. Je kunt in zo'n geval bijvoorbeeld een selecte groep missen die juist niet in het door de onderzoeksorganisatie gekozen modeprofiel past.

Ook de experts wijzen op de trade-off tussen selecte en aselekte toewijzing van modes aan respondenten. De voordelen van een selecte toewijzing zijn te vinden in het vergroten van de responskansen. Wel moet je:

bij een selecte toewijzing dus eigenlijk al van tevoren [moeten] kunnen voorspellen op basis van achtergrondkenmerken van mensen wat voor hen de beste modus van dataverzameling is. De vraag is dus of je die informatie al van tevoren hebt.

Of zoals een andere expert het verwoordt:

Je moet dus weten of men internet heeft en welke taal men wil ontvangen, of dat je iets schriftelijk moet sturen en in welke taal. Je kunt dan ook voorstellen om een interviewer langs te sturen. Ik denk dat dat best lastig wordt. Dan moet je heel goed na gaan denken over het script en hoe je dat doet. Je kunt natuurlijk voor een deel proberen om vooraf een goede inschatting te maken.

Behalve dat het dus lastig is om de responskansen bij de toewijzing van modes aan respondenten te optimaliseren, kleeft er nog een ander nadeel aan een selecte toewijzing van modes:

Het nadeel van een selecte toewijzing op een bepaald type mensen is dat het dan lastig is om een mode effect te herkennen.

De voordelen van een aselekte toewijzing liggen vooral daarin dat er gemakkelijker kan worden onderzocht of verschillen in schattingen door populatie- of mode effecten zijn ontstaan.

*Dan heb je werkelijk een random sample uit die groep twee keer getrokken zonder teruglegging. Dan heb je volgens mij de meest vergelijkbare groep en kun je later kijken naar hoe het zit als je mensen met elkaar gaat matchen en allerlei typen mensen hebt bereikt, schriftelijk of face-to-face.
Vanuit experimenteel oogpunt is dat het mooiste.*

3.2 Gevolgen van mixed-mode onderzoek

Een mixed-mode design heeft de potentie om de voordelen van de modes te gebruiken en de zwakheden van iedere mode afzonderlijk te compenseren (De Leeuw e.a. 2008). Vaak is er een expliciete trade-off tussen kosten en kwaliteit (cf. Leeuw e.a. 2008; Biemer en Lyberg 2003). Hierna worden de gevolgen van de keuze voor mixed-mode onderzoek besproken voor verschillende aspecten van de survey, te beginnen met consequenties voor de omvang en samenstelling van de respons.

3.2.1 Respons en responsamenstelling

In de literatuur wordt een hogere en betere responsamenstelling als potentieel voordeel van een mixed-mode design gezien. Het (op maat aanbieden) van een mode kan de responskansen van steekproefeenheden vergroten. Groepen die minder geneigd zijn deel te nemen aan een mondelinge mode, nemen wellicht wel deel aan een websurvey. De uiteindelijke responsamenstelling zou dan beter overeen kunnen komen met de populatie. Ook enkele experts zien in het gebruik van verschillende modes kansen om de respons te verbeteren.

Een voordeel van mixed-mode kan zijn dat je misschien een aantal groepen jongeren wel krijgt bij een webvragenlijst of een schriftelijke vragenlijst omdat het makkelijker is dan een interviewer thuis.

Wij [denken] ook wel eens aan dat wij de respons kunnen verhogen omdat je de mensen die je in CAP1 mist, de nerds, de yuppies, die krijg je wel met web. Dus je krijgt wel wat meer respons binnen.

Als je bij mixed-mode zegt dat je eerst face-to-face gaat doen en de mensen die dat echt niet willen daarna nog eens opbelt om te vragen of ze het misschien aan de telefoon willen doen of hen daarna nog een vragenlijst toestuurt dan kan je zeggen dat wat je binnenkrijgt extra is.

Eén respondent verwoordt het kernachtig:

Wie je in de ene mode niet bereikt, bereik je misschien wel in een andere mode.

Voor de nieuwe SIM zou dit een veelbelovend perspectief kunnen zijn. De totale respons van SIM2006 schommelt rond de 50%. Verhoging daarvan is dus zeker het nastreven waard. Een ander punt is dat in het onderzoek Leefsituatie Allochtone Stedelingen (LAS2004)⁸ en de SIM2006 met name jongeren en jongvolwassenen zijn

ondervertegenwoordigd in de respons (Schothorst 2005; Dagevos e.a. 2007). In die categorie lijkt het gebruik van bijvoorbeeld een websurvey kansrijk.

Toch moeten de verwachtingen over een verbetering van de responshoogte en -samenstelling bij toepassing van mixed-mode getemperd worden. Onlangs hebben onderzoekers van het Centraal Bureau voor de Statistiek (CBS) (Banning e.a. 2009) onderzoek gedaan naar de gevolgen van een mixed-mode aanpak voor de respons(samenstelling) (en -dekking). Er is gekeken naar verschillende edities van de Veiligheidsmonitor,⁹ naar het onderzoek Informele economie 2006 en naar de Enquête Beroepsbevolking (EBB2005).¹⁰ De studie wijst uit dat over het geheel genomen de omvang van de respons niet hoger is wanneer gebruik wordt gemaakt van een mixed-mode aanpak. De respons is gelijk of lager. Uitzondering vormt de EBB2005 waar met mixed-mode wel een hogere respons wordt gerealiseerd. Hetzelfde onderzoek (Banning e.a. 2009) laat zien dat respons-samenstelling niet hoger, maar eerder lager uitvalt. De mixed-mode benadering leidt tot meer van hetzelfde soort respondenten; slecht responderende groepen blijven slecht responderen. Dit geldt zelfs wanneer er een hogere respons wordt behaald. 'Gemakkelijke' respondenten die je al in de eerste mode had, krijg je er ook bij in de volgende mode. Specifiek voor jongeren wijzen Banning e.a. nog op de beperkingen van Computer Assisted Telephone Interviewing (CATI), omdat slechts weinig jongeren met hun telefoonnummer geregistreerd zijn. Hun aanbeveling is dat men ten aanzien van slecht responderende groepen de oplossing niet in de modes zoekt, maar in een aangepaste benaderingsstrategie, los van de keuzes van modes.

Het theoretische voordeel van verbetering van de respons-samenstelling wordt ook door geïnterviewde experts betwijfeld.

Maar ook daar moet je natuurlijk altijd mee oppassen. Om wat voor een respons je binnenkrijgt en dat dit niet meer van hetzelfde is. Het zou kunnen hoor, maar daar heb ik geen harde cijfers van.

Dan is het altijd nog de vraag of het de compositie van je uiteindelijke steekproef verbetert als je daarna telefonisch gaat enquêteren. Het kan ook best zijn dat hij daardoor achteruitgaat. Waarschijnlijk ga je de ouderen missen en mensen met taalproblemen, die zullen dus niet random zijn.

Onderzoek van Beukenhorst en Wetzels (2009) toont specifiek voor niet-westerse migranten (en personen uit de grote stad) aan dat de respons-samenstelling niet verbetert bij de introductie van meerdere modes.

In het kader van de gehouden interviewronde is met een onderzoeker gesproken die verantwoordelijk was voor de uitvoering van een survey onder tweedegeneratie Turkse en Marokkaanse Nederlanders in Amsterdam en Rotterdam. Dit is vanwege de jonge leeftijd en het feit dat ze in grote steden wonen bij uitstek een moeilijk responderende groep. In dit onderzoek was het aanvankelijk de bedoeling met een websurvey te werken.

Hiermee is gestart, maar de ervaringen waren zo slecht dat ze hiervan uiteindelijk hebben afgezien.

Dan konden ze gewoon inloggen met een code, en konden ze 10 of 20 euro aan een cadeaubon verdienen en het gewoon online doen. Dat was dramatisch, de respons was nog lager, dus dat moet je gewoon niet doen.

Naast de lage respons maken respondenten via websurveys de lijst niet af:

Er zijn veel te veel mensen die het onderweg afbreken. Die zijn er na een halfuur wel klaar mee. En dan houd jij ze dus niet vast. Dat was het grote probleem bij cAW1. Mensen hebben dan de vrijheid, maar dan haken ze dus ook af. Dan zijn ze het na twintig minuten zat en zeggen: 'doei'.

Een belangrijk deel van het onderzoek is uitgevoerd via Computer Assisted Personal Interviewing (CAPI) met een drop-off vragenlijst (achtergelaten na interview). De veronderstelling was dat een korte vragenlijst de respons zou verhogen. Door middel van een drop-off vragenlijst zou dan wel voor meer onderwerpen informatie verzameld kunnen worden. Hiermee waren de ervaringen evenmin erg bemoedigend. Men heeft zeer vaak en lang achter de respondent aan moeten zitten om de drop-off vragenlijst terug te krijgen. Uiteindelijk heeft ongeveer 70% tot 80% de drop-off lijst geretourneerd, met veel inspanning en lange (doorloop)tijd.

Dus het was puur drop-off, dat hadden we geprobeerd. En wat is dan de moeilijkheid: sturen mensen dat ding in? En maar recallen, en maar recallen. En natuurlijk de beloning uitstellen tot ze het hebben ingestuurd. Nou, dan krijg je toch nog 70 tot 80% terug, maar je mist toch ook weer een deel van de informatie voor een deel van de groep. Je moet enorm lange adem hebben, om dan dat laatste stuk binnen te halen. [...] En dat je dus héél vaak moet rappelleren. Je moet er maandenlang achteraan gaan. Het vertraagt uiteindelijk toch ook. Ik vraag me af of we er nou zoveel winst mee hebben geboekt.

Ook een andere informant heeft slechte ervaringen met drop-off en dan in het bijzonder bij niet-westerse migranten:

Dat je daar dan ook weer mensen verliest die zeggen tegen de interviewer dat ze het gaan invullen en het dan toch niet doen. Het is soms moeilijk om nee te zeggen, misschien ook wel cultuurgebonden, dat weet ik niet. Dat kwam met name bij de niet-westerse allochtonen voor.

In de survey onder de tweedegeneratie Turkse en Marokkaanse Nederlanders is op een gegeven moment besloten de drop-off vragenlijst niet langer te gebruiken en te kiezen voor een combinatie van face-to-face interviews waarbij een deel van de lijst door de respondent zelf werd ingevuld (CAPI en CASI; self-completion). Dat bleek uiteindelijk het beste te werken. De doelgroep van deze survey is het Nederlands voldoende machtig

om een deel van de vragenlijst zelfstandig in te vullen. Bovendien is de interviewer beschikbaar om toelichting te geven.

CASI kun je zo makkelijk maken. Het enige is dat de taal goed moet zijn. Kijk, er zit een interviewer bij, dus die kan desgewenst en desgevraagd uitleg geven over de vraag. Dus we hebben CASI alleen maar gedaan met een interviewer erbij die opheldering kan geven over vragen. [...] En bij CASI kan iemand vragen stellen, tenzij die zo verlegen is dat hij niet om opheldering durft te vragen. De sleutel is toch een goede interviewer.

De les van deze survey is, zo stelt deze respondent, dat wanneer je eenmaal binnen bent het interview moet afwerken. Gebruik van drop-off en van websurveys is niet voor herhaling vatbaar. Verder is de beschikbaarheid van voldoende tijd en inzet van goede interviewers cruciaal.

De les die ik heb geleerd: je hebt veel tijd nodig, veel contactpogingen nodig. En een groep interviewers die tegen een stootje kan, omdat het gewoon zo moeilijk is om binnen te komen. Dat is één. En binnenkomen en een afspraak maken en dan ook die afspraak ten uitvoer brengen is de grootste hobbel die genomen moet worden. Dus als je binnen zit, doe het dan meteen. Dat zou ik eigenlijk toch aanraden hoor. Zo van: zorgen dat je binnenkomt, zorg ervoor dat je daar gewoon ruim de tijd voor neemt. [...] Ik zou op deze groep aanraden: 'Neem de tijd om binnen te komen en zet er alles op in.' En als je binnen bent alles doen. En dan combineren van CAPI/CASI is prima. Maar niet dingen achterlaten. Ik zou het eigenlijk niet doen, want je hebt gewoon heel veel tijd nodig om dat veldwerk te doen. En hele goede interviewers heb je nodig. En de grootste moeilijkheid is: binnenkomen. Dus als je binnen bent, alles doen. Dat is eigenlijk kort en goed mijn belangrijkste advies.

Ook een andere respondent legt de nadruk op binnen zien te komen en wanneer dat is gelukt: alles afronden.

Wij zijn ervan overtuigd dat het belangrijkste is om het contact te leggen met die mensen. Als ze eenmaal meedoen dan kun je het beste meteen maar alles afmaken. Daar moet je de meeste energie in doen.

Voorgaande ervaringen worden bevestigd door weer een andere respondent met goede ervaringen met de mondelinge afnamemethode. De respons is hoger, minder selectief en je weet zeker dat de juiste persoon de vragen beantwoordt.

Bij de mondelinge variant krijgen we gewoon een veel hogere respons. Bij de schriftelijke respons kregen we 10 of 11% terug. Dat is gewoon heel laag [...] met een heel selectieve respons [...]. Bij de mondelinge respons merkten we, met het geselecteerde databestand, dat het gewoon vele malen hoger lag. Je kwam dan gewoon binnen en kreeg de juiste persoon te spreken en de respons was ook hoger omdat we gewoon ook gemakkelijker binnenkwamen voor het mondelinge gedeelte. Op deze manier was ook de urgentie duidelijker, waardoor je meer respons kreeg.

3.2.2 Meetfouten en vergelijkbaarheid tussen de modes

Een groot nadeel van een mixed-mode onderzoek zit in de vergelijkbaarheid van modes. Eenzelfde survey kan verschillende uitkomsten hebben afhankelijk van de mode waarin de survey is afgenomen (Van der Laan en Van Nunspeet 2009). Zoals in hoofdstuk 2 beschreven verschillen modes in de mogelijkheden om toegang tot de populatie te verkrijgen (dekking), variëren responscijfers tussen modes (non-respons) en kan de gekozen mode invloed hebben op de meting (meetfout). Hierbij kent iedere mode zijn eigen mode effect. In een mixed-mode onderzoek is het ontwarren van mode effecten geen sinecure, om niet te zeggen soms onmogelijk. Met name de effecten van modes op de responsstelling en op de meting zijn bijna niet uit elkaar te halen.

Inhoudelijke verschillen in survey uitkomsten kunnen dus veroorzaakt worden doordat in de verschillende modes andere groepen zijn bereikt of doordat respondenten in verschillende modes andere antwoorden geven op survey vragen. Dit probleem wordt ook door alle experts onderkend.

Ik denk dat het lastig is als bepaalde groepen mensen, bijvoorbeeld oudere allochtonen die geen Nederlands spreken of analfabeet zijn, mondeling geïnterviewd worden en andere schriftelijk geïnterviewd worden. Dan krijg je verschillende groepen die ook nog eens op verschillende manieren geïnterviewd worden. Als je dat in verband brengt met sociale wenselijkheid kan dat heel vreemde effecten hebben; bijvoorbeeld als je uit mondelinge interviews meer sociaal wenselijke antwoorden krijgt. Dan kan je heel vreemde relaties krijgen.

Dan krijg je dus wel weer een soort gesofisticeerde groep die de webvragenlijsten invult. Als je dat dan vergelijkt met een andere groep die het schriftelijk invult, en weer een andere groep die het mondeling invult – en dat zijn allemaal verschillende groepen –, dan weet ik niet wat je aan het eind zit te vergelijken. Als je bij de eerste en tweede generatie wilt vergelijken in hoeverre ze terug willen naar hun oude land, hun land van herkomst of tevreden zijn in Nederland en ze worden op een andere manier geënquêteerd en er zitten verschillen in de methode of de sociale wenselijkheid, of hoe snel je door een vragenlijst heen ragt, dan weet je eigenlijk drie keer niks.

Zijn ze nu echt progressiever of heb je alleen maar een sociaal wenselijk effect gehad? Dat de jongeren het wel online hebben durven in te tikken, maar zo gauw die interviewer voor hun neus zit toch nog steeds zeggen dat alcohol niet mag.

Het responseeffect en meeteffect zijn heel lastig, en in een niet-toevallige toewijzing van respondenten aan een mode eigenlijk niet uit elkaar te halen. Een recente studie van Beukenhorst en Wetzels (2009) naar mode effecten bij een sequentieel mixed-mode onderzoek van de Nederlandse Veiligheidsmonitor toont dit duidelijk aan.

De mate waarin een mixed-mode effect zich voordoet, is niet eenduidig vast te stellen. Het varieert in elk geval welke vraagstelling en welke populatie onderwerp van onderzoek is (De Leeuw 2009). Wel kan in algemene zin worden gesteld welke modes

het meest van elkaar verschillen. Pierzchala (2006) vat de factoren die de meeste invloed kunnen hebben op een mixed-mode effect als volgt samen: verbale of visuele presentatie, met of zonder interviewer en ten slotte computerondersteunde vragenlijsten of niet-computerondersteunde vragenlijsten. Hoe meer modes van elkaar verschillen ten opzichte van deze factoren, hoe waarschijnlijker en groter een mixed-mode effect (cf. Cobben 2009). Webenquêtes en mondelinge enquêtes zijn beide computerondersteunde interviews, maar verschillen in de presentatie (visueel versus verbaal) en in de aanwezigheid van een interviewer. Web- en mondelinge enquêtes verschillen dus op twee van de drie factoren van elkaar. Mondelinge en schriftelijke enquêtes verschillen op alle drie de factoren. Voor een schriftelijke enquête geldt namelijk dat deze gebruikmaakt van een visuele presentatie, er geen interviewer aanwezig is en het een niet-computerondersteund interview is. In dit geval is dus volgens Pierzchala (2006) de waarschijnlijkheid van het optreden van een verschil in meting tussen modes het grootst.

De keuze voor een bepaalde mix van modes heeft invloed op het mixed-mode effect. Er is echter geen consensus over welke mix tot geen of tot de minste mixed-mode effecten zal leiden, ook al omdat dit afhankelijk is van de vragenlijst en doelgroep. De geïnterviewde experts zeggen over deze kwesties het volgende:

Het minst ingrijpend is van telefonisch naar face-to-face, omdat je daar heel specifieke vraagvormen hebt. Zowel online als schriftelijk kan heel gemakkelijk achter elkaar, omdat je dan bij beide het visuele hebt. Van face-to-face naar online heb je dat de routing verandert en bij schriftelijk kun je een heel eenvoudige routing hebben. Ik weet niet hoe gecompliceerd de routing is, maar als de routing erg gecompliceerd is, dan is het misschien beter om het niet per post te doen.

Als je binnen face-to-face of binnen telefoon dingen op een andere manier vraagt, dan krijg je ook andere uitkomsten, andere verdelingen. Als je een vraag heel erg gaat aanpassen omdat hij opeens op het scherm moet. De ergste aanpassingen zijn natuurlijk als je naar telefoon gaat. Voor telefonische vragenlijsten moet je het meeste doen, omdat je daar alleen maar audio hebt.

Volgens mij wordt het allerbelangrijkste effect veroorzaakt door het al of niet aanwezig zijn van een interviewer. Ik kan me voorstellen dat je bij CAPI en CATI nog een beetje kan schipperen. Maar als je een CAPI tegenover web of mail zet, krijg je grote verschillen.

Veel is echter nog niet helder. Diverse experts geven aan dat nog lang niet alles duidelijk is over het effect op de meting bij het gebruik van meerdere modes.

Mixed-mode betekent een methodewisseling wat met je vergelijkbaarheid te maken heeft in de tijd. En ook binnen je meting kan het inzetten van verschillende methodes invloed hebben op de antwoordpatronen. Het is dus door de tijd heen, en binnen je meting dat het onvergelijkbaar wordt. De één zal aan de ene mode meedoen en de ander aan de andere mode. Wie zal zeggen wat dat voor effecten heeft? Daar is te weinig over bekend. Sowieso doen wij niet zoveel met mixed-mode omdat je het natuurlijk liever niet doet als het niet hoeft. Daar worstelt ook iedereen mee. Vandaar ook het grote aantal bijeenkomsten

over de gevolgen van mixed-mode aanpakken. Volgens mij is gewoon nog niet goed in kaart gebracht wat nou precies de gevolgen zijn van de ene en de andere methode.

3.2.3 Analyse van mixed-mode gegevens

Wanneer dus bij een survey verschillende modes worden ingezet bestaat er een gereede kans op een mixed-mode effect. Voor het verkrijgen van een kwalitatief goed bestand is het van belang om deze mode effecten – hoe moeilijk dit ook is – te ontwarren en te kwantificeren in termen van mode afhankelijk meetfouten (Van der Laan en Van Nunspeet 2009). Er zijn verschillende manieren om mixed-mode effecten op te sporen. Propensity score matching wordt genoemd (Lugtig en Lensvelt-Mulders 2007), experimenten al dan niet tijdens het veldwerk en cognitieve interviews als een soort van pretest.

Het idee is in ieder geval als volgt. Als je twee groepen mensen hebt en een deel daarvan heeft een schriftelijk onderzoek gedaan en een andere groep face-to-face, dan heb je eigenlijk populatie-effecten omdat je verwacht dat de mensen die schriftelijk antwoorden een ander type mensen is dan de mensen die face-to-face willen komen. Je verwacht dus een betere dekking van de totale populatie en alle soorten en maten die daarin voorkomen. Maar het nadeel is dat je gewoon te maken kunt hebben met ontzettende mode effecten. Dat wil je ook niet of je wilt minstens weten of die mode effecten er zijn. Een manier om dat te weten te komen is om te kijken welke mensen binnen die samples op elkaar lijken. Dat zijn de mensen die ik dan 'per toeval' noem. Deze zijn per toeval ergens aan toegewezen of terechtgekomen in een groep. Als je ze gevraagd had voor face-to-face, dan hadden ze ook meegedaan en als je ze gevraagd had voor paper and pencil, dan hadden ze ook meegedaan. Die mensen die met een heleboel kenmerken sterk op elkaar lijken, die zouden ook dezelfde antwoorden moeten geven. Dat is de aanname, want die mensen staan voor een bepaalde subgroep van de populatie die er hetzelfde over zou moeten denken. Die kun je matchen. Dus als die antwoorden van elkaar afwijken, dan kun je zien dat je te doen hebt met een mode effect. Dat is dat propensity matchen, waar je nu veel over hoort.

Behalve propensity score matching, dat oorspronkelijk uit de quasi-experimentele literatuur komt (Rosenbaum en Rubin 1983), kunnen er experimenten uitgevoerd worden.

Nou ja, eigenlijk is het naar mijn idee de beste oplossing om dan een kleine selectie van een paar vragen twee keer te stellen, bij dezelfde mensen, zowel in het face-to-face gedeelte als in het schriftelijk gedeelte. Dat is de beste manier om dat te controleren, zowel qua antwoordverdeling als kwaliteitschatters kan je dan iets doen. Je kunt het nog iets mooier maken natuurlijk in design, maar sowieso kan je de antwoorden vergelijken. Je hebt dezelfde mensen, dus je hebt uitgesloten dat er iets anders is en dan kan je zien wat ze face-to-face antwoorden en wat ze schriftelijk antwoorden. Je hebt wel een klein volgorde-effect dan.

Een experiment kan ook voor het veldwerk plaatsvinden.

Je kunt een experiment opzetten. Maar of je dat nou midden in het veldwerk moet doen? Misschien moet je nog wel een klein testje vooraf doen. Trek eens honderd allochtonen, of duizend en kijk eens met een klein vragenlijstje met wat typische vragen erin van het soort dat je in hoge mate zou willen stellen. In een onderzoekje van tien minuten of zo. Dan weet je nog steeds niet hoe het gaat met lange vragenlijsten, maar daar kun je wel iets meer achter komen. Mode effecten. Dan doe je er driehonderd per telefoon en driehonderd face-to-face en driehonderd met web. Dan ga je ook een idee krijgen hoeveel allochtonen wel of geen internet hebben of papieren vragenlijsten kunnen lezen.

Ten slotte kan een eventueel mixed-mode effect worden opgespoord via een pretest.

Ik zou in ieder geval een aantal gevallen random modes aan mensen toewijzen en dan eens kijken hoe dat gaat en of je die vragenlijst kunt afnemen. Dat kun je in een soort real life situatie doen waarbij je dat echt aan mensen toewijst. Dan kun je kijken hoe de respons is. Je kunt ook mensen van de straat halen en hen vragen deze vragenlijst eens in te vullen of te interviewen. Dan heb je meer een soort cognitieve laboratoriumsituatie en kun je kijken wat de verschillen zijn. Je kunt het mensen eerst schriftelijk laten doen en dan daarna een restgroep mondeling omdat het niet anders kan. Wat zijn dan de verschillen?

3.2.4 Vragenlijsten: dezelfde of een andere lijst in verschillende modes?

Een van de grote problemen bij het opzetten van een mixed-mode onderzoek is dat verschillende modes volgens een verschillend waarnemingsproces verlopen en verschillende eisen aan de respondent stellen (zie hoofdstuk 2). Een lijst die voor een bepaalde mode ontwikkeld is, kan niet zomaar worden gebruikt voor een andere mode. Dat betekent dat bij overgang van een uni-mode naar een mixed-mode design de vragenlijst grondig zal moeten worden herzien (al is het maar omdat in telefonische, schriftelijke en internetinterviews de vragenlijst veel korter zal moeten zijn dan in een face-to-face survey).

Vragenlijsten kunnen op verschillende manieren worden vormgegeven in een mixed-mode design en kunnen daarmee in principe een (gedeeltelijke) oplossing bieden om een mixed-mode effect te vermijden. Er zijn in de literatuur (Dillman 2007; zie ook Cobben 2009) drie manieren van vragenlijstontwikkeling voor een mixed-mode design geïdentificeerd, namelijk het *mode specific design*, het *generalised mode design* en het *uni-mode design*.

Vragenlijsten kunnen afzonderlijk voor modes worden ontwikkeld. Hierbij wordt dan optimaal gebruikgemaakt van de voordelen van iedere mode afzonderlijk en worden de beperkingen van iedere mode afzonderlijk geminimaliseerd. Dit wordt een *mode specific design* genoemd. Dit heeft voordelen voor bijvoorbeeld de gebruiksvriendelijkheid van de vragenlijst, maar tast de vergelijkbaarheid van modes aan. Een CAPI-interview

biedt bijvoorbeeld veel mogelijkheden om een ingewikkelde routing in de vragenlijst te implementeren. In een schriftelijke vragenlijst zijn de mogelijkheden om met routings te werken een stuk beperkter.

Een tweede mogelijkheid om de vragenlijst vorm te geven is dan ook om de modes gelijk te maken, een zogeheten *generalised mode design*. Dit houdt in dat gestreefd wordt om respondenten in elke mode een gelijke stimulus te geven. In dit geval wordt geprobeerd de beschikbare assistentie tussen modes gelijk te trekken. Een ander voorbeeld is de verschillen tussen verbale en visuele communicatie te minimaliseren om zo primacy en recency effecten te beperken. Ook kan ervoor worden gekozen om de laptop in een CAPI-interview aan de respondent te geven bij een blok met sensitieve vragen. Dit zal het gevoel van privacy bij de respondent vergroten en de vertekening als gevolg van sociale wenselijkheid zal dan waarschijnlijk gereduceerd worden tot het niveau van een self-completion interview.

Ten slotte kan ervoor worden gekozen een van de modes te optimaliseren (meestal de hoofdmode) en de andere modes hieraan aan te passen. Dit wordt een *uni-mode* of *uniform mode design* genoemd (Dillman 2000; zie ook Cobben 2009).

De experts zien verschillende knelpunten bij het gebruik van vragenlijsten in de verschillende modes.

Je probeert dus richtlijnen te geven over hoe je vragen zodanig moet bouwen dat de mode effecten zo gering mogelijk zijn. Dan zou je dus een soort uni-mode design moeten maken van je vragenlijst en die zou dan minst mode effect gevoelig zijn. Of dan mixed-mode effecten helemaal worden uitgesloten is echter onduidelijk.

Je begint met een vragenlijst waarvan je zegt dat je alles zo gaat formuleren dat hij in de drie modes op dezelfde manier afgenomen kan worden. Dat kan soms lastig en vervelend zijn omdat het niet bij een mode past. Maar dan krijg je wel hetzelfde. Het is natuurlijk ook nog de vraag of het in de verschillende modes nog wel hetzelfde betekent als iets hetzelfde is.

Omdat hier ook de originele vragenlijst voor zal moeten worden aangepast, heeft dit wel gevolgen voor de vergelijkbaarheid van onderzoeken over de tijd. Ook kan het gelijk maken van stimuli in modes ertoe leiden dat het minder prettig wordt een vragenlijst in te vullen.

Het probleem is dat de schriftelijke vragenlijst zoveel mogelijk gelijk wordt gehouden met de online vragenlijst, terwijl online heel veel mogelijkheden biedt tot routing en selecties.

Dat wilden ze bij die online vragenlijst ook. Daardoor zitten er bij de online vragenlijst heel veel onnodige vragen over dingen die al veel eerder zijn gevraagd en zouden kunnen worden overgeslagen. Dat wekt ergernis op bij mensen. Mensen vonden het een irritante vragenlijst.

Voor de SIM zou het aanpassen van de vragenlijst een majeure operatie zijn. Het geschikt maken van een lijst die opgezet is voor face-to-face onderzoek naar een voor web en andere modes equivalente lijst zal een forse (tijds)investering vergen. Hiervoor is al kort genoemd dat er niet aan valt te ontkomen de vragenlijst drastisch in te korten (of de lijst op te knippen en de delen aan verschillende respondenten voor te leggen). Iedere mode kent een maximale lengte voor een vragenlijst. Wanneer een webvragenlijst langer dan tien tot twintig minuten duurt, leidt dit tot meer non-respons en meer satisficing (zie hoofdstuk 2).

Telefoon zal misschien 30 minuten kunnen duren voordat respondenten gaan protesteren, maar zal zeker niet de duur kunnen hebben van 45 minuten tot een uur gesprekstijd zoals bij de huidige SIM-lijst. Dit betekent dat er een flink aantal onderwerpen geschrapt moet worden. Daarnaast zou overwogen kunnen worden om uit te gaan van een lijst met kernvragen die in elk geval afgewikkeld moet worden. Vervolgens zouden personen gevraagd kunnen worden of ze op hetzelfde of op een later moment resterende vragenblokken zouden willen beantwoorden via het web of een schriftelijke lijst. Ervaringen uit het onderzoek onder tweedegeneratie Turkse en Marokkaanse Nederlanders, dat eerder in dit hoofdstuk is besproken, stemmen echter niet erg optimistisch.

3.2.5 Trendgegevens: gevolgen voor de vergelijkbaarheid in de tijd

De SIM2010 past in een lijn surveys die vanaf het einde van de jaren tachtig in opdracht van de rijksoverheid zijn uitgevoerd. Het gaat hier om de surveys Sociale Positie en Voorzieningsgebruik van Allochtonen (SPVA's)¹¹ (1988, 1991, 1994, 1998, 2002) en de SIM2006. Hiermee is in Nederland een unieke serie metingen verricht onder niet-westerse migranten. Deze is in het bijzonder van belang voor het vaststellen van trends in de sociaal-culturele positie van niet-westerse groepen. In het laatste Jaarrapport integratie 2009 (Gijsberts en Dagevos 2009) zijn hiervan verschillende voorbeelden te vinden (o.a. sociale contacten, beheersing Nederlands, opvattingen en religie). De vergelijkbaarheid van deze metingen wordt door de introductie van nieuwe modes bemoeilijkt. Zo zijn de mode afhankelijke vertekeningen niet langer consistent (zoals sociaal wenselijke antwoorden vanwege de aanwezigheid van een interviewer). Ook komt de vergelijkbaarheid mogelijk in het gedrang doordat vanwege inkorting van de vragenlijst niet alle thema's kunnen worden bevroegd. Tegelijkertijd rijst de vraag of een heroriëntatie op de wijze van bevraging over deze onderwerpen niet gewenst is. De aanwezigheid van een interviewer lijkt zeker bij bepaalde items over sociaal-culturele integratie ook een versturende factor. Dit zou ertoe kunnen leiden dat deze vragen zonder een interviewer worden afgenomen. Behalve de al genoemde modes is ook self-completion in een CAPI een mogelijke vorm. In het algemeen kan worden gesteld dat het gebruik van mixed-mode onderzoek problematisch is voor de vergelijkbaarheid in de tijd. Iedere nieuwe meting zal namelijk leiden tot een unieke samenstelling van modes en dus tot een unieke samenstelling van responsgroepen en meetfouten.

3.2.6 Kosten en doorlooptijd

Face-to-face onderzoek is kostbaar, en dit is in het bijzonder het geval bij niet-westerse respondenten. Een belangrijke reden voor mixed-mode onderzoek zou dan ook een kostenreductie kunnen zijn. Die wordt vooral bewerkstelligd door het verminderen van de inzet van interviewers. Ook zou de doorlooptijd van het veldwerk verkorten, omdat de interviewercapaciteit minder een beperkende factor is.

Diverse experts zien de kostenreductie van mixed-mode als een duidelijk voordeel.

Maar we hebben toen wel in dezelfde pilot een vergelijking gemaakt tussen alles face-to-face en een combinatie van telefonisch en face-to-face en dat laatste is wel een enorme kostenbesparing.

Maar als je kijkt in termen van geld dan is online het allergeedkoopste. Je stuurt mensen een briefje en een link. De Veiligheidsmonitor wordt nu zo uitgevoerd. Eerst een briefje met een link. De mensen kunnen de schriftelijke vragenlijst zo opvragen. Dat gebeurt weinig overigens. Na verloop van tijd gaan we bellen. En het CBS gaat dan uiteindelijk nog face-to-face langs bij mensen. Dat is gewoon een toenemend kostenverhaal.

Wat voor de uitvoering van SIM2010 het precieze kostenvoordeel zal zijn, is op dit moment niet te bepalen. Daarvoor zijn gespecificeerde offertes van veldwerkbureaus nodig. Enige kanttekeningen bij al te hoge verwachtingen rondom besparingsmogelijkheden zijn echter wel te maken. Mixed-mode onderzoek leidt namelijk ook tot additionele uitgaven en op verschillende onderdelen tot langere doorlooptijd. De organisatie van het veldwerk is aanzienlijk complexer en afhankelijk van de toegepaste modes is het denkbaar dat er meer invoerwerkzaamheden nodig zijn dan bij CATI. Voorts is in het voorgaande toegelicht dat additionele onderzoekskosten voortkomen uit het opstellen van een complexere vragenlijst (voor verschillende modes) en uit de analyse van mixed-mode effecten:

Dan denk je dat een aantal dingen snel en goedkoop binnen is: je stuurt het schriftelijk of via het web en dan ben je klaar. Vaak wordt dan vergeten dat je daarna een niet zo simpel bestand hebt, want daar zitten allemaal extra factoren bij. Hoe is iemand geïnterviewd? Iemand is schriftelijk geïnterviewd. Iemand is mondeling geïnterviewd nadat hij schriftelijk niet meegedaan heeft. Als je het idee hebt dat er een bepaalde vermenging is van persoonskenmerken en in welke groep en in welke mode iemand terecht komt, dan moet je zorgen dat dit goed wordt vastgelegd. En dan moet je dat eigenlijk ook in je analyse meenemen.

Bij een mixed-mode is het voordeel dat je kosten misschien minder kunnen zijn, hoewel ik net zei dat je de kosten die je in het begin bespaart misschien later weer terugkrijgt omdat je eerst de data goed moet analyseren.

Behalve de verhoging van de kosten hebben deze aanvullende activiteiten vanzelfsprekend gevolgen voor de totale doorlooptijd.

In hoofdstuk 2 is aangegeven dat behalve naar de hoogte van de kosten gekeken zou moeten worden naar kosteneffectiviteit: hoe krijg je met zo weinig mogelijke kosten het beste bestand. Die vraag is zeer relevant, maar zonder inzicht in de kosten van uitvoering van surveys met verschillende designs is het op dit moment ondoenlijk hierover een zinnige uitspraak te doen. We kunnen op dit moment eigenlijk alleen wijzen op de al eerder genoemde risico's van het uitvoeren van mixed-mode onderzoek onder niet-westerse migranten, zonder dat we dit goed kunnen afwegen tegen de kosten van de verschillende designs. Toepassing van CATI onder niet-westerse migranten levert problemen op met betrekking tot de dekking, de respons is laag en de vragenlijst zal aanzienlijk

moeten worden ingekort en aangepast. Bij Computer Assisted Web Interviewing (CAWI) is de situatie niet veel beter. Zelfs het CBS krijgt met veel moeite (veel reminders e.d.) de respons onder een aselechte Nederlandse bevolkingssteekproef niet boven de 30%. De respons zal zeer waarschijnlijk veel lager uitpakken wanneer de mode alleen wordt aangeboden aan zeer moeilijk bereikbare en weigerachtige respondenten zoals jongeren. Dit blijkt ook uit ervaringen van andere onderzoekers met websurveys onder tweedegeneratie jongeren van Turkse en Marokkaanse herkomst. Consequenties van die lage respons is dat een zeer grote steekproef moet worden getrokken onder niet-westerse migranten en hun kinderen, wat zorgt voor hogere kosten vooraf (aankondigingbrief) en tijdens het veldwerk (reminder). In het geval van een zeer lage respons zijn waarschijnlijk meerdere reminders nodig. Dit zal ook weer consequenties hebben voor de doorlooptijd van de survey. Een lage respons heeft verder als risico de mogelijke uitputting van de steekproef voor bepaalde subgroepen of in sommige gemeenten.

Gezien de lengte van de SIM-vragenlijst, de doelpopulatie en de verwachte zeer lage respons van webenquêtes (en de daarbij behorende selectiviteit) zal de kosteneffectiviteit van webenquêtes naar alle waarschijnlijkheid aanzienlijk lager uitpakken dan het kostenvoordeel dat vaak wordt verwacht. Hoe het precies met de kosteneffectiviteit staat, valt – het zij herhaald – op dit moment niet met zekerheid te zeggen. Offertes van veldwerkbureaus moeten hierover meer duidelijkheid geven.

3.3 Samenvatting en conclusies

In dit hoofdstuk is op basis van de literatuur en interviews nagegaan welke voor- en nadelen verbonden kunnen worden aan mixed-mode onderzoek. Tabel 3.1 vat de bevindingen samen.

Tabel 3.1

Samenvatting voor- en nadelen mixed-mode onderzoek

genoemde voordelen	genoemde nadelen
<i>betere responsamenstelling</i> Door het gebruik van meerdere modes zou men tot een betere responsamenstelling komen.	Empirisch onderzoek vindt geen positief effect van meerdere modes op de responsamenstelling.
<i>hogere respons</i> Door het gebruik van meerdere modes zou men tot een hogere respons komen.	Het positieve effect van meerdere modes op de respons wordt empirisch niet ondersteund. Inzet van websurveys bij tweede-generatie Turkse en Marokkaanse Nederlanders levert nauwelijks respons op. Ervaringen met onderzoek onder migranten met schriftelijke en drop-off lijsten zijn slecht: lage en selectieve respons. Respons in mixed-mode valt alleen hoger uit wanneer f-t-f onderdeel is van de mixed-mode benadering.
<i>lagere kosten</i> vanwege inzet van minder enquêteurs	Vragen bij effect op totale onderzoekskosten. In hoeverre wegen lagere interviewkosten op tegen extra kosten grotere steekproef-omvang, extra invoer- en onderzoekscapaciteit (in analyse en veldwerkondersteuning)? Mixed-mode effect. Verschillende modes leiden tot verschillende uitkomsten.
<i>kortere doorlooptijd</i>	Doorlooptijd wordt verlengd. Door o.a. complexere vragenlijst-constructie, invoer bij schriftelijke mode, ingewikkeld logistiek veldwerk en analyse mixed-mode effect. Vergelijkbaarheid van reeksen en binnen reeksen wordt door mixed-mode onderzoek aangetast.

De inventarisatie van voor- en nadelen laat volgens ons zien dat aan de toepassing van een mixed-mode design voor de SIM2010 belangrijke nadelen zijn verbonden. Er bestaat gerede vrees voor een lagere respons, voor het ontstaan van meetfouten, een totaal andere vragenlijst en – met de voorgaande punten verbonden – een breuk met de voorgaande SIM. Dit zijn belangrijke criteria waaraan de kwaliteit van een bestand mag worden afgemeten. De winst van een eventueel kortere doorlooptijd gaat mogelijkerwijs verloren door meer werk vooraf – opstellen van nieuwe en waarschijnlijk meerdere vragenlijsten – en door meer werk achteraf – analyse van de mixed-mode effecten. Een grote onbekende zijn de kosten: wat levert een aangepast design aan besparingen op. Verder is de inschatting van de precieze gevolgen van een designverandering moeilijk te maken. Dat maakt een keuze op dit moment niet gemakkelijk, al is de richting naar ons idee wel duidelijk.

Ook de door ons geïnterviewde respondenten zijn niet eensluidend in hun oordeel. Er is een categorie die zich op het standpunt stelt dat er eigenlijk geen houden meer aan is: mixed-mode moet, of je nu wilt of niet en bij die ontwikkeling kun je maar het best aansluiten.

Ik denk dat je er gewoon niet onderuit gaat komen. Het gaat gewoon gebeuren en dan kun je maar beter vóór lopen.

Ik denk dat in de toekomst het face-to-face onderzoek er helemaal uitgaat. Het is onbetaalbaar en de respons wordt zo laag, dat het niet meer rechtvaardigt ten opzichte van andere methoden. Ik zou het zeker doen. Ik zou het juist nu doen. Ik zou het half om half doen. Half via het originele design en half via het nieuwe design.

Vernieuwing betekent dat je je ook internationaal blijvend kunt onderscheiden.

Maar je zou een voorloper kunnen worden in de vernieuwing op dit gebied. Want Nederland doet het al heel bijzonder goed en hiermee zou je echt iets heel moois kunnen hebben.

Anderen adviseren echter zeer terughoudend te zijn met de toepassing van mixed-mode bij onderzoek onder niet-westerse migranten.

Ik zou toch heel voorzichtig zijn met de mixed-mode. En ik zou alleen de CAPI/CASI-variant echt uitproberen. Dus ik zou voor deze groepen toch een persoonlijk interview doen. Er is niet zo veel alternatief. En die CAWI op het web moet je gewoon goed uittesten. Want wij hebben gezien dat mensen dat niet zo lang volhouden, die gaan niet een uur lang een vragenlijst invullen. [...] Misschien dat de conclusie is dat je niet te veel van de mixed-mode moet verwachten hoor. Bij ons zijn we daar nou niet echt heel enthousiast over, als we het over mochten doen zou ik het anders doen. Als je eenmaal binnen bent, moet je zorgen dat je het dan doet. En daar moet je op inzetten in je ontwerp. [...] Je moet heel veel tijd reserveren voor de benadering en daar moet je heel creatief in zijn. Dat je interviewers samen laat optrekken, want die interviewers gemotiveerd houden is een hele belangrijke opdracht, daar moet je veel aan doen. En die moet je elkaar laten helpen, die moet je ondersteunen, opdat ze binnenkomen. En als ze binnenkomen, komt het wel. [...] Om die interviewers gemotiveerd te houden, moet je ze heel goed trainen en heel goed motiveren. Dat is duur hè, dat is duur. Die alternatieven hebben nog niet zo veel overtuigend succes opgeleverd bij deze groep, bij ons dan. Maar nogmaals, we hebben echt grote steden en die leeftijdsgroepen. En dat is dus gewoon het moeilijkste van het moeilijkste. Het is de moeilijkste leeftijdsgroep en het is de moeilijkste locatie.

Om nou blindelings niet alleen het design (vergeleken met vroeger) te veranderen, maar ook een nieuw design te kiezen dat naar mijn weten bij etnische minderheden nog nooit is toegepast? Het gaat om heel erg veel geld. Dat is alsof je een nieuw schip gaat ontwikkelen wat sneller, en mooier, en goedkoper naar Amerika gaat varen en het schip in het water zet en wel ziet of het onderweg zinkt of niet. Dat kost een soort van risico's, en wat dat oplevert moet je ook in de kosten meenemen.

De keuze wordt verder bemoeilijkt doordat lang niet alles duidelijk is van het effect op de meting bij het gebruik van meerdere modes.

Het is gewoon heel ingewikkelde, complexe materie, er is nog geen eenduidig antwoord.

Met zoveel onzekerheid en onduidelijkheid zijn er eigenlijk maar twee mogelijkheden: of je doet het niet en gaat op de oude voet verder of je voert een pilot met experimenten uit. Dit laatste zou in elk geval het advies van bijna alle geïnterviewde experts zijn.

Ik zou eigenlijk adviseren om eerst maar eens hele kleine experimentjes te gaan doen om te kijken of wat ik zeg nu echt waar is of niet. Misschien valt het wel mee, of misschien is het nog veel erger. Iedereen roept maar wat. Het hoeft allemaal niet groot van omvang te zijn. Trek eens honderd allochtonen, of duizend en kijk eens met een klein vragenlijstje met wat typische vragen erin van het soort dat je in hoge mate zou willen stellen. In een onderzoekje van tien minuten of zo. Dan weet je nog steeds niet hoe het gaat met lange vragenlijsten maar daar kun je wel iets meer achter komen, achter mode effecten.

Dat zou ik doen, als ik niet hoeft na te denken over budgetten. Wanneer ik het mag zeggen, dan zou ik ervoor kiezen om de steekproefiets te verhogen om te zorgen dat je een goede vergelijking binnen dat jaar kan maken en een mixed design maken met de meest toekomstbestendige methode.

Het is een behoorlijke overgang. Ik zou het doen, maar met een pilot. Misschien moet je zelfs een aantal kleinere pilots doen, die op elkaar voortbouwen voor de beslissingsfase. Ik denk dat je wel de hele tijd kunt roepen dat je het niet moet doen. Dat is aan de ene kant het gemakkelijkst, maar hoelang kan men dat zich nog veroorloven? Dat is de rock bottom. Waarom begin je niet gewoon met een aantal pilots om te zien hoe het gaat? Daarna kun je beslissen welke kant je op gaat. Je kunt beter een aantal kleine pilots doen, omdat je dan weet waar je over praat.

Je hebt heel veel onbekende factoren die overal op ingrijpen. Als je iets gaat veranderen, ga dan eerst een paar dingen uittesten. Want als je dingen niet uittest en kiest voor dit dan kan het zijn dat het allemaal heel goed en heel mooi loopt, maar aan het eind zit je toch met die combinatie van allerlei effecten. Het kan ook zijn dat je resultaten vergeleken met de vorige keer ineens heel erg veranderd zijn; dat ze dan beter zijn. Je weet in ieder geval niet iets over de vergelijkbaarheid. Je doet een stap in het duister.

In het volgende hoofdstuk maken we de balans van dit alles op.

4 Conclusies en varianten

4.1 Variant 1: behoud van het huidige design

Uit het voorgaande mag worden afgeleid dat de introductie van een mixed-mode design moeilijkheden met zich meebrengt. Naast de introductie van een mixed-mode effect brengt vooral het (gedeeltelijk) afstappen van een mondelinge mode problemen met zich mee bij het verzamelen van data onder niet-westerse migranten. Enquêteonderzoek onder deze groep is al ingewikkeld genoeg. Zo is het niet eenvoudig respons te verkrijgen. Ook de meting onder niet-westerse migranten levert meer problemen op dan onder autochtone Nederlanders. De mondelinge afnamemethode is zowel voor tot het verkrijgen van een goede en evenwichtige respons als voor een zo goed mogelijke meting de mode die onder deze groep de beste kwaliteit garandeert. De aanwezigheid van een interviewer helpt de non-respons onder niet-westerse migranten te minimaliseren en kan een respondent helpen in het responsproces bij (lichte) taalproblemen.

Juist onder niet-westerse migranten zijn er sterke aanwijzingen dat mixed-mode onderzoek niet goed werkt. Het (gedeeltelijk) afstappen van een mondelinge mode voor de al moeilijke groep in enquêteonderzoek zal de respons nog verder verlagen. Ook zal de afwezigheid van een interviewer tot extra problemen leiden bij degenen die de taal niet goed spreken of lager opgeleid zijn.

Ook is de complexe en lange vragenlijst van SIM lastig naar een andere dan mondelinge mode om te zetten. Bovendien zal het veldwerk bij het invoeren van een mixed-mode design een stuk complexer worden en is het onduidelijk wat het effect zal zijn op de lengte van onderzoekstijd. Door een extra slag die in de analysefase zal moeten worden gemaakt, kan deze langer duren. Wanneer het mondeling verzamelen van data dus (gedeeltelijk) verlaten wordt, zal dit een waarschijnlijk een negatief effect hebben op de kwaliteit van de enquêtedata, namelijk een verslechtering van de respons en meting. Bovendien wordt bij een mixed-mode design een mixed-mode effect geïntroduceerd. Er zijn dus goede redenen om het design van de SIM niet te veranderen (variant 1).

Genoemde kwaliteitsargumenten wegen zwaar. Niettemin bestaat er ook oog voor de argumenten die hier tegenin kunnen worden gebracht, namelijk de kosten en een welhaast onvermijdelijk lijkende ontwikkeling in de richting van mixed-mode designs. Naast het behoud van het huidige design gaan we dan ook in op andere varianten en onze opvattingen daarover. Er kan gekozen worden voor een mixed-mode design waar het behoud van kwaliteit vooropstaat. Hierbij start het veldwerk met de beste mode, gevolgd door een andere mode. Deze variant wordt besproken in paragraaf 4.2. De mixed-mode variant waarin het primaat ligt bij de verlaging van kosten begint met de goedkoopste mode (§ 4.3). In beide mixed-mode varianten is sprake van een sequentieel design. In zo'n design kan echter niet geanalyseerd worden wat het effect is van de verandering in het design op de trend en of respondenten in verschillende modes daadwerkelijk anders antwoorden dan op enquêtevragen. Dit kan alleen geanalyseerd worden

in een design waarin respondenten random worden toegewezen aan een mode. Dit biedt de beste mogelijkheden om populatie- en meeteffecten uit elkaar te halen. Een random toewijzing van respondenten aan modes kan alleen in een concurrent design (§ 4.4). Voor alle mixed-mode varianten geldt dat de huidige vragenlijst van de SIM zal moeten worden aangepast. Zoals eerder al is aangegeven kan de vragenlijst door de lengte en complexiteit ervan niet gebruikt worden in een andere dan een mondelinge mode.

4.2 Variant 2: mixed-mode onderzoek met kwaliteit als criterium

Als de verbetering van kwaliteit het belangrijkste uitgangspunt is voor de keuze voor een mixed-mode design, ligt de start van het veldwerk bij de mode die de hoogste data-kwaliteit oplevert. In dit design wordt begonnen met face-to-face onderzoek gevolgd door telefonisch en dan webonderzoek. In zo'n mixed-mode design zijn er waarborgen ingebouwd dat de kwaliteit en daarmee de vergelijkbaarheid door de jaren heen niet dramatisch verandert; waarschijnlijk is dat het grootste gedeelte mondeling zal responderen. Dit is uiteraard wel afhankelijk van hoe lang de mondelinge mode in het veld blijft staan. Zoals in vorig hoofdstuk toegelicht, is het de vraag of via dit design de responsamenstelling wordt verbeterd. De aanduiding mixed-mode onderzoek met 'kwaliteit als criterium' moet dus vooral worden gelezen als de keuze voor de volgorde van het inzetten van modes in het veldwerk. De logistiek van een sequentieel mixed-mode design (zowel bij de keuze voor kwaliteit als voor kosten als criterium) is een stuk complexer dan bij een uni-mode design. De overstap van de ene mode naar een andere mode vraagt een gedegen veldwerkverantwoording. Van een mondelinge mode (gedeeltelijk) overstappen naar een andere mode zal kostenvoordelen met zich meebrengen, maar extra kosten zullen ontstaan in het veldwerkmanagement en de analyse. Groot nadeel bij een sequentieel design (zowel bij kwaliteit als kosten als gekozen criterium) is dat door het gebruik van meerdere modes populatie- en meeteffecten ontstaan. Dit betekent ook dat verschillen in uitkomsten ten opzichte van de eerdere SIM lastig te interpreteren zijn.

4.3 Variant 3: mixed-mode onderzoek met kosten als criterium

Er kan ook gekozen worden voor een sequentiële benadering waarin kosten als criterium gebruikt worden. In dit geval wordt met de goedkoopste mode begonnen: het web. Daarna volgt telefonische afname en pas aan het eind wordt de mondelinge mode ingezet. Dit design kan de interviewkosten reduceren. Zoals eerder toegelicht zijn wel additionele kosten te verwachten. Belangrijk is dat in dit design de kwaliteit van de enquêtedata ernstig onder druk komt te staan. Ten eerste is het lastig een goede en representatieve websurvey uit te voeren (o.a. vanwege problemen rondom dekking en non-respons). Deze levert vaak een lage en selectieve respons op, en dan met name onder migranten. Dit moet dan in een latere fase worden rechtgetrokken in de mondelinge mode. Het is maar de vraag of dat dan nog lukt. Juist het contact leggen met niet-westerse migranten is lastig en vraagt veel contactpogingen. Wanneer slechts in een deel van het veldwerk de mondelinge mode wordt ingezet, zouden er problemen kunnen ontstaan omdat er te weinig tijd is om een adequate respons te realiseren. Ten slotte geldt ook voor deze

benadering dat de vergelijkbaarheid door de jaren heen met andere SIM-studies heel lastig wordt. Deze variant is het minst aan te bevelen.

4.4 Variant 4: experiment met mixed-mode

Een verandering in het design doorkruist de wens om met de lange reeks van onderzoeken onder niet-westerse minderheden (sinds de jaren tachtig van de vorige eeuw) een trend neer te zetten. De vergelijkbaarheid van de verschillende metingen is lastig met de introductie van nieuwe modes. Het is dan namelijk onduidelijk welke verschillen in schattingen zijn toe te schrijven aan veranderingen in het design en welke aan werkelijke veranderingen. Deze effecten kunnen ontrafeld worden wanneer gekozen wordt voor een studie waarin ook een experiment met de designveranderingen is ingebouwd. Het oude design blijft dan gehandhaafd en naast het reguliere onderzoek loopt een experiment of een aantal experimenten waarin het mixed-mode design wordt uitgetest. In zo'n experiment kan dan gekeken worden wat de effecten van een mixed-mode onderzoek zijn op de responsamenstelling en of er een mixed-mode effect optreedt. Om deze effecten goed in kaart te brengen is een toevallige toewijzing van respondenten aan een mode nodig. Dit kan alleen in een concurrent design. Voordeel is dat het oorspronkelijke design in ieder geval bruikbare gegevens oplevert. Zonder problemen kan de trendvergelijking worden voortgezet en kan goed in kaart worden gebracht wat de effecten van een mixed-mode onderzoek op de kwaliteit en kosten zullen zijn. Een nadeel is dat een experiment met een random toewijzing van modes aan respondenten niet efficiënt is. In dit geval zullen door de toevallige toewijzing aan modes ook mensen in een mode vallen van wie je van tevoren al weet dat de kans van deelname van die persoon aan die mode niet hoog is. Dat geldt bijvoorbeeld voor de toevallige toewijzing van een webmode aan een oudere migrant. Een ander nadeel is dat het uitvoeren van een experiment geld en tijd kost.

4.5 Conclusies

Een mixed-mode design implementeren in de SIM is door de bijzondere doelgroep en doelstelling van deze studie geen stap die lichtvaardig gezet kan worden. Niet-westerse migranten en hun kinderen laten zich lastig in een andere dan de mondelinge mode enquêteren. Een mixed-mode design zal dus waarschijnlijk problemen opleveren met de respons en responsamenstelling bij deze groep. Ook zullen er meer problemen in de meting optreden wanneer er geen interviewer aanwezig is. Bij een verandering van het design doet zich ook het probleem voor dat het lastig is de nieuwe cijfers te vergelijken met die van de voorgaande SIM. Of er een gunstig effect op de kosten zal zijn bij de invoering van een mixed-mode design is onduidelijk. Er bestaan toch al veel onduidelijkheden over de effecten van mixed-mode designs op kwaliteit en kosten (behalve dat er dus sterke aanwijzingen dat het voor de doelgroep van de SIM lastig is een mixed-mode design te gebruiken). Om meer duidelijkheid in deze effecten te krijgen is het aan te bevelen mixed-mode varianten te testen. Er kan voor gekozen worden om dit experimenteel te doen, zoals besproken in paragraaf 4.4. Het is dan een concurrent design

nodig waarin (relatief kleine groepen) steekproefeenheden random aan een mode kunnen worden toegewezen. De mondelinge mode kan dan als hoofdmode blijven bestaan en waarborgen dat er in ieder geval een groot deel van de data gebruikt kan worden. Een andere relatief ‘veilige’ optie is om voor een sequentieel design te kiezen waarin een mondelinge mode het eerst wordt ingezet in het veld, zoals besproken in paragraaf 4.2. Ook in dit geval is er een soort waarborg dat een gedeelte van de data in ieder geval goed gebruikt kan worden. Wil men de interviewkosten reduceren, dan kan voor een mixed-mode design gekozen worden waarin het veldwerk begint met een webmode (zie § 4.3). De kwaliteitsrisico’s zijn dan echter een stuk groter. Naar onze mening brengt deze variant te veel risico’s met zich mee. Naast de zorgen die er zijn over de respons en meting, – met name bij de SIM-doelgroep – bij het verlaten van een exclusief mondelinge mode, geldt voor alle mixed-mode varianten dat het veldwerk een stuk complexer wordt en dat de vragenlijst (die nu alleen geschikt is voor een mondelinge mode) ingekort dient te worden.

Noten

- 1 Met de geïnterviewden is afgesproken dat hun antwoorden niet op naam in dit rapport worden verwerkt.
- 2 Er kan ook sprake zijn van overdekking, bijvoorbeeld als in een adressensteekproef ook vakantiehuizen worden meegenomen en mensen via twee adressen in de steekproef kunnen vallen. Bij een schriftelijke enquête kan een nieuwe bewoner de vragenlijst voor de oude bewoner invullen, maar misschien ook wel de lijst die naar het eigen oude adres was gestuurd, en bij een internetenquête kan de persoon met de meeste internetvaardigheden de lijst invullen in plaats van de geadresseerde persoon. Een andere vorm van overdekking kan voorkomen bij telefonische enquêtes (Bethlehem 2009) bij huishoudens met meerdere vaste aansluitingen, waardoor zo'n huishouden meer dan eenmaal kan meedoen.
- 3 Gehouden in 2006.
- 4 Survey Integratie Nieuwe Groepen (SING) is een survey gehouden in 2009 onder zes grote groepen migranten en hun kinderen (van Afghaanse, Irakese, Iranese, Somalische, Poolse en Chinese herkomst) en autochtone Nederlanders.
- 5 Onder een 'zachte' weigeraar wordt onder anderen iemand verstaan die bij de weigering als reden opgaf 'geen tijd' te hebben of geen gelegenheid.
- 6 Er kan ook gevarieerd worden met modes in de verschillende fasen van het surveyproces: Steekproef-eenheden kunnen bijvoorbeeld telefonisch gecontacteerd en mondeling geïnterviewd worden.
- 7 Bij (I) dezelfde respondent tijdens één interview; (I1) bij dezelfde respondent over meerdere meetmomenten (een panelonderzoek); (I11) bij groepen van respondenten met één meetmoment; of (IV) bij groepen respondenten over meerdere meetmomenten (zie ook De Leeuw 2008).
- 8 Het LAS 2004 is een survey gehouden in 2004 onder de vier grote groepen niet-westerse migranten en hun kinderen (van Turkse, Marokkaanse, Surinaamse en Antilliaanse herkomst) en autochtone Nederlanders woonachtig in de 50 grootste gemeenten van Nederland.
- 9 De Veiligheidsmonitor is een grootschalig onderzoek naar de gevoelens van (on)veiligheid onder de bevolking.
- 10 Enquête Beroepsbevolking (EBB) is een periodieke survey uitgevoerd door het CBS onder de Nederlandse beroepsbevolking.
- 11 SPVA's zijn een reeks surveys gehouden onder de vier grote groepen niet-westerse migranten en hun kinderen (van Turkse, Marokkaanse, Surinaamse en Antilliaanse herkomst) en autochtone Nederlanders.

Bijlage A Interviewguide

LOGBOEK

- **Naam interviewer:**
- **Naam respondent:**
- **Datum interview:**
- **Begintijd interview:**
- **Eindtijd interview:**
- **Duur interview:**
- **Locatie interview:**

Introductie

- 1 Introduceer jezelf.
- 2 **DOEN:** controle opnameapparatuur.
- 3 Introduceer project.

Zoals in de e-mail aangegeven voert het Sociaal en Cultureel Planbureau de Survey Integratie Minderheden (SIM) uit. In deze studie worden data verzameld onder allochtonen. Traditioneel gebeurt dit door middel van een mondelinge dataverzamelingmode. Bij deze studie wordt een random steekproef uit het GBA getrokken. Als van steekproefunits een telefoonnummer bekend is, kan dat gebruikt worden.

Vanuit de opdrachtgever van deze studie bestaat nu het verzoek te onderzoeken of er en zo ja, welke mogelijkheden er zijn om de data in de toekomst te verzamelen met behulp van een mixed-mode design. Voordat er een eventuele verandering in het design plaatsvindt, wil het SCP eerst de meningen van experts op het gebied van mixed-mode en/of allochtonenonderzoek over de mogelijke voor- en nadelen van zo'n verandering in het design inventariseren.

De Survey Integratie Minderheden wil een representatieve beschrijving van een breed veld van integratie-indicatoren geven. Deze worden gebruikt om verklarende analyses te maken en zijn ook een bron voor trends: de nieuwe dataverzameling zal ook als vergelijking met SIM2006 gebruikt worden. Verder is het van belang op te merken dat een deel van de Turkse en Marokkaanse respondenten geen of onvoldoende Nederlands spreekt en dat met name onder ouderen nogal wat analfabeten te vinden zijn.

- 1 Waarom deze respondent geselecteerd?
- 2 Introductie respondent.
- 3 Leg procedure interview uit:
 - a tijdsduur (60 min., natuurlijk ook mogelijkheid tot pauze);
 - b transcriptie;
 - c vertrouwelijkheid;
 - d eventuele vragen van respondent.

Inhoud

1 Mixed-mode onderzoek onder allochtonen

Er wordt vaak een onderscheid gemaakt tussen een mixed-mode design in de benaderingsstrategie en een mixed-mode design in de dataverzameling. In een mixed-mode design in de dataverzameling worden met meerdere dataverzamelingmodes data verzameld onder respondenten. In eerste instantie wil ik het met u hebben over een mixed-mode onderzoek waarin verschillende modes in de dataverzameling gebruikt worden. De vragen hebben hier uitdrukkelijk betrekking op de Survey Integratie Minderheden. Dus een studie met allochtonen als onderzoekspopulatie en een survey die een representatief beeld probeert te geven.

Schriftelijk

- a Welke mogelijkheden ziet u bij de enquêtering onder allochtonen als er naast een mondelinge dataverzameling een schriftelijke mode geïmplementeerd zou worden?
 - i Responscijfer:
 - 1 (contact: mode mondeling, schriftelijk, telefonisch, contactpogingen etc.)
 - 2 (participatie: incentives)
 - 3 (in staat: vertaling vragenlijst, tweetalige interviewers)
 - ii Responssamenstelling
 - iii Kwaliteit antwoorden / meting
 - iv Vergelijkbaarheid SIM2006
- b Welke mogelijkheden ziet u bij een websurvey? Een websurvey zal overigens betekenen dat steekproefeenheden schriftelijk / telefonisch verzocht wordt naar een website te gaan om een vragenlijst in te vullen. Dat betekent al dat de eerste benadering nooit per e-mail/web kan plaatsvinden.

En welke moeilijkheden ziet u bij het switchen van een mondelinge mode naar een mixed-mode met een mondelinge en schriftelijke dataverzameling:

- i Effect op responscijfer
- ii Effect op responssamenstelling
- iii Effect op kwaliteit van antwoorden
- iv Vergelijkbaarheid reeksen
 - 1 Respons: van alleen face-to-face naar mixed-mode mondelinge en schriftelijke mode
 - 2 Benadering door interviewer
 - 3 Hulp van interviewer bij deelname
 - 4 Meting van alleen face-to-face naar mixed-mode
 - 5 Hulp interviewer bij afname interview

Welke moeilijkheden ziet u bij een websurvey?

- c We proberen ook meer zicht te krijgen op de doorlooptijd van een enquête, waarin niet alleen gebruikgemaakt wordt van een mondelinge dataverzameling, maar waarin naast deze mondelinge enquêtes ook schriftelijke enquêtes afgenomen worden. Daarbij wordt in het veldwerk gestreefd naar een bepaalde steekproefomvang.

Hoe zou u de doorlooptijd inschatten bij een mixed-mode onderzoek bestaande uit een mondelinge mode en schriftelijke enquêtes? En hoe zit dat voor een mix van een mondelinge en web dataverzameling?

- d Ook proberen we een inschatting te krijgen van de te verwachten totale survey kosten wanneer in een mixed-mode design met mondelinge en schriftelijke interviews wordt gewerkt. Hoe schat u deze in?

En in vergelijking met een survey waarin alleen mondelinge enquêtes worden afgenomen?

Ook hier zijn we geïnteresseerd in uw inschatting van de te verwachten totale survey kosten van een mixed-mode onderzoek bestaande uit een mondelinge en webvariant. Hoe schat u deze in?

- e Er bestaan verschillende mogelijkheden om de dataverzamelingsmethoden vorm te geven.
- 1 Modes kunnen geoptimaliseerd worden / op maat gesneden (mode-specific design); dit betekent dat de voordelen van elke mode geoptimaliseerd worden.
 - 2 De modes kunnen gelijk gemaakt worden (een generalized mode), wat betekent dat daarnaar gestreefd wordt een gelijke stimulus in elke mode te bereiken.
 - 3 Een van de modes kan geoptimaliseerd (uni-mode of uniform-mode design). Hierin wordt een van de modes geoptimaliseerd en de andere modes worden dan aan deze mode aangepast. Wat zou u adviseren?
- f Het gebruikmaken van meerdere modes in een design kan met zich meebrengen dat data verzameld in de ene mode er anders uitzien dan de data verzameld in een andere mode, juist door gebruik te maken van die andere mode, het mixed-mode effect. Hoe schat u de eventuele aanwezigheid van een mixed-mode effect en de gevolgen ervan in?
- g Het gebruik van meerdere modes in de dataverzameling kan dus met zich meebrengen dat er een mixed-mode effect optreedt. Welke analysemogelijkheden zouden gebruikt kunnen worden om mixed-mode effecten te evalueren?
- h Tot nu toe hebben we het alleen gehad over mixed-mode onderzoek in de dataverzameling. Ik zou het nu willen hebben over een mixed-mode onderzoek in de benadering. Een voorbeeld van een mixed-mode design in de benaderingsstrategie is bijvoorbeeld een mondelinge dataverzameling laten volgen door een telefonische herbenadering. Welke mogelijkheden en moeilijkheden ziet u in het gebruik van mixed-mode onderzoek bij de benadering van allochtonen?
- 2 Advisering design SIM
- a Wat zouden argumenten zijn om *wel* een mixed-mode design in de dataverzameling te implementeren en welke mix zou u dan adviseren?

En hoe zit dat voor een mixed-mode onderzoek met een mix van modes in de benaderingsstrategie?

- b En welke argumenten zijn er om *geen* mixed-mode in de dataverzameling te implementeren?

En opnieuw voor de benaderingsstrategie?

- c Dit alles gezegd hebbende, wat zou u adviseren gegeven dit onderzoek met een al-luchtone onderzoekspopulatie?
 - i *mixed-mode doen*
 - ii *niet doen*
 - iii *pilot studie*

3 Andere designoverwegingen

Tot slot zijn er nog twee andere mogelijke veranderingen in het design die ik kort met u zou willen bespreken:

- a Er bestaat ook de mogelijkheid om in het nieuwe design gebruik te maken van personen die hebben aangegeven dat zij herbenaderd willen worden voor vervolgonderzoek. Hoe ziet u de voor- en nadelen van zo'n methode in vergelijking met het trekken van een verse steekproef?
- b Er wordt overwogen om gebruik te maken van zogenoemde wisselende blokken in het nieuwe design van de Survey Integratie Minderheden. Dit wil zeggen dat om de interviewtijd voor de respondent te beperken slechts een gedeelte van de vragenlijst aan de respondent wordt voorgelegd. Aan een andere respondent worden in dat geval de andere vragen van de vragenlijst gesteld. Later kunnen de data dan samengevoegd worden. Welke voor- en nadelen ziet u hier?

Afsluiting

Respondent vragen:

- 1 of er nog vragen zijn bij respondent;
- 2 of we de respondent nog een keer mogen benaderen als we bepaalde zaken achteraf onduidelijk zijn.

Dank respondent.

Literatuur

- Aitkin, J.F., P.H. Youl, M. Janda, E. Elwood, I.T. Ring en J.B. Lowe (2004). Comparability of skin screening histories obtained by telephone interviews and mailed questionnaire: A randomized crossover study. In: *American Journal of Epidemiology*, jg. 160, nr. 6, p. 598-604.
- Ait Moha, A. (2009). *Online onderzoek onder Nieuwe Nederlanders – Illusie of haalbaar?* Presentatie op de lezingenmiddag over enquêteonderzoek onder migranten en etnische minderheden, maart 2009, Den Haag.
- Aquilino, W.S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. In: *Public Opinion Quarterly*, vol. 58, nr. 2, p. 210-240.
- Ariel, A., D. Giessen, F. Kerssemakers en R. Vis-Visschers (2008). *Literature review on mixed-mode studies*. Heerlen/Voorburg: Centraal Bureau voor de Statistiek (CBS intern rapport, BPA-nummer DMH-2008-04-16-RVCS).
- Banning, R., B. Schouten en W. Wetzels (2009). *Effectiviteit van mixed-mode onderzoek in het verminderen van onderdekking en non-respons*. Presentatie op discussiemiddag over mixed-mode surveys, 7 december 2009, Den Haag.
- Bethlehem, J. (2004). *Technieken voor survey-onderzoek*. Voorburg/Heerlen: Centraal Bureau voor de Statistiek.
- Bethlehem, J. (2009). *Applied survey methods: A statistical perspective*. Hoboken, New Jersey: John Wiley & Sons (Wiley series in survey methodology).
- Beukenhorst, D. en W. Wetzels (2009). *A comparison of two mixed-mode designs of the Dutch safety monitor: mode effects, costs, logistics*. Presentatie op de ESRA-conferentie, juni 2009, in Warschau, Polen.
- Biemer, P.B. en L.E. Lyberg (2003). *Introduction to survey quality*. Hoboken, New Jersey: John Wiley & Sons.
- Bronner, A.E. en T. Kuijlen (2006). De levende of digitale interviewer: Een vergelijking tussen CASI, CAPI en CATI. In: A.E. Bronner, P. Dekker, E. de Leeuw, L.J. Paas, K. de Ruyter, A. Smidts en J.E. Wieringa (red.), *Ontwikkelingen in het Marktonderzoek, Jaarboek 2006* (p. 77-94). Haarlem: SpaarenHout.
- Cobben, F. (2009). *Nonresponse in sample surveys, methods for analysis and adjustment*. Heerlen/Voorburg: Centraal Bureau voor de Statistiek.
- Dagevos, J., M. Gijsberts, J. Kappelhof en M. Vervoort (2007). *Survey Integratie Minderheden 2006*. Den Haag: Sociaal en Cultureel Planbureau (SCP-special 8).
- Dillman, D.A. (2000). *Mail and internet surveys: The total design method*. New York: John Wiley & Sons.
- Dillman, D.A. (2007). *Mail and internet surveys: The tailored design method (2007 update)*. New York: John Wiley & Sons.
- Dillman, D.A. en J. Tarnai (1991). Mode effects of cognitive designed recall questions: A comparison of answers to telephone and mail surveys. In: P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, en S.J. Sudman (red.), *Measurement errors in surveys*. Hoboken, New Jersey: John Wiley & Sons.
- Eurostat (2005). *European Statistics Code of Practice*. Luxemburg: Eurostat. Feskens, R.C.W., J.J. Hox, G.J.L.M. Lensvelt-Mulders en J.J.G. Schmeets (2006). Collecting data among ethnic minorities in an international perspective. In: *Field Methods*, jg. 18, nr. 3, p. 284-304.
- Feskens, R.C.W., J.J. Hox, G.J.L.M. Lensvelt-Mulders en J.J.G. Schmeets (2007). Nonresponse among ethnic minorities: A multivariate analysis. In: *Journal of Official Statistics*, jg. 23, nr. 3, p. 387-408.
- Gijsberts, M. en J. Dagevos (2009). *Jaarrapport Integratie 2009*. Den Haag: Sociaal en Cultureel Planbureau.
- Greenfield T.K., L.T. Midanik en J.D. Rogers (2000). Effects of telephone versus face-to-face interview modes on reports of alcohol consumption. In: *Addiction*, jg. 95, nr. 2, p. 277-284.

- Groves, R.M., E. Singer en A. Corning (2000). Leverage-saliency theory of survey participation: Description and an illustration. In: *Public Opinion Quarterly*, jg. 64, nr. 3, p. 299-308.
- Groves, R.M. (1989). *Survey errors and survey costs*. New York: Wiley.
- Groves, R.M. en R.L. Kahn (1979). *Surveys by telephone: A national comparison with personal interviews*. New York: Academic Press.
- Herzog, R.A. en J.G. Bachman (1980). Effects of questionnaire lengths on response quality. In: *Public Opinion Quarterly*, jg. 45, nr. 4, p. 489-504.
- Hox, J.J., E.D. de Leeuw en D.A. Dillman (2008). The cornerstones of survey research. In: E.D. de Leeuw, J.J. Hox en D.A. Dillman (red.), *International handbook of survey methodology* (p. 1-17). New York: Taylor en Francis, Psychology Press (EAM series).
- Jäckle, A., C. Roberts en P. Lynn (2006). *Telephone versus face-to-face interviewing: Mode effects on data quality and likely causes. Report on phase II of the ESS-Gallup Mixed Mode Methodology Project*. Colchester: University of Essex (ISER Working Paper 2006-41) (<http://www.iser.essex.ac.uk/pubs/workpaps/>).
- Ingen, E. van, J. de Haan en M. Duimel (2007). *Achterstand en afstand. Digitale vaardigheden van lager opgeleiden, ouderen, alloctonen en inactieven*. Den Haag: Sociaal en Cultureel Planbureau (SCP-publicatie 2007/24).
- Kampen, J.K. (2006). The impact of survey methodology and context on central tendency, nonresponse and associations of subjective indicators of government performance. In: *Quality and Quantity*, jg. 41, nr. 6, p. 793-813.
- Kappelhof, J. (2010). *Op maat gemaakt? Een evaluatie van enkele responsverbeterende maatregelen onder Nederlanders van niet-westerse afkomst*. Den Haag: Sociaal en Cultureel Planbureau (SCP-special 53).
- Krosnick, J.A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. In: *Applied Cognitive Psychology*, vol. 5, p. 213-36.
- Krosnick, J.A. (1999). Survey research. In: *Annual Review of Psychology*, jg. 50, p. 537-567.
- Krosnick, J.A. (2009). *The end of agree/disagree rating scales: Acquiescence bias and other flaws suggest a popular measurement method should be abandoned*. Keynote speech op ESRA-conferentie, juni 2009, in Warschau, Polen.
- Krosnick, J.A. en D.F. Alwin (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. In: *Public Opinion Quarterly*, vol. 51, p. 201-219.
- Krysan, M., H. Schuman, L.J. Scott en P. Beatty (1994). Response rates and response content in mail versus face-to-face surveys. In: *Public Opinion Quarterly*, vol. 58, p. 381-399.
- Laan, P. van der en W. van Nunspeet (2009). *Modernising household surveys in the Netherlands: Design, efficiency gains and perspectives*. Heerlen/Voorburg: Centraal Bureau voor de Statistiek (Discussion paper 09044).
- Leeuw, E.D. de (1992). *Data quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT-Publicaties.
- Leeuw, E.D. de (2005). To mix or not to mix? Data collection modes in surveys. In: *Journal of Official Statistics*, jg. 21, nr. 2, p. 233-255.
- Leeuw, E.D. de (2008). Choosing the method of data collection. In: E.D. de Leeuw, J.J. Hox, en D.A. Dillman (red.), *International handbook of survey methodology* (p.113-135). New York: Taylor en Francis, Psychology Press (EAM series).
- Leeuw, E.D. de (2009). *Mixing survey modes: Challenges for questionnaire construction*. Presentatie op discussiemiddag over mixed-mode surveys, 7 december 2009, Den Haag.
- Leeuw, E.D. de, D.A. Dillman en J.J. Hox (2008). Mixed-mode surveys: When and why? In: E.D. de Leeuw, J.J. Hox en D.A. Dillman (red.), *International handbook of survey methodology* (p. 299-316). New York: Taylor en Francis, Psychology Press (EAM series).

- Link, M.W. en A.H. Mokdad (2005). Alternative modes for health surveillance surveys: An experiment with web, mail, and telephone. In: *Epidemiology*, jg. 16, nr. 5, p. 701-704.
- Lugtig, P. en G.J.L.M. Lensvelt-Mulders (2007). *Analyzing self-selection bias in internet panel surveys*. Paper gepresenteerd op de tweede conferentie van de European Survey Research Association, juni 2007, in Praag, Tsjechië.
- Lyberg, L. en D. Kasprzyk (1991). Data collection methods and measurement error. In: P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz en S.J. Sudman (red.), *Measurement errors in surveys* (p. 237-257). New Jersey: Wiley & Sons.
- Nicolaas, G. en S. Tipping (2004). *HSE social capital questions: Comparison of interview-administered questions and self-administered*. Manuscript voorbereid voor Socio-Economic Inequalities Branch, Office for National Statistics. Niet gepubliceerd.
- Pierzchala, M. (2006). *Disparate modes and their effect on instrument design*. Proceedings of the 10th International Blaise Users Conference (p. 199-209) in Papendal, Arnhem.
- Roberts, C. (2007). *Mixing modes of data collection in surveys: A methodological review*. Southampton (VK): ESRC National Centre for Research Methods (NCRM Methods Review Papers, NCRM/008).
- Rosenbaum, P.R. en D.B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. In: *Biometrika*, jg. 70, nr. 1, p. 41-55.
- Schmeets, H. (2005a). De leefsituatie van allochtonen. In: H. Schmeets en R. van der Bie (red.), *Over het omgaan met allochtonen in enquêteonderzoek* (p. 169-176). Heerlen/Voorburg: Centraal Bureau voor de Statistiek.
- Schmeets, H. (2005b). Slecht bereikbare allochtonen, autochtone weigeraars. In: H. Schmeets en R. van der Bie (red.), *Over het omgaan met allochtonen in enquêteonderzoek* (p. 81-88). Heerlen/Voorburg: Centraal Bureau voor de Statistiek.
- Schonlau, M., K. Zapert, L.P. Simon, K.H. Sanstad, S.M. Marcus, J. Adams, M. Spranca, H. Kan, R. Turner en S.H. Berry (2004). A comparison between responses from a propensity-weighted web survey and an identical RDD survey. In: *Social Science Computer Review*, jg. 22, nr. 1, p. 128-138.
- Schothorst, Y. (2002). Onderzoek onder allochtonen: wat mag, wat moet en wat kan? In: H. Houtkoop-Steenstra en J. Veenman (red.), *Interviewen in de multiculturele samenleving. Problemen en oplossingen*. Assen: Koninklijke Van Gorcum.
- Schothorst, Y. (2005). *Veldwerkverslag Leefsituatie Allochtone Stedelingen*. Amsterdam: Veldkamp.
- Schouten, B. en F. Cobben (2007). *R-indexes for the comparison of different fieldwork strategies and data collection modes*. Heerlen/Voorburg: Centraal Bureau voor de Statistiek (Discussion paper 07002).
- Stoop, I.A.L. (2005). *The hunt for the last respondent* (proefschrift). Utrecht: Universiteit van Utrecht.
- Tourangeau, R., L. Lips en K. Rasinski (2000). *The psychology of survey response*. Cambridge (VK): Cambridge University Press.
- Veenman, J. (2002). Interviewen in multicultureel Nederland. In: H. Houtkoop-Steenstra en J. Veenman (red.), *Interviewen in de multiculturele samenleving. Problemen en oplossingen*. Assen: Koninklijke Van Gorcum.
- Zouwen, J. van der (2006). De interviewer, hulp of hindernis? In: A.E. Bronner, P. Dekker, E. de Leeuw, L.J. Paas, K. de Ruyter, A. Smidts, en J.E. Wieringa (red.), *Ontwikkelingen in het Marktonderzoek, Jaarboek 2006* (p. 63-76). Haarlem: SpaarenHout.
- Zouwen, J. van der, en E.D. de Leeuw (1990 en 1991). The relationship between mode of administration and quality of data in survey research. In: *Bulletin de Méthodologie Sociologique*, nr. 29, p. 3-14 (First Part) en nr. 31, p. 49-60 (Final Part).

Publicaties van het Sociaal en Cultureel Planbureau

Werkprogramma

Het Sociaal en Cultureel Planbureau stelt twee keer per jaar zijn Werkprogramma vast.

De tekst van het lopende programma is te vinden op de website van het scp: www.scp.nl.

scp-publicaties

Onderstaande lijst bevat een selectie van publicaties van het Sociaal en Cultureel Planbureau. Deze publicaties zijn verkrijgbaar bij de boekhandel, of via de website van het scp.

Een complete lijst is te vinden op deze website: www.scp.nl.

Sociale en Culturele Rapporten

Investeren in vermogen. Sociaal en Cultureel Rapport 2006. ISBN 90-377-0285-6

Betrekkelijke betrokkenheid. Studies in sociale cohesie. Sociaal en Cultureel Rapport 2008.

ISBN 978-90-377-0368-9

scp-publicaties 2009

2009/1 *Nooit meer dezelfde. Gevolgen van misdrijven voor slachtoffers* (2009). Willemijn Lamet en Karin Wittebrood. ISBN 978-90-377-0402-0

2009/2 *Emancipatiemonitor 2008* (2009). Ans Merens en Brigitte Hermans (red.). ISBN 978-90-377-0406-8

2009/3 *Goede bureu kun je niet kopen. Over de woonconcentratie en woonpositie van niet-westerse allochtonen in Nederland* (2009). Jeanet Kullberg, Miranda Vervoort en Jaco Dagevos. ISBN 978-90-377-0401-3

2009/4 *De jeugd een zorg. Ramings- en verdeelmodel jeugdzorg 2009* (2009). John Stevens, Evert Pommer, Hetty van Kempen, Elke Zeijl, Isolde Woittiez, Klarita Sadiraj, Rob Gilsing, Saskia Keuzenkamp. ISBN 978-90-377-0355-9

2009/5 *Mantelzorg. Een overzicht van de steun van en aan mantelzorgers in 2007* (2009). Alice de Boer, Marjolein Broese van Groenou en Joost Timmermans (red.). ISBN 978-90-377-0408-2

2009/6 *At a glance. Summaries of 16 scp research projects in 2008.* ISBN 978-90-377-0413-6

2009/7 *Vrijwilligerswerk in meervoud. Civil society en vrijwilligerswerk 5* (2009). Paul Dekker en Joep de Hart. ISBN 978-90-377-0348-1

2009/8 *Cultuurbewonderaars en cultuurbeoefenaars. Trends in cultuurparticipatie en media-gebruik* (2009). Andries van den Broek, Jos de Haan en Frank Huysmans. ISBN 978-90-377-0400-6

2009/9 *Making up the Gap. Migrant Education in the Netherlands* (2009). Lex Herweijer. ISBN 978-90-377-0433-4

2009/10 *Gelukkig voor de klas?. Leraren voortgezet onderwijs over hun werk.* Ria Vogels. ISBN 978-90-377-0340-5

2009/11 *Rules of Relief. Institutions of social security, and their impact* (2009). J.C. Vrooman. ISBN 978-90-377-0218-7

2009/12 *Overlast en verloedering ontsleuteld* (2009). Lonneke van Noije en Karin Wittebrood. ISBN 978-90-377-0436-5

- 2009/13 *Crisis in aantocht? Verdiepingsstudie Continu Onderzoek Burgerperspectieven 2008* (2009). Paul Dekker, Tom van der Meer, Peggy Schyns en Eefje Steenvoorden. ISBN 978-90-377-0404-4
- 2009/14 *De sociale staat van Nederland 2009* (2009). Rob Bijl, Jeroen Boelhouwer, Evert Pommer, Peggy Schyns (red.). ISBN 978-90-377-0434-1
- 2009/15 *Deeltijd (g)een probleem. Mogelijkheden om de arbeidsduur van vrouwen met een kleine deeltijdbaai te vergroten* (2009). Saskia Keuzenkamp (red.), Carlien Hillebrink, Wil Portegijs, Babette Pouwels. ISBN 978-90-377-0448-8
- 2009/16 *De toekomst van de mantelzorg* (2009). Klarita Sadiraj, Joost Timmermans, Michiel Ras, Alice de Boer. ISBN 978-90-377-0435-8
- 2009/17 *Vergrijzing, verpleging en verzorging. Ramingen, profielen en scenario's 2005-2030* (2009). Isolde Woittiez, Evelien Eggink, Jedid-Jah Jonker, Klarita Sadiraj. ISBN 978-90-377-0422-8
- 2009/18 *Jaarrapport integratie 2009* (2009). Mérove Gijsberts, Jaco Dagevos (red.). ISBN 978-90-377-0446-4
- 2009/19 *Werkloos in crisistijd. Baanverliezers, inkomensveranderingen en sociale gevolgen; een verkenning* (2009). Cok Vrooman (red.). ISBN 978-90-377-451-8
- 2009/20 *Maten voor gemeenten. Een analyse van de prestaties van de lokale overheid* (2009). B. Kuhry en J.J.J. Jonker. ISBN 978-90-377-0452-5

SCP-publicaties 2010

- 2010/1 *Liever mark dan Mohammed? Onderzoek naar arbeidsmarktdiscriminatie van niet-westerse migranten via praktijktests* (2010). Iris Andriessen, Eline Nievers, Laila Faulk, Jaco Dagevos. ISBN 978-90-377-0421-1
- 2010/2 *Op weg met de Wmo. Evaluatie van de Wet maatschappelijke ondersteuning 2007-2009* (2010). Mirjam de Klerk, Rob Gilsing en Joost Timmermans (red.), Gijs van Houten, Anna Maria Marangos, Mathijs Tuynman, Martha Dijkgraaf, Jennifer van den Broeke, Judith van der Veer, Jelmer Shalk, Jurjen Iedema, Alice de Boer. ISBN 978-90-377-0463-1
- 2010/4 *Steeds meer verstandelijk gehandicapten? Ontwikkelingen in vraag en gebruik van zorg voor verstandelijk gehandicapten 1998-2008* (2010). Michiel Ras, Isolde Woittiez, Hetty van Kempen, Klarita Sadiraj. ISBN 978-90-377-0468-6
- 2010/5 *Een baanloos bestaan. De betekenis van werk voor werklozen, arbeidsongeschikten en werkenden* (2010). Patricia van Echtelt. ISBN 978-90-377-0350-4
- 2010/6 *The social state of the Netherlands 2009* (2010). Rob Bijl, Jeroen Boelhouwer, Evert Pommer, Peggy Schyns (red.). ISBN 978-90-377-0466-2
- 2010/7 *The minimum agreed upon. Consensual budget standards for the Netherlands* (2010). Stella Hoff, Arjan Soede, Cok Vrooman, Corinne van Gaalen, Albert Luten, Sanne Lamers. ISBN 978-90-377-0472-3
- 2010/8 *Sociale uitsluiting bij kinderen: omvang en achtergronden* (2010). Annette Roest, Anne marie Lokhorst, Cok Vrooman. ISBN 978-90-377-0493-8
- 2010/9 *Bepert aan het werk. Rapportage ziekteverzuim, arbeidsongeschiktheid en arbeidsparticipatie* (2010). Gerda Jehoel-Gijsbers (red.). ISBN 978-90-377-0489-1
- 2010/10 *Minder werk voor laagopgeleiden? Ontwikkelingen in baanbezit en baankwaliteit 1992-2008* (2010). Edith Josten. ISBN 978-90-377-0474-7
- 2010/11 *At home in the Netherlands* (2010). Mérove Gijsberts, Jaco Dagevos. ISBN 978 90 377 0487 7

- 2010/12 *FAQ's over kunstbeoefening in de vrije tijd* (2010). Andries van den Broek. ISBN 978 90 377 0455 6
- 2010/13 *Mogelijkheden tot kunstbeoefening in de vrije tijd* (2010). Andries van den Broek (red.). ISBN 978 90 377 0456 3
- 2010/14 *Toekomstverkenning kunstbeoefening* (2010). Andries van den Broek. ISBN 978 90 377 0491 4
- 2010/15 *Steeds gewoner, nooit gewoon. Acceptatie van homoseksualiteit in Nederland* (2010). Saskia Keuzenkamp et al. ISBN 978 90 377 0501 0
- 2010/16 *Publieke dienstverlening in perspectief. SCP-memorandum voor de kabinetsformatie 2010* (2010). Evert Pommer en Evelien Eggink (red.). ISBN 978 90 377 0513 3
- 2010/17 *Wellbeing in the Netherlands. The SCP life situation index since 1974* (2010). Jeroen Boelhouwer. ISBN 978 90 377 0345 0
- 2010/18 *Minderheden in de mixed-mode? Een inventarisatie van voor- en nadelen van het inzetten van verschillende dataverzamelmethode onder niet-westerse migranten* (2010). R. Feskens, J. Kappelhof, J. Dagevos, I. Stoop. ISBN 978 90 377 0517 1
- 2010/20 *Discriminatiemonitor niet-westerse migranten op de arbeidsmarkt 2010* (2010). Eline Nievers en Iris Andriessen (red.). ISBN 978 90 377 0438 9
- 2010/21 *Iemand moet het doen. Ervaringen van verzorgers van partners* (2010). Judith van Male, Marion Duimel en Alice de Boer. ISBN 978 90 377 0518 8
- 2010/23 *Het werken waard. Het arbeidsaanbod van laagopgeleide vrouwen vanuit een economisch en sociologisch perspectief* (2010). Mariëlle Cloïn. ISBN 978 90 377 0514 0
- 2010/24 *Zorgen voor zorg. Ramingen van de vraag naar personeel in verpleging en verzorging tot 2030* (2010). Evelien Eggink, Debbie Oudijk en Isolde Woittiez. ISBN 978 90 377 0512 6
- 2010/25 *Alle kanalen staan open. De digitalisering van mediagebruik* (2010). Frank Huysmans en Jos de Haan. ISBN 978 90 377 0521 8

SCP-essays

- 1 *Voorbeelden en nabeelden* (2005). Joep de Hart. ISBN 90-377-0248-1
- 2 *De stem des volks* (2006). Arjan van Dixhoorn. ISBN 90-377-0265-1
- 3 *De tekentafel neemt de wijk* (2006). Jeanet Kullberg. ISBN 90-377-0261-9
- 4 *Leven zonder drukte* (2006). Tjirk van der Ziel met een naschrift van Anja Steenbekkers en Carola Simon. ISBN 90-377-0262-7
- 5 *Otto Neurath en de maakbaarheid van de betere samenleving* (2007). Ferdinand Mertens. ISBN 978-90-5260-260-8

Overige publicaties

- Continu Onderzoek Burgerperspectieven. Kwartaalbericht 2009. Deel 1* (2009). Eefje Steenvoorden, Peggy Schyns en Tom van der Meer. ISBN 978-90-377-0417-4
- Continu Onderzoek Burgerperspectieven. Kwartaalbericht 2009. Deel 2* (2009). Tom van der Meer, Paul Dekker en Eefje Steenvoorden. ISBN 978-90-377-0418-1
- Continu Onderzoek Burgerperspectieven. Kwartaalbericht 2009. Deel 3* (2009). Eefje Steenvoorden, Tom van der Meer en Paul Dekker. ISBN 978-90-377-0447-1
- Wmo Evaluatie. Eerste tussenrapportage. De invoering van de Wmo: gemeentelijk beleid in 2007* (2008). Gijs van Houten, Mathijs Tuynman en Rob Gilsing. ISBN 978-90-377-0390-0

- Wmo Evaluatie. Tweede tussenrapportage. Ondersteuning en participatie van mensen met een lichamelijke beperking* (2009). Anna Maria Marangos, Mieke Cardol en Mirjam de Klerk. ISBN 978-90-377-0399-3
- m/v. SCP-nieuwjaarsuitgave 2009* (2009). ISBN 978-90-377-0411-2
- Definitief advies over het Wmo-budget huishoudelijke hulp voor 2009* (2009). Evert Pommer, Ab van der Torre, Evelien Eggink. ISBN 978-90-377-0415-0
- TBO/eu en TBO/nl. Een vergelijking van twee methoden van tijdbestedingsonderzoek* (2009). Carlijn Kamphuis, Remko van den Dool, Andries van den Broek, Ineke Stoop, Patty Adelaar, Jos de Haan. ISBN 978-90-377-0423-5
- Kunnen alle kinderen meedoen? Onderzoek naar de maatschappelijke participatie van arme kinderen* (2009). Gerda Jehoel-Gijsbers. ISBN 978-90-377-0416-7
- Werken en weldoen. Kiezen voor betaalde en onbetaalde arbeid* (2009). Ingrid Ooms, Jedid-Jah Jonker, Ab van der Torre. ISBN 978-90-377-0403-7
- Genoeg om van te leven. Focusgroepen in discussie over de minimale kosten van levensonderhoud* (2009). Stella Hoff, Arjan Soede, Cok Vrooman, Corinne van Gaalen, Albert Luten, Sanne Lamers. ISBN 978-90-377-407-5
- Profielen van vragers naar AWBZ-GGZ* (2009). Cretien van Campen. ISBN 978-90-377-0444-0
- Sociale uitsluiting: een meetinstrument* (2009). Gerda Jehoel-Gijsbers (SCP), Wendy Smits (CBS), Jeroen Boelhouwer (SCP) en Harry Bierings (CBS). ISBN 978-90-377-0427-3
- Sociale samenhang in de wijk. nsv actualiteitencollege 2008* (2009). Maurice Gesthuizen en Vic Veldheer (red.). ISBN 978-90-377-0445-7
- Meten wat leeft? Achtergrondstudie bij het Continu Onderzoek Burgerperspectieven* (2009). Paul Dekker (red.). ISBN 978-90-377-0419-8
- VeVeRa-III. Ramingen verpleging en verzorging 2005-2030 modelbeschrijving* (2009). Evelien Eggink, Jedid-Jah Jonker, Klarita Sadiraj, Isolde Woittiez. ISBN 978-90-377-0356-6
- Sporten gemeten. Methodologische aspecten van het onderzoek naar sportdeelname* (2009). Koen Breedveld, Annet Tiessen-Raaphorst. ISBN 978-90-377-0358-0
- Hoe het ons verging... Traditionele nieuwjaarsuitgave van het SCP* (2010). Paul Schnabel (red.). ISBN 978-90-377-0465-5
- Continu Onderzoek Burgerperspectieven. Kwartaalbericht 2009. Deel 4* (2010). Paul Dekker, Tom van der Meer en Irene de Goede. ISBN 978-90-377-0464-8
- Wmo Evaluatie. Vierde tussenrapportage. Ondersteuning en participatie van mensen met een lichamelijke beperking; twee jaar na de invoering van de Wmo* (2010). A. Marangos, M. Cardol, M. Dijkgraaf, M. de Klerk. ISBN 978-90-377-0470-9
- Op weg met de Wmo. Journalistieke samenvatting door Karolien Bais. Mirjam de Klerk, Rob Gilsing en Joost Timmermans. Samenvatting door Karolien Bais* (2010). ISBN 978-90-377-0469-3
- NL Kids online. Risico's en kansen van internetgebruik onder jongeren* (2010). Jos de Haan. ISBN 978-90-377-0430-3
- Kortdurende thuiszorg in de AWBZ. Een verkenning van omvang, profiel en afbakening* (2010). Maaïke den Draak. ISBN 978-90-377-0471-6
- De publieke opinie over kernenergie* (2010). Paul Dekker, Irene de Goede, Joop van der Pligt. ISBN 978-90-377-0488-4
- Op maat gemaakt? Een evaluatie van enkele responsverbeterende maatregelen onder Nederlanders van niet-westerse afkomst* (2010). J. Kappelhof. ISBN 978-90-377-0495-2
- Oudere tehuisbewoners. Landelijk overzicht van de leefsituatie van ouderen in instellingen 2008/2009* (2010). Maaïke den Draak. ISBN 978-90-377-0499-0
- Kopers in de knel? Een scenariostudie naar de gevolgen van de crisis voor huiseigenaren met een hypotheek* (2010). Michiel Ras, Ingrid Ooms, Evelien Eggink. ISBN 978-90-377-0498-3

Gewoon anders. Acceptatie van homoseksualiteit in Nederland (2010). Saskia Keuzenkamp.

ISBN 978 90 377 0502 7

De aard, de daad en het Woord. Een halve eeuw opinie- en besluitvorming over homoseksualiteit in protestants Nederland, 1959-2009 (2010). David Bos. ISBN 978 90 377 0506 5

Werkloosheid in goede banen. Bijdragen aan de SCP-studiemiddag 2010 (2010). Patricia van Echtelt (red.). ISBN 978 90 377 0516 4