

Vragenlijstontwikkeling

10

Deirdre Giesen, Vivian Meertens, Rachel Vis-Visschers en Dirkjan Beukenhorst

Statistische Methoden (10005)



Verklaring van tekens

.	= gegevens ontbreken
*	= voorlopig cijfer
**	= nader voorlopig cijfer
x	= geheim
–	= nihil
–	= (indien voorkomend tussen twee getallen) tot en met
0 (0,0)	= het getal is kleiner dan de helft van de gekozen eenheid
niets (blank)	= een cijfer kan op logische gronden niet voorkomen
2008–2009	= 2008 tot en met 2009
2008/2009	= het gemiddelde over de jaren 2008 tot en met 2009
2008/'09	= oogstjaar, boekjaar, schooljaar enz., beginnend in 2008 en eindigend in 2009
2006/'07–2008/'09	= oogstjaar, boekjaar enz., 2006/'07 tot en met 2008/'09

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

Colofon

Uitgever
Centraal Bureau voor de Statistiek
Henri Faasdreef 312
2492 JP Den Haag

Prepress
Centraal Bureau voor de Statistiek - Grafimedia

Omslag
TelDesign, Rotterdam

Inlichtingen
Tel. (088) 570 70 70
Fax (070) 337 59 94
Via contactformulier: www.cbs.nl/infoservice

Bestellingen
E-mail: verkoop@cbs.nl
Fax (045) 570 62 68

Internet
www.cbs.nl

ISSN: 1876-0333

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen, 2010.
Verveelvoudiging is toegestaan, mits het CBS als bron wordt vermeld.

Inhoudsopgave

1.	Inleiding op het thema	5
1.1	Algemene beschrijving en leeswijzer	5
1.2	Afbakening en relatie met andere thema's	6
1.3	Plaats in het statistisch proces	6
1.4	Definities	7
2.	Vragenlijstontwikkeling als proces	8
2.1	Conceptualisering en onderzoeksontwerp	8
2.2	Vragenlijstontwerp	10
2.3	Testen	11
2.4	Herziening	12
2.5	Dataverzameling	13
2.6	Proces-monitoring en evaluatie kwaliteit vragenlijst	13
3.	Vragenlijstontwerp	14
3.1	Inleiding	14
3.2	Algemene uitgangspunten bij het ontwerpen van vragenlijsten	15
3.3	Ontwerpen van enquêtevragen	23
3.4	Ontwerpen van antwoordmogelijkheden	29
3.5	Gebruik van toelichtingen en instructies	32
3.6	Richtlijnen voor de structuur van de vragenlijst	33
3.7	Vormgeving van vragenlijsten	34
4.	Methoden voor het ontwikkelen, testen en evalueren van vragenlijsten	37
4.1	Inleiding	37
4.2	Welke methode voor welke fase?	37
4.3	Informele tests	38
4.4	Expertraadpleging	39
4.5	Focusgroepen	41
4.6	Cognitieve interviews	43
4.7	Vignetten en invulling van een vragenlijst aan de hand van een praktijkvoorbeeld	48
4.8	Observeren van respondenten en interviewers	51
4.9	Analyse van kwalitatief onderzoeksmateriaal	53
4.10	Technisch testen	54

4.11	Experimenten.....	55
4.12	Analyses van procesgegevens	56
4.13	Analyses van surveydata	58
5.	Afsluiting	59
6.	Literatuur	61
	Bijlage 1. Taalniveaus.....	65
	Bijlage 2. Checklist Vragenlijstontwikkeling	67
	Bijlage 3. Scoreformulier expertraadpleging vraagstelling	74
	Bijlage 4. Methode Interviewerevaluatie.....	76
	Bijlage 5a. Voorbeeld observatieformulier CAPI / CATI interview	80
	Bijlage 5b. Voorbeeld observatieformulier CAWI vragenlijst	81
	Bijlage 6. Standaardprotocol bedrijfsbezoeken Vragenlijstevaluatie	82

1. Inleiding op het thema

1.1 Algemene beschrijving en leeswijzer

Dit deel van de Methodenreeks behandelt de ontwikkeling en evaluatie van vragenlijsten, hierna kortweg vragenlijstontwikkeling genoemd. Vragenlijstontwikkeling kan plaatsvinden bij het ontwikkelen van nieuwe statistische producten, maar ook bij een evaluatie of herontwerp van bestaande producten en processen. De vragenlijstontwikkeling die hier besproken wordt, is van toepassing op alle binnen het CBS gangbare modes van onderzoek (CAPI, CATI, webvragenlijsten en papieren vragenlijsten) en voor alle doelgroepen (vragenlijsten voor personen en huishoudens en voor bedrijven en instellingen).

Grofweg worden bij de ontwikkeling van een vragenlijst drie doelen nagestreefd, die soms tegenstrijdig zijn. In de eerste plaats moet de vragenlijst leiden tot het verzamelen van de gegevens die nodig zijn om de onderzoeksvragen te beantwoorden - daartoe dienen deze gegevens valide, betrouwbaar en relevant te zijn (zie hoofdstuk 3). Daarnaast dient de enquête voor de respondent en interviewer zo prettig en gemakkelijk mogelijk te zijn. Tot slot dient het ontwikkelen, verwerken en aanpassen van vragenlijsten efficiënt te kunnen gebeuren. Het kan belangrijk zijn dat snel op de actualiteit kan worden gereageerd. Dat geldt niet alleen voor nieuwe thema's maar ook voor wetswijzigingen die aanpassingen noodzakelijk maken (denk aan de zorgverzekeringen).

De basis voor het opstellen van dit onderdeel van de Methodenreeks is naast de uitgebreide literatuur en de ervaring binnen het CBS het "Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System" (Brancato et al., 2006). De richtlijnen in dit handboek zijn een specificatie van de vereisten zoals die zijn beschreven in de Europese Praktijkcode voor Statistische Autoriteiten (Commissie voor de Europese Gemeenschappen, 2005). Hierin is ondermeer als kwaliteitsindicator opgenomen: "Bij statistische onderzoeken worden de vragenlijsten systematisch getest voordat de gegevensverzameling begint".

Dit onderdeel van de Methodenreeks is als volgt opgebouwd: Hoofdstuk 2 beschrijft kort het gehele proces van vragenlijstontwikkeling. Vervolgens komt in hoofdstuk 3 het onderdeel vragenlijstontwerp uitvoerig aan bod. We behandelen daarbij enkele centrale theoretische uitgangspunten bij het vragenlijstontwerp, zoals het vraagantwoordproces en geven vervolgens concrete richtlijnen voor het formuleren van enquêtevragen en het structureren en vormgeven van vragenlijsten. Hoofdstuk 4 gaat uitgebreid in op de methoden en technieken die gebruikt worden bij het ontwikkelen van vragenlijsten. Hierin behandelen we zowel klassieke methoden die gebruikt worden in een laboratoriumsetting, zoals cognitieve interviews, als ook methoden die bruikbaar zijn voor vragenlijsten die al in het veld zijn, zoals het evalueren van vragenlijsten met interviewers.

Tot slot is er aan dit deel van de Methodenreeks een aantal bijlagen toegevoegd met gedetailleerde informatie. Hierin is onder andere een checklist opgenomen voor het ontwerpen van vragenlijsten.

1.2 Afbakening en relatie met andere thema's

Dit thema van de Methodenreeks beschrijft het proces van vragenlijstontwikkeling. Ook opgenomen zijn de beschrijvingen van methoden die gebruikt worden bij het ontwikkelen van vragenlijsten, waaronder diverse testmethoden. Deze methoden zijn een essentieel onderdeel van de vragenlijstontwikkeling.

De specifieke werkprocessen binnen het CBS en technische detailbeschrijvingen zijn niet opgenomen. Deze vallen buiten het bereik van de Methodenreeks.

Een nieuw proces van vragenlijstontwikkeling start pas wanneer men de gewenste informatie niet op basis van bestaande enquêtes of registers kan verkrijgen. Het waarneembeleid van het CBS is immers dat in principe alleen primaire gegevens worden verzameld als er geen secundaire informatie voorhanden is (Beukenhorst, van Dam, Göttgens, Snijkers & Verbruggen, 2005). Zaken die verband houden met ontsluiten van gegevens uit registers et cetera kan men vinden in de andere (deel)thema's van de Methodenreeks zoals in het onderdeel Secundaire waarneming en XBRL. Overige gerelateerde onderdelen van de Methodenreeks zijn Organisatie van het veldwerk en Benaderingsstrategieën (waar onder Mixed Mode benadering).

1.3 Plaats in het statistisch proces

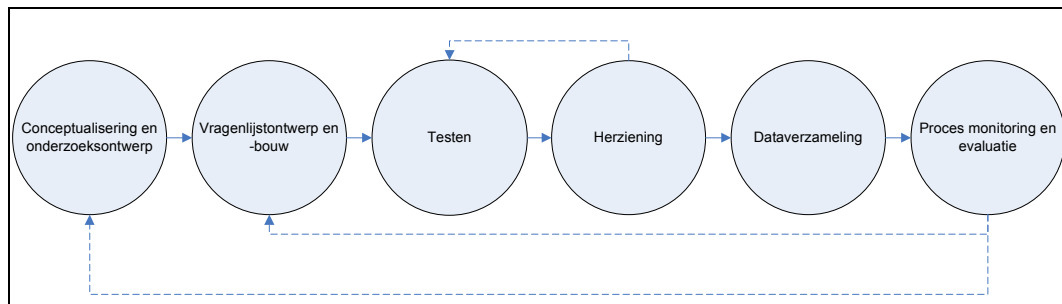
In termen van het Statistisch Proces zoals dat binnen de Methodenreeks is beschreven door Willeboordse (2008) speelt de vragenlijstontwikkeling vooral in de ontwerpfase. Vragenlijstontwikkeling start met het specificeren van de benodigde output, dat wil zeggen dat men eerst moet vaststellen welke statistieken uiteindelijk opgeleverd moeten worden, en eindigt met de evaluatie van het proces van gegevensverzameling en van de kwaliteit van de daadwerkelijk verzamelde gegevens. Zie ook hoofdstuk 2 voor een uitgebreidere beschrijving van vragenlijstontwikkeling als proces.

1.4 Definities

Begrip	Omschrijving
Berichtgever	Terminologie die gebruikt wordt bij bedrijfsenquêtes: persoon of instantie die de vragen beantwoordt. Het equivalent van respondent bij persoonswaarneming.
CAPI	Computer Assisted Personal Interviewing, ook bekend als face-to-face interviews. Interviews waarbij de interviewer persoonlijk aanwezig is en met behulp van een computergestuurde vragenlijst de enquêtevragen stelt. Doorgaans vinden CAPI enquêtes bij de respondent thuis plaats.
CATI	Computer Assisted Telephone Interviewing, telefonisch enquêteren met behulp van een computergestuurde vragenlijst.
CAWI	Computer Assisted Web Interviewing, computer gestuurde vragenlijst waarbij de respondent zelf de vragen invult via het internet.
Enquête	De term wordt tweeledig gebruikt: <ul style="list-style-type: none"> •Onderzoek met gebruikmaking van een gestandaardiseerde vragenlijst. •Synoniem voor de (gestandaardiseerde) vragenlijst zelf.
Non-respons	<ul style="list-style-type: none"> • Unit non-respons: waarneembaarheid waarover geen vragenlijstgegevens worden verkregen • Item non-respons: vraag in de vragenlijst waarop geen antwoord wordt verkregen (terwijl dat bij een of meer andere vragen wel het geval is d.w.z zonder dat er sprake is van unit non-respons)
Onderzoekspersoon (OP)	De persoon over wie data worden verzameld. Dit kan de respondent zijn, maar dit is niet noodzakelijk. Zie respondent.
PAPI	Paper and Pencil Interviewing. Schriftelijke vragenlijsten die de respondent zelf invult. Ze kunnen worden toegestuurd over de post of worden verspreid door een interviewer.
Primaire data	Binnen het CBS: Gegevens die door het CBS zelf of in opdracht daarvan zijn verzameld.
Respondent	De persoon die de vragen beantwoordt. Doorgaans is dit de onderzoekspersoon (OP) maar de respondent kan ook als proxy optreden voor OP of vragen betreffende het huishouden beantwoorden.
Statistische eenheid	Eenheid waarover de statistiek rapporteert
Survey	Onderzoeksvorm waarbij geselecteerde personen of andere eenheden op de gestandaardiseerde wijze worden onderzocht.
Secundaire data	Binnen het CBS: Gegevens die door anderen dan het CBS zijn verzameld (en niet in opdracht van het CBS).
Vragenlijst	Het geheel van vragen dat aan een respondent gesteld wordt.
Vragenlijstontwerp	Het maken van de teksten, structuur en vormgeving van de vragenlijst.
Vragenlijstontwikkeling	Het hele proces van het opstellen van een vragenlijst: van conceptualisering en onderzoeksontwerp tot aan de uiteindelijke evaluatie.
Waarneembaarheid	De eenheden waarover data worden verzameld. Deze eenheden kunnen later tot statistische eenheden worden geaggregeerd of gedesaggregeerd.

2. Vragenlijstontwikkeling als proces

Vragenlijstontwikkeling behelst meer dan alleen het opstellen van vraagteksten en vormgeving hiervan. Vragenlijstontwikkeling is een proces waarbij de volgende zes stappen onderscheiden kunnen worden (zie figuur 2.1, gebaseerd op Brancato et al., 2006):



Figuur 2.1 Vragenlijstontwikkeling als proces

De zes stappen en de verbindingspijlen laten zien dat het proces van vragenlijstontwikkeling een iteratief proces is waarbij informatie uit elke fase in het proces als input kan dienen voor een volgende of eerdere fase. Bij dit proces zijn niet alleen vragenlijstspecialisten betrokken maar ook inhoudelijk deskundigen en andere specialisten.

2.1 Conceptualisering en onderzoeksontwerp

De eerste stap bij het ontwikkelen van een vragenlijst is de conceptualisering: het afbakenen van het onderwerp en de te meten variabelen. Theoretische concepten van onderzoek zijn vaak complex, zoals armoede, toegang tot de gezondheidszorg, outsourcing, of slachtofferschap van criminaliteit. Dergelijke concepten kunnen niet één op één worden vertaald in enquêtevragen. Eerst moet nauwkeurig worden bepaald wat de opdrachtgevers en/of gebruikers van de data precies willen weten en vervolgens moet verkend worden in hoeverre deze begrippen inderdaad meetbaar zijn bij de doelgroep. Uiteindelijk moet dit leiden tot concreet meetbare begrippen. Deze in principe meetbare begrippen worden indicatoren genoemd. Bij het vaststellen van de indicatoren zijn de inhoudelijk deskundigen nauw betrokken. De indicatoren moeten vervolgens in één of meer enquêtevragen worden vertaald. Dit laatste deel van het proces wordt operationalisering genoemd. Daar wordt de rol van de vragenlijstontwikkelaars belangrijker. Een voorbeeld: bij het begrip armoede zal het nodig zijn om te definiëren wat precies onder armoede verstaan wordt. Een mogelijke operationalisering hiervan is bijvoorbeeld: een alleenstaande is arm indien deze een inkomen heeft dat lager is dan 50% van het minimumloon, of een gezin is arm indien dit een inkomen geniet dat lager is dan 70% van het minimumloon. In de vragenlijst zal dan dus een vraag opgenomen moeten worden over het inkomen van de respondent en over de samenstelling van het huishouden. Overigens ontleent het

CBS al enige tijd het inkomen van personen en huishoudens aan gegevens van de Belastingdienst.

Voor een goede operationalisering is het zinvol om bestaande vakliteratuur, interne documenten en deskundigen te raadplegen (door middel van een review van bestaande documenten of een expertraadpleging). Denk daarbij aan inhoudelijke en methodologische vakliteratuur, maar ook aan interne documenten met bijvoorbeeld verslagen van interviewerevaluaties of kwaliteitsanalyses van gerelateerde reeds bestaande vragenlijsten, rapporten van internationale zusterbureaus en deskundigen die kennis hebben van het onderwerp of het veld van onderzoek. Indien er intern of extern al onderzoek is gedaan naar het onderwerp, probeer dan de daarvoor gebruikte vragenlijsten op te sporen en erachter te komen hoe goed deze hebben gefunctioneerd.

Bij geheel nieuwe onderwerpen of nieuwe doelgroepen is het zinvol om in deze fase ook onderzoek bij potentiële respondenten te doen, om te verkennen in hoeverre het beoogde onderzoeksontwerp en de concepten die gemeten moeten worden aansluiten bij hun beleving. Dit kan gebeuren met gestructureerde groepsgesprekken (focusgroepen) of diepte-interviews met proefrespondenten (cognitieve interviews). Deze beide methoden komen uitgebreid aan bod in hoofdstuk 4.

Nadat vastgesteld is wat gemeten gaat worden, moet besloten worden bij wie dit gemeten gaat worden en hoe. Eerst moet bepaald worden welke statistische eenheid relevant is. Heeft de informatie die men wil weten betrekking op bijvoorbeeld een persoon of op een huishouden, op een bedrijfseenheid of op een afdeling binnen een bedrijf? De statistische eenheid is de eenheid waarop de statistiek wordt gebaseerd. Verder moet vastgesteld worden op welke niveau de data worden waargenomen. Bij bedrijfswaarneming kan deze waarneem eenheid groter of kleiner zijn dan de statistische eenheid waarover wordt gerapporteerd (moederconcern, holding, vestiging etc). De zogenaamde berichtgever is degene die daadwerkelijk de vragenlijst beantwoordt. Bij persoonswaarneming kan een persoon (de OnderzoeksPersoon OP) de waarneem eenheid zijn of een huishouden. De respondent kan verschillen van de onderzoekspersoon (OP kan een kind zijn voor wie de vragen worden beantwoord door een ouder/verzorger). Ook kan de respondent informatie verstrekken over hele huishouden of alle leden daarvan. Een volgende belangrijke keuze is of er een interviewer wordt ingezet en welke waarneemmodus wordt gebruikt. Om tot een goede keuze te komen kunnen ervaringen met eerder uitgevoerde enquêtes behulpzaam zijn. Bekijk bijvoorbeeld responsoverzichten en statistische analyses van (andere) eerder uitgevoerde enquêtes. Voor het kiezen van een onderzoeksontwerp kan het ook zinvol zijn het veld in te gaan. Zo is bijvoorbeeld voor een nieuw onderzoek naar het uitbesteden van bedrijfsactiviteiten bij potentiële berichtgevers gevraagd wie in hun bedrijf de vragenlijst het beste kon beantwoorden (Vis-Visschers, 2007).

Samengevat omvat de fase van conceptualisering en onderzoeksontwerp dus de afbakening van het onderzoeksterrein. Bij de conceptualisering wordt bepaald wat er precies gemeten wordt met de vragenlijst, met andere woorden: welke variabelen de

vragenlijst moet opleveren. Uit het onderzoeksontwerp volgt een aantal voorwaarden waaraan de vragenlijst moet voldoen.

Tenslotte: keuzes documenteren!

Bij statistische bureaus blijven enquêtes vaak lang in het veld of worden met lange tussenpozen herhaald. Bovendien is er binnen de organisatie mobiliteit onder de medewerkers. Hierdoor verdwijnt de kennis over het ontwikkelproces. Een goede documentatie van de keuzes en afwegingen die in deze fase gemaakt zijn is daarom van groot belang. Zowel voor latere gebruikers van de data als voor vragenlijstontwikkelaars is bij een herziening van het onderzoek dergelijke documentatie bijzonder nuttig.

2.2 Vragenlijstontwerp

In de fase vragenlijstontwerp en -bouw wordt een tekstversie van de vragenlijst gemaakt en wordt de vormgeving ervan ontworpen. Het onderzoeksontwerp, en met name de waarnemmodus, is daarbij uiteraard een belangrijke randvoorwaarde, evenals de onderzoekspopulatie. Zo biedt bijvoorbeeld een papieren vragenlijst veel minder mogelijkheden voor complexe routing dan een elektronische vragenlijst. Zie voor meer informatie over het kiezen van de modus van waarneming het Methodenreeksdeel Benaderingsstrategieën.

Idealiter gebeurt vragenlijstontwerp in teamverband. In dat team dienen deskundigen te zitten op de volgende gebieden: de inhoud van de enquête (doorgaans de statistische afdeling), het vragenlijstontwerp en de logistiek van de dataverzameling en -verwerking. Bij complexe en nieuwe PAPI en CAWI vragenlijsten is het bovendien aan te raden advies in te winnen bij een grafisch vormgever.

Een eerste stap in deze fase is het bepalen van een - voor de respondent!- logische structuur voor de vragenlijst. Hierin wordt de volgorde vastgelegd waarin de verschillende onderwerpen (en eventueel respondenten) aan bod zullen komen. Vervolgens wordt de volgorde van de vragen per onderwerp bepaald. De variabelen die zijn gedefinieerd in de vorige fase worden daarbij vertaald naar concrete enquêtevragen. Bij het kiezen van de variabelen en vragen moet steeds bedacht worden welke vraag voor welke respondent relevant is. Door middel van routing worden respondenten naar de voor hen relevante vragen geleid.

Zodra een versie van de routing en de teksten beschikbaar is, wordt ook de vormgeving van de vragenlijst ontwikkeld. Teksten worden immers in een bepaald format gepresenteerd, afhankelijk van de modus van waarneming en de in de organisatie gebruikte standaarden. Denk hierbij bijvoorbeeld aan interviewer-instructies, toelichtingen op vraagteksten, kopjes boven bij elkaar horende blokken van vragen en informatie over de routing. Zeker bij PAPI en CAWI vragenlijsten, waarbij de respondenten zelf de vragenlijst zien, is vormgeving een essentieel onderdeel van de ontwerpfase die in een zo vroeg mogelijk stadium aan bod moet komen. Soms kan het bijvoorbeeld nodig zijn om teksten of vraagvolgordes aan te passen aan de beschikbare plek op een pagina of beeldscherm.

Bij het aanpassen van een bestaande vragenlijst dienen – indien voor handen - documenten over monitoring en evaluaties met interviewers of respondenten van de reeds geïmplementeerde enquête aan een review te worden onderworpen. Deze informatie kan aanleiding zijn het vragenlijstontwerp en/of de vragenlijst te wijzigen. Ook verslagen van focusgroepen met interviewers en respondenten dienen dit doel. Vragenlijstontwerp komt uitgebreid aan bod in hoofdstuk 3, focusgroepen worden beschreven in hoofdstuk 4.

2.3 Testen

Het testen van een conceptvragenlijst kan zeer verschillende doelen hebben. Men kan bijvoorbeeld testen of een computergestuurde vragenlijst correct gebouwd is, of cruciale vragen door de respondenten goed begrepen en beantwoord kunnen worden, of de vragenlijst prettig loopt voor interviewer en respondent en hoe lang het afnemen van de vragenlijst duurt. Een minimumvereiste van Eurostat is dat een nieuwe vragenlijst in ieder geval één keer met potentiële respondenten getest moet zijn (Brancato et al., 2006, pagina XI). Bij reeds in het veld zijnde vragenlijsten of bij aanpassingen van de vragenlijst dienen ten aanzien van dit vereiste kosten en baten te worden afgewogen. Hieronder wordt een kort overzicht gegeven van de meest gebruikte methoden in deze fase van vragenlijstontwikkeling. De belangrijkste methoden die hier genoemd worden, worden in hoofdstuk 4 nader uitgewerkt.

Informele tests

De informele test is het eenvoudigst uit te voeren. Hierbij vult de vragenlijstmaker bijvoorbeeld zelf eens serieus de vragenlijst in, leest alle teksten eens hardop voor of interviewt een bekende. Deze manier van testen geeft een eerste inzicht in de opleesbaarheid, begrijpelijkheid en beantwoordbaarheid van de vragen.

Expert raadpleging

Bij een expertraadpleging of ‘expert review’ wordt een groep deskundigen gevraagd de vragenlijst te beoordelen.

Kleinschalige, kwalitatieve tests

In de testfase worden vaak kleinschalige, kwalitatieve tests gedaan om de respondent- en interviewervriendelijkheid van sommige vragen of van de gehele vragenlijst te onderzoeken. Dit kan bijvoorbeeld door middel van *focusgroepen* en/of *cognitieve interviews* met proefpersonen. *Observaties* van proefpersonen die een vragenlijst invullen of interviewers die de vragenlijst afnemen zijn in dit stadium ook zinvol en kunnen in combinatie met de eerder genoemde methoden worden gebruikt.

Functioneel of technisch testen

Bij computergestuurde vragenlijsten is het van belang dat goed wordt getest of de elektronische vragenlijst volgens de specificaties gebouwd is en in de juiste volgorde verloopt. Gecontroleerd moet worden of alle routings kloppen en specifieke groepen

alle voor hen bedoelde vragen krijgen. Ook is belangrijk na te gaan of op de antwoordmogelijkheden ‘weet niet’ of ‘weigert’ de juiste vervolgvragen worden aangeboden. Hiervoor is de functionele of technische test noodzakelijk.

Kleinschalige, operationele tests

In een operationele test, ook wel pilot of “dress rehearsal” genoemd, wordt de complete dataverzamelingsprocedure kleinschalig uitgeprobeerd. Hierbij wordt met name getest of de logistieke procedures goed werken in het veld.

Grootschalige pilots en experimenten

In een grootschalige pilot en bij experimenten worden gegevens verzameld die ook kwantitatief geanalyseerd kunnen worden. Daarbij kan gekeken worden naar non-respons op vragen of onderdelen van de vragenlijst en naar de kwaliteit van de verzamelde data. Experimenten zijn zinvol wanneer er wordt getwijfeld tussen alternatieve vraagstellingen of alternatieve methoden en twee of meer vragenlijsten met elkaar vergeleken moeten worden.

2.4 Herziening

Uit de resultaten van het testen van de vragenlijst en (eventueel) de dataverzamelingsprocedure komen doorgaans allerlei wijzigingsvoorstellen naar voren. Of en hoe deze worden geïmplementeerd vergt soms uitgebreide discussies met alle betrokkenen. Hierbij moeten vaak lastige afwegingen gemaakt worden tussen verschillende belangen, bijvoorbeeld de vergelijkbaarheid met andere vragenlijsten of eerdere waarnemingen (tijdreeksen!) enerzijds en het gebruiksgemak van een vragenlijst voor respondenten en interviewers anderzijds. De aangepaste vragenlijst moet idealiter opnieuw inhoudelijk en zeker technisch worden getest. Het is namelijk goed mogelijk dat wijzigingen in de vragenlijst nieuwe problemen opleveren.

Voor een verantwoord ontwikkel- en testtraject is veel tijd en geld nodig. Beide zijn echter in de praktijk vaak schaars. Daarom moet veel zorg worden besteed aan de planning van zowel de vragenlijstontwikkeling als het testtraject. Een inderhaast samengestelde vragenlijst gaan testen waarvan bij voorbaat al vaststaat dat er veel problemen zullen optreden, heeft weinig zin. Anderzijds is zo veel tijd besteden aan de vragenlijstontwikkeling dat geen tijd overblijft voor testen evenmin verstandig. Indien een vragenlijst ongetest het veld ingaat en bij evaluaties geen problemen naar voren komen, betekent dit weliswaar dat het geen slechte vragenlijst is maar zeker niet dat het de best mogelijke vragenlijst is. Een uitgebreid testtraject kan veel problemen voorkómen maar kan evenmin garanderen dat de vragenlijst de best mogelijke is geworden. Om verantwoord te kunnen ontwikkelen en testen moet voldoende tijd worden ingeruimd waarbij ermee rekening moet worden gehouden dat er ook tijd genoeg is om wijzigingen te ontwikkelen, te testen en door te voeren.

2.5 Dataverzameling

Als de vragenlijst gereed is kan deze worden geïmplementeerd in het veld. In deze fase worden de data verzameld. Eventueel kunnen in de vragenlijst aanvullende, evaluerende vragen worden opgenomen voor respondenten of interviewers. Denk daarbij aan vragen over de interview- of invultijd, hoe makkelijk of moeilijk de vragen waren, hoe zeker men is over een bepaald antwoord of hoe de interviewer vond dat het gesprek is verlopen. Voor met name veldinterviewers kan dit echter lastig zijn. Daarom moet dit instrument zuinig worden ingezet. Er bestaan methoden waarbij interviewers per vraag alleen hoeven aan te geven of er ‘een probleem’ optrad. In een latere fase wordt dan nagegaan hoe vaak en waarom deze problemen zich voordeden.

2.6 Proces-monitoring en evaluatie kwaliteit vragenlijst

Na implementatie van de vragenlijst in het veld is het proces van vragenlijst-ontwikkeling nog niet afgelopen. Zeker bij nieuwe vragenlijsten is het belangrijk om tijdens de dataverzameling de vinger aan de pols te houden. Het kan gebeuren dat ondanks een uitgebreid ontwikkel- en testtraject kleine aanpassingen nodig zijn of dat tijdens de dataverzameling een aanvullende instructie voor de interviewers, helpdesk, respondenten of dataverwerking worden gemaakt om de gevolgen van de opgetreden problemen te beperken.

Een evaluatie van hoe de vragenlijst het doet in het veld kan nuttige informatie opleveren voor degenen die de data analyseren. Bovendien is het bij het herhaald gebruik van vragenlijsten, zoals dat bij het CBS meestal het geval is, nuttig om de ervaringen uit het veld te gebruiken om de vragenlijst bij een volgende revisie te verbeteren. Daarbij moet zowel gekeken worden naar het *proces van dataverzameling* als naar de resultaten daarvan: *de verzamelde data*.

Belangrijke bronnen van informatie over het functioneren van de vragenlijst zijn in de eerste plaats de respondenten en medewerkers die contact hebben met respondenten: dit zijn de interviewers, de buitendienstmedewerkers en de medewerkers van de helpdesk. De kwaliteit van de data kan onderzocht worden met analyses van bijvoorbeeld de non-respons, de wijze waarop antwoordschalen ingevuld zijn, de plausibiliteit van de antwoorden of de verdeling van de antwoorden over de antwoordcategorieën. Ingevulde papieren formulieren zijn ook nuttige bronnen: let hierbij op doorhalingen en commentaren. Tenslotte leveren de door de helpdesk geregistreerde vragen en opmerkingen van respondenten een belangrijke bron van informatie.

Zeker bij langdurig, herhaald onderzoek zoals het CBS dat uitvoert, is het belangrijk dat er een terugkoppeling plaatsvindt van de fase van het verzamelen van de gegevens naar de ontwerpfase. Meer over het gebruik van deze evaluatiemethoden is te vinden in hoofdstuk 4.

3. Vragenlijstontwerp

3.1 Inleiding

In de fase van het vragenlijstontwerp worden de teksten, structuur en vormgeving van de vragenlijst gemaakt. Dit behelst vaak lastige afwegingen tussen verschillende belangen. Zo is het bijvoorbeeld goed om de vermogenspositie van huishoudens te meten aan de hand van een aantal vragen, die de respondent herinneren aan allerlei mogelijke soorten van vermogen. Maar meer vragen betekent ook meer kosten (meer ontwerptijd, meer interviewtijd) en soms meer respondentbelasting.¹ Een te hoge respondentbelasting kan leiden tot onzorgvuldige antwoorden en zelfs non-respons.

Er bestaat geen eenduidig recept voor het maken van dit soort afwegingen en daardoor blijft vragenlijsten maken ook altijd een creatief proces waarin verschillende teams tot verschillende goede oplossingen kunnen komen. Er is lang niet altijd één beste oplossing aan te wijzen. In dit hoofdstuk geven we een overzicht van het belangrijkste gereedschap om zorgvuldig vragenlijsten te maken. Allereerst komen een aantal algemene uitgangspunten aan bod, zoals het theoretische model van het vraag-antwoord-proces en de belangrijkste bronnen van meetfouten (paragraaf 3.2). Vervolgens gaan we concreet in op het formuleren van vragen en antwoordopties (paragraaf 3.3 en 3.4). Tenslotte komen het gebruik van toelichtingen en instructies en de structuur en vormgeving van vragenlijsten aan bod (paragraaf 3.5 tot en met 3.7).

Er is veel geschreven over het ontwerpen van vragenlijsten en er komt in dit hoofdstuk dan ook veel aan bod. Als dit alles complex lijkt, onthoud dan dat de basisregel voor vragenlijstontwerp simpel is: zorg ervoor dat de respondent begrijpt wat je vraagt, in staat is een antwoord te geven en ook bereid is dat te doen (zie kader 3.1).

Kader 3.1 Basisregel voor vragenlijstontwerp

De respondent moet

- de vraag begrijpen
- in staat zijn antwoord te geven
- bereid zijn dat te doen.

¹ In een aantal gevallen kan een serie samenhangende vragen het voor de respondent juist makkelijker maken de complexe vraagstelling op een juiste wijze te beantwoorden.

3.2 Algemene uitgangspunten bij het ontwerpen van vragenlijsten

3.2.1 *Betrouwbaarheid en validiteit*

Vragenlijsten zijn meetinstrumenten. Betrouwbaarheid en validiteit zijn de maten waarmee we beschrijven hoe goed dit instrument zijn werk doet.

Betrouwbaarheid geeft aan in hoeverre je bij herhaalde metingen van hetzelfde concept hetzelfde meetresultaat krijgt: als je bij één persoon twee keer iets meet terwijl het aspect dat je meet niet is veranderd, krijg je bij een betrouwbare meting twee keer hetzelfde resultaat. Datzelfde geldt als je twee verschillende personen meet die niet van elkaar verschillen op het gemeten kenmerk. Alleen als aan die voorwaarde wordt voldaan kunnen we statistieken maken met de antwoorden op enquêtevragen.

Validiteit betekent dat ook daadwerkelijk het concept gemeten wordt dat de onderzoeker bedoeld heeft. Je kunt wel keurig heel betrouwbaar de appels tellen, maar als je dan een statistiek wil maken over peren heb je geen valide meting gedaan.

Om betrouwbare en valide metingen te kunnen doen moeten enquêtevragen dus aan twee voorwaarden voldoen 1) iedere respondent moet de vraag op dezelfde wijze opvatten 2) iedere respondent moet de vraag zo begrijpen als de onderzoeker deze bedoeld heeft.

Om dit te bereiken werken we met gestandaardiseerde enquêtes: we proberen er voor te zorgen dat iedere respondent de enquêtevragen zoveel mogelijk op dezelfde manier hoort of ziet. Het mag bijvoorbeeld niet zo zijn dat sommige respondenten vanwege de instellingen van hun beeldscherm wel de antwoordmogelijkheden van een vraag zien en anderen niet. De theoretische achtergrond van de gestandaardiseerde enquête is het stimulus-respons-model. Als je onder gelijke omstandigheden een identieke stimulus toedient, is volgens deze theorie een verschil in respons het gevolg van een werkelijk verschil. Om een voorbeeld te geven, wanneer je een holle ruimte vermoedt achter een muur, klop je met je knokkels tegen de wand tot je een hol geluid hoort. Dat werkt alleen als je met je knokkels steeds op identieke wijze op de wand klopt. Houd je je hand ineens anders, hoor je een verschillend geluid maar dan weet je niet of dat door een holle ruimte komt of door de andere manier van kloppen.

Vanuit het streven naar gestandaardiseerde enquêtevragen, zoals die bij het CBS gebruikt worden, is het bij door interviewers afgenomen vragenlijsten ook erg belangrijk dat interviewers de vragenlijst altijd op precies dezelfde manier afnemen en, in het geval dat er uitleg nodig is, altijd op precies dezelfde manier de vragen uitleggen. Ook als de interviewer vervolgvragen moet stellen om het antwoord van de respondent beter te kunnen duiden dient dit altijd op een uniforme manier gebeuren. Vragenlijst en interviewer vormen samen het ‘meetinstrument’ dat identieke stimuli moet genereren. De kwaliteit ervan wordt uiteindelijk bepaald door de interviewer: een goede vragenlijst die slecht wordt afgenomen geeft geen juiste

meting (zie Vosmer en Engelen 2008). Vanwege deze cruciale rol van de interviewer is opleiding en instructie van interviewers erg belangrijk. Zie hiervoor het deelthema ‘Interviewertraining’ binnen Organisatie veldwerk.

Overigens moet opgemerkt worden dat de keuze voor een uniforme stimulus (vraagtekst) niet altijd de beste oplossing is, zie hiervoor bijvoorbeeld Houtkoop-Steenstra (2000). Zij pleit ervoor een interview binnen de strenge marges van standaardisering toch enigszins te laten verlopen als een normaal gesprek. Dit is ook de CBS-praktijk. Het kan bijvoorbeeld vervelend zijn wanneer een respondent al gezegd heeft dat hij een zoon en een dochter heeft, toch nog twee vragen te stellen over het aantal en de kunne van eventuele kinderen. Dit druist volledig in tegen de conventies van een gesprek en laat de respondent denken dat de interviewer geen enkele interesse heeft en nauwelijks luistert. In genoemd geval kan men de interviewer toestaan deze twee vragen in te vullen zonder ze te stellen. Bij veel complexere begrippen kan dit echter niet: ook als een respondent al heeft gezegd werkloos te zijn moet de interviewer alle vervolgvragen stellen om de werkloosheid zowel volgens de Nederlandse als de definitie van de International Labour Organization te kunnen bepalen. Zie ook paragraaf 3.2.3. Op basis van routing en tekstimputaties kunnen vragen onder andere wat ‘persoonlijker’ worden gemaakt. Bijvoorbeeld door afhankelijk van eerdere informatie uit het interview vragen te stellen over “uw echtgenoot” dan wel “uw partner”. Dit dient dan wel op uniforme wijze te geschieden. Vandaar dat tekstimputaties de voorkeur verdienen boven dergelijke initiatieven van individuele interviewers. Het maakt het werk van de interviewers ook eenvoudiger.

Voor de goede orde, een uniforme stimulus betekent niet dat *alle* respondenten altijd precies dezelfde stimulus moeten krijgen, het betekent dat alle respondenten bij wie een zekere doelvariabele moet worden gemeten een gelijke stimulus dienen te krijgen. Om verschillende zaken te meten, gebruikt men verschillende stimuli: om een holte te ontdekken achter een muur klopt men met de knokkels, om te kijken of een ei gekookt is, laat men het tollen. Routing zorgt ervoor dat respondenten de juiste uniforme stimulus krijgen.

3.2.2 *Het responsproces*

Het correct beantwoorden van een vragenlijst is een taak die doorgaans voor de respondent moeilijker is dan de onderzoeker denkt. Het responsproces model, afkomstig uit de cognitieve psychologie (o.a. Tourangeau, Rips & Rasinski, 2000) onderscheidt een viertal stappen die een respondent kan doorlopen bij het beantwoorden van een vraag: het begrijpen van de vraag, het ophalen van de benodigde informatie (uit het geheugen of uit documenten), het beoordelen van de informatie en het formuleren van een antwoord. Een respondent zal lang niet altijd alle stappen bewust doorlopen. Op de vraag “Hoe oud bent u?” hebben de meeste mensen direct een antwoord klaar. Bij ingewikkelde vragen kan het nodig zijn dat de respondent sommige stappen in het responsproces moet herhalen: als bijvoorbeeld uit zijn antwoord blijkt dat hij de vraag niet goed heeft begrepen en de interviewer

een extra toelichting geeft op de vraag. In kader 3.2 worden de stappen van het proces uitgewerkt aan de hand van een voorbeeld.

Kader 3.2 Voorbeeld van een responsproces bij persoonsenquêtes

Enquêtevraag

De volgende vragen gaan over contacten met de huisarts. Het gaat hierbij zowel om bezoeken aan de huisartsenpraktijk als om visites van de huisarts aan huis en telefonische consulten. Wanneer heeft u voor het laatst contact gehad met de huisarts voor uzelf? Was dat:

- minder dan een half jaar geleden
- een half jaar tot een jaar geleden
- of langer dan een jaar geleden?

Stappen in responsproces

Begrip

De respondent twijfelt of het bellen met de receptenlijn ook moet worden geteld als een contact met de huisarts. Het betreft namelijk alleen het inspreken op een antwoordapparaat. Het uitschrijven van het recept wordt echter wel gefactureerd door de huisarts en op grond daarvan beslist de respondent dat het dus wel zal meetellen als contact.

Ophalen van benodigde informatie

De respondent zoekt in het geheugen naar de datum van het laatste contact met huisarts, maar weet dit niet meer. Hij weet wel dat het vlak voor de zomervakantie was, omdat de aanleiding voor het telefoontje naar de receptenlijn was dat hij voor de reis nog voldoende medicijnen wilde hebben. De respondent herinnert zich dat de vakantie ergens begin juli was.

Beoordelen informatie

De respondent schat op basis van de globale vakantiedatum dat het laatste contact met de huisarts ongeveer eind juni heeft plaatsgevonden. De respondent bepaalt vervolgens dat "ergens eind juni" minder dan een half jaar geleden was.

Meedelen antwoord

Respondent antwoordt: "Minder dan een half jaar geleden" .

Bij bedrijfsenquêtes spelen er naast cognitieve aspecten – vaker dan bij personen of huishoudensenquêtes - ook organisatorische en administratieve aspecten een rol. Vandaar dat daar vaak een wat aangepast model voor het responsproces wordt gebruikt, afkomstig van Willimack en Nichols (2001). Zie kader 3.3 voor een overzicht van het responsproces voor personen en bedrijven. Voorafgaand aan het cognitieve proces zoals hierboven beschreven voor personen, zijn voor bedrijfsenquêtes drie stappen toegevoegd: de opslag van informatie, het selecteren

van de berichtgever (respondent) en het door de organisatie vaststellen van de prioriteit. Deze drie stappen vormen de context waarbinnen het cognitieve proces van het invullen zich vervolgens afspeelt. Aan het eind van het responsproces is nog een stap toegevoegd, namelijk het vrijgeven van de informatie. Hiermee wordt bedoeld dat bij sommige bedrijven de berichtgevers eerst toestemming van hogerhand moeten hebben voor zij een enquêteformulier mogen insturen.

Kader 3.3 Responsproces voor personen en bedrijven

Responsproces	
Persoonsenquêtes	Bedrijfsenquêtes
<ul style="list-style-type: none"> ▪ Begrijpen vraag ▪ Ophalen van benodigde informatie ▪ Beoordelen informatie ▪ Meedelen antwoord 	<ul style="list-style-type: none"> ▪ Opslag in administratie/geheugen ▪ Selectie berichtgever ▪ Vaststellen prioriteit ▪ Begrijpen vraag ▪ Ophalen van benodigde informatie ▪ Beoordelen informatie ▪ Meedelen antwoord ▪ Formeel vrijgeven van antwoord

3.2.3 Een vragenlijst als gesprek

Een vragenlijst is een instrument waarmee de onderzoeker eigenlijk een soort gesprek voert met de respondent. Dit is een belangrijk gegeven, want dat betekent dat rekening gehouden moet worden met normen die in het dagelijks leven gelden bij een gesprek. Zo reageren we meestal op de veronderstelde betekenis van een opmerking of vraag, en niet zo zeer op de letterlijke inhoud. Het is heel normaal om op de vraag “Ga je straks mee naar de bioscoop?” te antwoorden met “Ik moet vanavond werken”. Letterlijk gezien is dat geen antwoord op de vraag, maar iedereen zal begrijpen dat je niet mee kunt gaan.

Grice (1975) onderscheidt vier principes die ten grondslag liggen aan succesvolle conversatie (zie kader 3.4):

Kader 3.4 Grice's principes van conversatie

Principe	Toelichting
Kwantiteit	Geef precies de hoeveelheid informatie die nodig is voor het doel van het gesprek.
Kwaliteit	Spreek de waarheid, zeg niets waarvoor je geen onderbouwing hebt.
Relevantie	Spreek over zaken die relevant zijn voor het doel en onderwerp van het gesprek.
Wijze	Maak de bijdrage aan het gesprek kort en duidelijk.

Respondenten gaan ervan uit dat de onderzoeker zich bij het opstellen van de vragenlijst ook aan deze principes houdt (zie o.a. Maynard et al., 2002). Respondenten verwachten dus dat de vragen die aan de orde komen zo zijn geformuleerd dat ze makkelijk te begrijpen en relevant zijn. Afwijkingen worden als storend ervaren: als iemand eerst antwoordt dat hij in vaste dienst is en vervolgens de vraag krijgt of hij als oproepkracht werkt, dan wekt dat irritatie op.

Respondenten zullen zich doorgaans ook volgens bovenstaande regels gedragen. Volgens het principe van de kwantiteit zullen zij daarom geneigd zijn om informatie die overbodig lijkt of al eerder in het vraaggesprek is vermeld niet nogmaals te vermelden. Indien een algemene vraag over tevredenheid wordt vooraf gegaan door een vraag over de tevredenheid met een specifiek onderwerp (werk, huwelijk etc.), dan wordt het antwoord op de algemene vraag beïnvloed door de eerdere vraag: de vraag over algemene tevredenheid wordt geïnterpreteerd als 'algemeen minus het specifiek genoemde aspect' (Sudman en Bradburn, 1983, 143-144). Dit soort mechanismen speelt ook bij het rapporteren van "harde" financiële informatie. Wanneer bijvoorbeeld in een bedrijfsenquête eerst de kosten voor uitzendkrachten worden gevraagd en vervolgens de arbeidskosten, zullen de meeste respondenten de arbeidskosten exclusief kosten voor uitzendkrachten rapporteren.

3.2.4 Vier soorten taal in vragenlijsten

Respondenten gebruiken - vaak onbewust - alle informatie op de (schriftelijke of web) vragenlijst om hun taken uit te voeren. Er kunnen vier soorten taal worden onderscheiden op vragenlijsten: verbale, numerieke, symbolische en grafische taal (Redline & Dillman, 2002; zie kader 3.5).

Er zijn diverse experimenten waaruit blijkt dat kleine wijzigingen in één van deze vormen van taal grote effecten kunnen hebben op het antwoordgedrag van respondenten (o.a. Dillman & Redline 2004). De verschillende soorten taal moeten daarom op elkaar worden afgestemd en dienen samen houvast te geven bij het invullen van de vragenlijst. De uiteindelijke tekst en de opmaak van de vragenlijst moeten daarom altijd in samenhang worden ontwikkeld.

Kader 3.5 Vier soorten taal in vragenlijsten

Soort taal	Voorbeelden
Verbaal	Tekst van bijvoorbeeld kopjes, vragen, toelichtingen.
Numeriek	Nummering van bijvoorbeeld blokken, vragen, antwoordcategorieën en schaalpunten.
Symbolisch	Opteltekens, pijltjes die verwijzen naar een volgende vraag, symbool waarop geklikt kan worden voor hulptekst.
Grafisch	Lettergrootte, plaats op de pagina, afstand tussen vragen, kleurgebruik.

Het is belangrijk binnen één vragenlijst steeds consistent te zijn in alle vormen van taalgebruik. Zo dienen toelichtingen bijvoorbeeld steeds op dezelfde manier te worden aangegeven en voor hetzelfde concept (bijvoorbeeld ‘personeel op de loonlijst’) moeten steeds dezelfde woorden worden gebruikt.

3.2.5 Belangrijke bronnen van meetfouten

Er is veel onderzoek gedaan naar meetfouten die ontstaan bij het verzamelen van gegevens met gestandaardiseerde enquêtes. De belangrijkste bronnen van meetfouten worden hieronder kort op een rij gezet. Meer informatie hierover is onder andere te vinden in Tourangeau, Rips en Rasinski (2000) en Biemer en Lyberg (2003).

- ‘Satisficing’: het suboptimaal doorlopen van het vraag-antwoord proces

Respondenten hebben soms geen zin of tijd om veel moeite te steken in het correct begrijpen en beantwoorden van een enquêtevraag. Ze doorlopen dan niet alle stappen van het responsproces of doen dit niet zorgvuldig genoeg. Bijvoorbeeld door bij een vraag die eigenlijk een berekening vergt, snel een gok te doen of te kiezen voor ‘weet niet’. Dit verschijnsel wordt in de literatuur “satisficing” genoemd (Krosnick, 1991).

- Sociaal wenselijk antwoorden

We spreken van sociaal wenselijk antwoorden wanneer de respondenten hun ware antwoord op een vraag aanpassen aan hetgeen zij denken dat sociaal wenselijk is. Dit kan leiden tot onderrapportage van gedrag (bijvoorbeeld gebruik van alcohol) maar ook tot overrapportage (bijvoorbeeld scheiden van afval).

Uit de literatuur blijkt dat vertekening door sociaal wenselijk antwoorden een rol speelt bij tal van onderwerpen, ook bij zaken die op het eerste gezicht niet zo gevoelig lijken, zoals ervaren gezondheidsproblemen of tevredenheid met de buurt. Het blijkt voor onderzoekers vaak lastig vooraf te bepalen in hoeverre items gevoelig zijn voor sociaal wenselijk antwoorden. Vragen naar de huidige situatie

lijken gevoeliger voor vertekening dan vragen naar het verleden (Fowler, Roman & Di, 1998).

- ‘Acquiescence’

Respondenten blijken eerder geneigd om in te stemmen met een uitspraak, onafhankelijk van de inhoud ervan dan het er niet mee eens te zijn. Deze neiging wordt in de literatuur ‘acquiescence’ (i.e. instemming) genoemd. Bij het formuleren van vragen kan hiermee rekening worden gehouden, door deze gebalanceerd te formuleren. Een voorbeeld hiervan is: Bent u het eens of oneens met... .

- Geheugeneffecten

Bij vragen die een beroep doen op het lange-termijngeheugen van de respondent kunnen geheugeneffecten optreden. Niet alle informatie wordt even goed opgeslagen in het lange-termijngeheugen en informatie die is opgeslagen is niet altijd even gemakkelijk te vinden.

Het belangrijkste geheugeneffect is uiteraard het vergeten van zaken, hetgeen leidt tot onderrapportage. Een ander bekend geheugeneffect ontstaat door ‘forward’ of ‘backward’ *telescoping*: respondenten plaatsen ten onrechte bepaalde gebeurtenissen binnen of buiten de referentieperiode waarop de vraag betrekking heeft.

Tourangeau, Rips en Rasinski (2000) hebben de bevindingen uit onderzoek over geheugeneffecten en de implicaties voor vragenlijstontwerp op een rij gezet:

- Gebeurtenissen die langer geleden gebeurd zijn, zijn moeilijker te herinneren. Beperk daarom de referentieperiode voor zover dit mogelijk is vanwege de (lage) frequentie van de waar te nemen gebeurtenissen. Voor ziekenhuisopnames lijkt een referentieperiode van een jaar mogelijk, voor huisartscontacten is veertien dagen aan te bevelen.
- Gebeurtenissen rondom belangrijke mijlpalen in iemands leven zijn makkelijker te dateren. Door gebruik te maken van deze persoonlijke mijlpalen kan de herinnering gestimuleerd worden, waardoor de gebeurtenissen correct in de tijd geplaatst worden. Om op deze wijze goed te dateren wordt eerst door interviewer en respondent samen een soort kalender met belangrijke data samengesteld. Daarop worden vervolgens de gebeurtenissen geplaatst die de enquête meet. Dit kan uiteraard alleen bij door interviewers afgenomen vragenlijsten. Voor webvragenlijsten lijkt nader onderzoek nodig.
- Belangrijke, emotionele gebeurtenissen zijn makkelijker te herinneren dan minder belangrijke. Gebruik daarom kortere referentieperioden voor gebeurtenissen die weinig belang hebben voor de respondent. Dit voor zover de frequentie van de onderzochte gebeurtenissen dit toelaat. Zie ook het eerste punt.
- Het gebruik van diverse geheugensteuntjes vergemakkelijkt de herinnerings-taak. Gebruik daarvoor meer typen geheugensteuntjes (en bij voorkeur in deze volgorde): vraag naar het ‘wat’ (type gebeurtenis), naar ‘wie’ of ‘waar’ (de betrokken personen of de locatie) en als laatste naar het ‘wanneer’ (de

tijdsperiode). Deel de herinneringstaak op in deeltaken. Vraag bijvoorbeeld bij frequent gedrag zoals “drinken” niet naar totalen voor een hele dag maar naar consumpties per verschillende dagdelen.

- Als er meer tijd wordt besteed aan de taak, herinnert men zich meer informatie. Geef de respondent meer tijd door een langere vraagintroductie en het verlagen van het interviewtempo (langzamer spreken en meer tijd tot nadenken gunnen). Het hogere tempo van telefonische interviews maakt deze modus minder geschikt voor moeilijke retrospectieve vraagstellingen.

- Context-effecten

Identieke enquêtevragen kunnen in een andere context anders worden beantwoord. Dit kan bijvoorbeeld komen door de volgorde van de vraag in een vragenlijst of door de vormgeving, de vorm van een antwoordmogelijkheid (waarbij bijvoorbeeld het aantal posities van een open vraag suggereert of een lang of kort antwoord wordt verwacht) of de titel van een (schriftelijke) vragenlijst of de tekst van de aanschrijfbrief. Bij het vergelijken van resultaten uit verschillende enquêtes is het dus belangrijk om precies te kijken naar de context waarin een vraag is gesteld. Bovendien dient men bij onderzoek waar trends in kaart gebracht worden zo min mogelijk te wijzigen aan de vragenlijst: een verandering in de vraagvolgorde kan immers al een andere respons opleveren, waardoor de vergelijkbaarheid afneemt. Een belangrijke context is uiteraard de modus waarin een vragenlijst wordt aangeboden. Vult de respondent zelf de vragenlijst in op de computer of wordt deze door een interviewer thuis aan de keukentafel afgenomen? Met name bij vragen die cognitief belastend zijn (en dus veel motivatie vergen) en bij vragen die gevoelig zijn voor sociaal wenselijk antwoorden treden modus-effecten op.

- Interviewereffecten

De interviewer kan onbedoeld een grote invloed hebben op hoe de respondent een vraag interpreteert en beantwoordt. Dit kan gebeuren als de interviewer de vragenlijst niet correct afneemt, bijvoorbeeld omdat hij/zij de vraag- en antwoordteksten niet correct opleest of de respondent stuurt naar een bepaald antwoord. Soms kunnen ook alleen al de eigenschappen van een interviewer van invloed zijn op de antwoorden van een respondent. Zo kan het voorkomen dat bij interviewers van middelbare leeftijd die zelf angstig zijn eerder negatief wordt gerapporteerd over de veiligheid in de buurt. Indien een vraag niet volstrekt eenduidig is maar voor meer interpretaties in aanmerking komt, neemt de kans op interviewereffecten toe. De interviewer moet dan immers zelf een oplossing zoeken. Ditzelfde gebeurt indien de vragenlijst niet met alle mogelijkheden en situaties rekening houdt. Dan moet de interviewer zelf wat bedenken om verder te kunnen. Ook als de vraag vaak extra toelichting nodig heeft die niet in de vragenlijst zelf in de vorm van een interviewerinstructie is opgenomen, neemt de kans toe dat de interviewer een eigen interpretatie kiest. Dit soort zaken kunnen niet bij voorbaat bij de instructie worden opgelost. Als ze bekend waren geweest, had men immers wel van te voren het probleem opgelost.

3.2.6 Speciale aandachtspunten voor bedrijfsenquêtes

- Contactpersonen wisselen

De bedrijfsenquêtes van het CBS worden vaak met een zekere regelmaat (bijvoorbeeld ieder jaar of kwartaal) toegezonden aan een vaste contactpersoon binnen een organisatie. Deze contactpersonen kunnen echter wisselen of wel eens op vakantie gaan. De opvolger of vervanger van een contactpersoon heeft niet automatisch de kennis van zijn voorganger. De vervanger kan concepten anders interpreteren. Afspraken met contactpersonen dienen daarom altijd te worden vastgelegd. Alle informatie die nodig is om een vragenlijst correct te kunnen invullen moet in principe meegestuurd worden met een vragenlijst of in ieder geval beschikbaar zijn als de contactpersoon wisselt.

- De vragenlijst gaat naar meer berichtgevers binnen een organisatie

Bij bedrijfsenquêtes komt het vaak voor dat niet één berichtgever alle vragen kan beantwoorden. Soms worden stukken van de vragenlijst gekopieerd en doorgezonden naar andere afdelingen. Vragen over personeel worden bijvoorbeeld beantwoord door Personeelszaken, vragen over bedrijfsvoering door de financiële administratie et cetera. Toelichtingen die op een apart blad staan worden dan niet altijd meegestuurd. Dit is een reden om toelichtingen op vragen altijd zo dicht mogelijk bij de betreffende vraag te plaatsen.

- Men gebruikt bij het invullen voorgaande opgaven

Veel bedrijven houden een archief bij met kopieën van alle ingevulde CBS-vragenlijsten. Bij het invullen van een vragenlijst wordt vaak gekeken hoe men de vragenlijst de vorige keer heeft ingevuld. Sommige bedrijven maken koppelingen in de administratie om snel gegevens voor een regelmatig terugkerende vragenlijst te kunnen genereren. Voor deze bedrijven is het vervelend als een vragenlijst plotseling wijzigt. Frequente wijzigingen moeten worden voorkomen en als er wijzigingen zijn dan moeten die expliciet worden aangegeven.

3.3 Ontwerpen van enquêtevragen

In deze paragraaf gaan we in op het formuleren van vragen en geven we aan waar je op moet letten om tot goede vragen te komen. Het leidend principe bij het opstellen van een vragenlijst is dat het vraaggesprek inderdaad een gesprek is, voor zover de noodzakelijke standaardisatie dit toelaat. Zie ook paragraaf 3.2.3. Om zo'n gesprek prettig te laten verlopen dient de vragenlijst respondentvriendelijk te zijn. Belangrijke aspecten van deze respondentvriendelijkheid zijn enerzijds dat je alleen zaken vraagt die de respondent weet en anderzijds dat je zaken vraagt die voor de respondent relevant zijn en waarvan de respondent het belang inziet. Uiteraard dient een vragenlijst zo kort mogelijk te zijn en op efficiënte wijze de benodigde informatie te verzamelen. Bij het formuleren van de vragen dient men met deze aspecten rekening te houden. In deze paragraaf bouwen we voort op de meer theoretische overwegingen uit de paragrafen 3.2.3, 3.2.4 en 3.2.5. Zie ook Vosmer en Engelen (2008).

3.3.1 Vraag alleen naar zaken die de respondent weet

Vraag alleen naar zaken die de respondent redelijkerwijs kan beantwoorden. Gegevens die niet gemakkelijk te vinden zijn in het hoofd of de administratie van de respondent kunnen niet zomaar met een enquête worden gemeten. Dit soort informatie kan mogelijk wel met een logboek of dagboek speciaal voor het onderzoek worden bijgehouden. Wees extra alert op de beantwoordbaarheid van vragen bij:

- Proxi-vragen. Dit zijn vragen waarbij de respondent antwoordt voor iemand anders, de onderzoekspersoon (OP);
- Vragen die betrekking hebben op het verre verleden (bijvoorbeeld welke inentingen iemand allemaal gehad heeft);
- Vragen met een groot detailniveau (bijvoorbeeld bij de enquête Wegvervoer een specificatie van de factuurwaarde van alle verzonden goederen);
- Vragen die een bepaalde deskundigheid vereisen (bijvoorbeeld als gevraagd wordt naar de bekwaamheid van de eigen huisarts);
- Bij het vragen naar motieven en oorzaken moet je er rekening mee houden dat respondenten meestal niet goed kunnen inschatten of aangeven waarom zij iets wel of niet doen. Je kunt beter vragen of bepaalde omstandigheden of motieven al dan niet aanwezig zijn;
- Hypothetische vragen (“wat zou u doen als”). Dit soort vragen levert doorgaans weinig betrouwbare informatie op.

3.3.2 Definieer eenheid en referentieperiode

Zorg ervoor dat voor de vragenlijst in zijn geheel én per vraag duidelijk is over welke eenheid en periode de respondent moet rapporteren. Voor het beantwoorden van vragen is het immers van belang dat men duidelijk voor ogen heeft wat precies het onderwerp is. Gaat het enquêteformulier over het gehele concern of slechts over een enkele bedrijfseenheid? Heeft de vragenlijst betrekking op de persoon zelf of het hele huishouden? Gaat de vraag over de laatst genoemde werkgever of de belangrijkste? Bedoelt het CBS ‘doorgaans’, ‘op dit moment’ of ‘afgelopen jaar’? Iedere vraag waarvan verwacht kan worden dat het antwoord over de tijd kan variëren dient voorzien te zijn van een referentieperiode (Fowler & Cosenza, 2008; zie kader 3.6).

Kader 3.6 Voorbeeld referentieperiode

Niet :

“Hoe vaak voelt u zich neerslachtig” ?

Wel:

“Hoe vaak heeft u zich in de afgelopen vier weken neerslachtig gevoeld? Was dat altijd, meestal, soms of nooit? “

De referentieperiode van een vraag dient zorgvuldig gekozen te worden. Deze kan immers zowel de interpretatie van de vraag beïnvloeden als het beantwoorden bemoeilijken. Als je bijvoorbeeld vraagt naar hoe vaak iemand boos is geweest in de afgelopen week zullen mildere vormen van boosheid worden gerapporteerd dan wanneer over het afgelopen jaar wordt gevraagd. Immers, lichtere vormen van ergernis of boosheid zullen sneller worden vergeten dan heftige driftbuien. Zie ook paragraaf 3.2.5. Vragen naar hetzelfde gedrag over verschillende referentieperioden tellen daarom soms niet op. Fowler & Cosenza geven de volgende tips voor het kiezen van referentieperiodes:

- Kies de referentieperiode in relatie tot het belang van de gebeurtenis waarnaar gevraagd wordt. Hoe onbelangrijker de gebeurtenis voor de respondent, hoe korter de referentieperiode moet zijn.
- Deel geheugentaken op in verschillende deelstappen. Vraag bijvoorbeeld niet in één keer hoeveel contacten met artsen iemand heeft gehad in het afgelopen jaar, maar vraag specifiek naar verschillende soorten artsen.

3.3.3 Stel slechts één vraag per vraag

Als er meer dan één vraag tegelijk wordt gesteld is niet duidelijk op welke vraag het antwoord betrekking heeft. Een mogelijke aanwijzing voor de aanwezigheid van dubbele vragen zijn woordjes als ‘omdat’, ‘vanwege’, ‘of’ en ‘en’(zie kader 3.7).

Kader 3.7 Voorbeeld dubbele vraagstelling

Niet:

“Let u op het energieverbruik vanwege de kosten?”

Wel:

“Let u op het energieverbruik?”

Indien ja:

“Doet u dat vanwege de kosten?”

3.3.4 *Formuleer vragen gebalanceerd*

Respondenten zijn vaak geneigd om instemmend te antwoorden (“acquiescence” genoemd in de literatuur). Door vragen zo gebalanceerd mogelijk te stellen wordt deze vorm van vertekening gereduceerd. Vraag dus niet “Bent u het eens met het verhogen van de pensioengerechtigde leeftijd van 65 naar 67 jaar?” maar vraag “Bent u het eens of oneens met ...”. Zie ook paragraaf 3.2.5 voor de bespreking van deze meetfout.

3.3.5 *Gebruik geen verkeerde aannames in de vraagstelling*

Een bekende bron van problemen bij enquêtevragen is dat bij het ontwerp wordt uitgegaan van verkeerde aannames over de kennis of het gedrag van de respondent.

Als zonder nadere toelichting wordt gevraagd naar de mening over een bepaald feit (“de nieuwe ziektekostenregeling”) wordt verondersteld dat de respondent hiermee bekend is. Uit verschillende experimenten blijkt dat veel respondenten ook een antwoord geven op vragen waar zij geen kennis over hebben. Hetzij omdat ze niet willen toegeven dat ze ergens nog nooit van gehoord hebben, hetzij omdat ze de vraag op een manier interpreteren waardoor ze wel tot een voor hen zinvol antwoord kunnen komen. Zie kader 3.8 voor een voorbeeld hiervan.

Krosnick et al. (2002) concluderen op basis van diverse studies waarbij respondenten een mening werd gevraagd over fictieve issues (bijvoorbeeld een niet bestaande regeling) dat circa 30% van de respondenten een antwoord geeft op dit soort vragen terwijl ze die feitelijk niet zouden kunnen beantwoorden. Indien te verwachten is dat veel respondenten niet precies weten waar een vraag over gaat kan gekozen worden voor een uitgebreide introductie op het onderwerp, voor een filtervraag (“Bent u bekend met ...”) of voor het expliciet aanbieden van een optie ‘weet niet / geen mening’.

Kader 3.8 Voorbeeld van interpretatie van een onduidelijk begrip

In een CBS experiment (Vis-Visschers, Arends-Tóth, Giesen en Meertens, 2008) naar het gebruik van “weet niet” en toelichtingen bij internet vragenlijsten werden met opzet vraagstellingen met moeilijke woorden gebruikt. Eén van de vragen was: “Eet u of één van uw gezinsleden wel eens producten die retinol bevatten?”

Voor uitleg over het begrip retinol moest men een toelichting aanklikken. Lang niet alle respondenten maakten daar gebruik van.

Zo antwoordde één van de respondenten op deze vraag dat hij nooit retinol eet. Toen hem werd gevraagd om een toelichting op dat antwoord bleek dat hij (ten onrechte) veronderstelde dat retinol iets ongezonds was en dat zijn vrouw er wel op zou letten dat hij dat niet zou eten.

Naast een verkeerde inschatting van het kennisniveau zijn enquêtevragen soms gebaseerd op verkeerde aannames over het gedrag van respondenten. Als bijvoorbeeld aan ouders van jonge kinderen wordt gevraagd hoe vaak hun kind een fietshelm draagt als het op de fiets wordt vervoerd, wordt al verondersteld dat het kind wel eens op de fiets zit. Als dat niet het geval is en de antwoordcategorie “niet van toepassing” ontbreekt, kunnen respondenten hun antwoord niet kwijt.

3.3.6 Kies eenduidige woorden; geef waar nodig definities

Voor valide en betrouwbare metingen dienen de gebruikte begrippen eenduidig te zijn. Zelfs schijnbaar eenvoudige begrippen kunnen heel uiteenlopend worden opgevat. Zo kan de term ‘afgelopen week’ worden geïnterpreteerd als ‘vorige week maandag tot en met afgelopen zondag’ of ‘de afgelopen maandag tot afgelopen vrijdag’ of, als de vraag op donderdag gesteld wordt: ‘afgelopen donderdag tot en met afgelopen woensdag’. Alle begrippen die abstract of ambigu zijn dienen daarom gedefinieerd te worden. Respondenten gaan er in dit soort gevallen vaak vanuit dat hun eigen interpretatie de juiste is en informeren daarom niet naar wat wordt bedoeld.

Definities en toelichtingen dienen bij voorkeur voorafgaand aan de vraag gegeven te worden of te worden geïntegreerd in de vraagtekst. Dit geeft de respondenten direct de nodige informatie bij het interpreteren van de vraag en voorkomt dat zij de interviewer onderbreken of stoppen met lezen zodra de vraag gesteld is. Om voortijdig onderbreken te voorkomen dient er nooit middenin een vraagtekst een vraagteken gebruikt te worden.

Soms wordt ervan afgezien een definitie aan de respondent voor te schrijven. Dit kan zijn omdat de CBS-definitie te ver afdrikt van die van respondenten of omdat respondenten zeer verschillend denken over het begrip. Te denken valt aan de definitie van part-time werken. Hierover denken respondenten zeer verschillend en de CBS-definitie (minder dan dertig uur) wordt zeker niet door iedereen gedeeld. In dit geval is het beter te vragen naar het aantal uren dat men werkt en dan zelf bepalen of men dit als part-time of full-time beschouwt.

Ook kan het een expliciet onderzoeksdoel zijn na te gaan wat respondenten onder een bepaald begrip verstaan. Dan wordt uiteraard geen definitie gegeven.

3.3.7 Gebruik eenvoudige taal

Enquêtes van het CBS zijn altijd bedoeld voor een brede groep mensen. Daarom is het belangrijk eenvoudige taal te gebruiken. Het advies is om eenvoudig Nederlands op taalniveau B1 aan te houden (zie Bijlage 1 voor een beschrijving van de verschillende taalniveaus). Taalniveau B1 kan bijna iedereen begrijpen, ook mensen die geen hoge opleiding hebben gehad en voor hun werk nooit hoeven te lezen. Hieronder volgen de belangrijkste tips om eenvoudige taal te schrijven. Meer informatie over het schrijven van eenvoudige taal is onder andere te vinden in de Schrijfwijzer begrijpelijke formulieren (Bureau Taal, 2006) of op de website van de Gemeente Den Haag. Op deze website staan veel schrijftips en is zelfs een

woordenboek opgenomen met moeilijke woorden en alternatieven daarvoor (“Het Helder Haags Woordenboek”).

- Maak korte vragen en zinnen.

Korte vragen kunnen beter worden begrepen en onthouden dan lange vragen. Kort is echter niet altijd beter, soms is voor een goed begrip een toelichting nodig of geeft een langere inleiding op een vraag respondenten juist de gelegenheid te focussen op het gewenste onderwerp of periode. Te lange inleidingen of toelichtingen schieten echter hun doel voorbij omdat de respondent niet alles kan onthouden. Probeer in ieder geval de zinnen binnen een vraagtekst zo kort mogelijk te houden. Vijfentwintig woorden per vraag is een richtlijn (Vosmer en Engelen 2008).

- Gebruik veel voorkomende en concrete woorden.

Jargon, figuurlijk taalgebruik en abstracte termen zijn moeilijk te begrijpen voor veel respondenten. Zowel voor de bedrijfsenquêtes als persoonsenquêtes van het CBS geldt dat de doelgroep vaak zeer heterogeen is. Het taalgebruik moet ook voor de deelpopulatie met een lage taalkennis begrijpelijk zijn.

- Teksten met een dubbele ontkenning zijn vaak moeilijk te begrijpen.

“Ik ben het er niet mee eens dat je dat niet doet” is al een lastige tekst. Dubbele ontkenningen in enquêtevragen ontstaan vaak bij stellingen waarbij respondenten moeten aangeven in welke mate ze het ermee eens of oneens zijn. Bij negatief geformuleerde stellingen ontstaat dan een dubbele ontkenning. Want wat betekent “oneens” als antwoord op de stelling “Ik ben niet tevreden over mijn werk?” Respondenten die wel tevreden zijn met hun werk moeten hier *oneens* antwoorden. Hierbij kan men zich makkelijk vergissen.

- Formuleer vraagteksten en instructies positief.

Positief geformuleerde teksten (“gebruik een pen”) zijn makkelijker te begrijpen dan negatief geformuleerde teksten (“gebruik geen potlood”). Pas op met het gebruik van woorden als “niet” of “exclusief” - deze worden vaak gemist door respondenten.

3.3.8 *Maak duidelijk wat het gewenste antwoordformaat is*

Maak zoveel mogelijk al in de vraag duidelijk welk soort antwoord gewenst is. Dit maakt de responstaak voor de respondent makkelijker. Dit geldt vooral voor interviewgestuurde vragenlijsten waarbij de respondent de antwoordmogelijkheden meestal niet ziet (zie kader 3.9).

Kader 3.9 Maak in de vraag het antwoordformaat duidelijk

Niet:
“Wanneer bent u hier gaan wonen?”
Wel:
“In welk jaar bent u hier gaan wonen?”.

3.3.9 Vragenlijstontwerp voor internationaal gebruik

Vaak willen onderzoekers de uitkomsten van onderzoek uit verschillende landen met elkaar vergelijken. Dit stelt bijzondere eisen aan het vragenlijstontwerp: men wil dan immers in ieder land zo veel mogelijk hetzelfde meten. Er zijn verschillende procedures ontwikkeld om concepten en vraagstellingen in internationaal verband te harmoniseren. Zie voor een uitgebreide behandeling daarvan hoofdstuk 4 in Brancato et al. (2006).

3.4 Ontwerpen van antwoordmogelijkheden

Bij het ontwerpen van antwoordmogelijkheden die na het stellen van een enquêtevraag aan de respondent worden voorgelegd spelen verschillende soorten overwegingen een rol. De belangrijkste zijn:

- **Gewenst meetniveau**

De soort analyses die kunnen worden uitgevoerd op de data hangen af van het meetniveau van de variabelen. Het hangt af van het gewenste meetniveau of je leeftijd of inkomen bijvoorbeeld meet in klassen of in jaren en euro's. Keuzes met betrekking tot de antwoordcategorieën dienen dus in overleg met de beoogde gebruikers van de data te worden gemaakt.

- **Kosten voor verwerking en dataverzameling**

Gesloten vragen zijn makkelijker te verwerken dan open vragen. Gesloten vragen zijn vragen met vaste antwoordmogelijkheden en open vragen zijn vragen waarbij de respondent zelf het antwoord formuleert. Het coderen van antwoorden op open vragen is tijdrovend en maar in beperkte mate te automatiseren. Ook het door respondenten of interviewers laten noteren van antwoorden op open vragen is tijdrovend.

- **Gemak van beantwoorden**

De keuze van het soort antwoordmogelijkheden is van invloed op hoe moeilijk het is een vraag te beantwoorden. Het is makkelijker globaal aan te geven in welke van vijf inkomensklassen je valt dan tot op de euro nauwkeurig een jaarinkomen op te geven.

- Kwaliteit van het antwoord

De respondent gebruikt de antwoordmogelijkheden bij het interpreteren van het doel en de betekenis van een vraag en als een overzicht van wat volgens de onderzoekers blijkbaar de verwachte verdeling in de populatie is. Respondenten kunnen hun antwoord aanpassen aan de antwoordschaal omdat zij een sociaal wenselijk antwoord willen geven. Maar het kan ook zijn dat zij de schaal gebruiken om gemakkelijk tot een antwoord te kunnen komen: het middelste punt in een schaal wordt vaak opgevat als een gemiddeld antwoord. Als iemand denkt dat hij een bepaalde activiteit iets vaker zal doen dan gemiddeld, is het makkelijker om snel een antwoordoptie boven het midden te selecteren dan eerst precies het antwoord te bedenken en deze dan te matchen met de antwoordschaal. Zie kader 3.10 voor een voorbeeld van het effect van antwoordschalen.

Kader 3.10 Effect van verschillende antwoordschalen

Hoeveel uur per dag kijkt u televisie?												
Optie 1	$\leq \frac{1}{2}$	$\frac{1}{2} - 1$	1-1½	1½-2	2-2½	$> 2\frac{1}{2}$						
Optie 2	$\leq 2\frac{1}{2}$	2½-3	3-3½	3½-4	4-4½	$> 4\frac{1}{2}$						
Optie 3	Open vraag											
<p>% respondenten dat meer dan 2,5 uur per dag TV kijkt:</p> <table border="1"> <tbody> <tr> <td>Optie 1</td> <td>22%</td> </tr> <tr> <td>Optie 2</td> <td>54%</td> </tr> <tr> <td>Optie 3</td> <td>52%</td> </tr> </tbody> </table>							Optie 1	22%	Optie 2	54%	Optie 3	52%
Optie 1	22%											
Optie 2	54%											
Optie 3	52%											
<p><i>Bron: Toepoel, Vis, Das en Van Soest, 2009.</i></p>												

In tabel 3.1 staat een overzicht van veel voorkomende antwoordmogelijkheden en tips over het gebruik daarvan. Dit overzicht is gedeeltelijk gebaseerd op Jansen et al. (1989). Het valt buiten het bestek van deze Methodenreeks om gedetailleerd in te gaan op alle afwegingen bij het ontwikkelen van antwoordcategorieën. Een goed overzicht hiervan is te vinden in hoofdstuk 3 van Brancato et al. (2006) en in Billiet (2006). Deze laatste is Nederlandstalig en bevat veel voorbeelden.

Tabel 3.1. Soorten antwoordmogelijkheden

Antwoordmogelijkheid	Wanneer gebruiken	Tips
<p><i>Open</i> De invuller moet gegevens zelf invullen met woorden, codes of cijfers</p>	<ul style="list-style-type: none"> - als het aantal mogelijke antwoorden groot is of niet bekend is welke antwoorden zijn te verwachten; - mits het de invuller duidelijk is wat voor soort antwoord wordt verwacht. 	<p>Lok een exact en bondig antwoord uit:</p> <ul style="list-style-type: none"> - stel de vraag zo concreet mogelijk; - begin de vraag met een vraagwoord (bijv. hoe, wie, welke); - maak antwoordvakken niet onnodig groot.
<p><i>Half-open</i> Een aantal vaste antwoordmogelijkheden is voorgegeven, maar een antwoordmogelijkheid 'anders, namelijk.....' maakt het mogelijk een niet opgesomd antwoord te geven.</p>	<ul style="list-style-type: none"> - Als de meest voor de hand liggende antwoordmogelijkheden bekend zijn maar men niet zeker is of dit alle relevante mogelijkheden zijn 	<p>Gebruik deze mogelijkheid bij nieuwe bij nieuwe vraagstellingen. Als de vraag een tijd in het veld geweest is, kan dan worden nagegaan of sommige frequent voorkomende antwoorden als aparte antwoordmogelijkheid opgenomen moeten worden. De minder vaak voorkomende kunnen dan in de gesloten categorie 'overig' worden genoteerd.</p>
<p><i>Enkelvoudige keuze</i> De invuller moet kiezen uit twee alternatieven, bijvoorbeeld ja of nee, groter of kleiner etc.</p>	<ul style="list-style-type: none"> - als van belang is of één bepaalde categorie van toepassing is. 	<ul style="list-style-type: none"> - Ga na of de vraag altijd relevant is of dat een categorie 'niet van toepassing' moet worden toegevoegd.
<p><i>Meerkeuzevragen</i> De invuller moet ofwel één antwoord kiezen uit diverse mogelijkheden of kan een aantal antwoorden aankruisen.</p>	<ul style="list-style-type: none"> - als van belang is welke van een serie categorieën van toepassing is; - als het aantal mogelijke antwoorden beperkt is; - nb: indien meer zaken uit een lijst van toepassing kunnen zijn moet overwogen worden deze stuk voor stuk te vragen als een ja/nee vraag. Dit is meer werk voor de interviewer en respondent, maar levert doorgaans een hogere (en betere) rapportage op van de onderzochte aspecten. 	<ul style="list-style-type: none"> - zorg voor een uitputtende lijst; - zorg ervoor dat de antwoordmogelijkheden elkaar niet overlappen; - maak duidelijk of één of meer antwoordmogelijkheden mogen worden aangekruist; - als antwoordmogelijkheid X een bijzonder geval van Y is, plaats X dan vóór Y; let op dat bij dit soort gevallen de antwoordcategorieën zo geformuleerd zijn dat zij elkaar uitsluiten - Als gekozen moet worden uit een lange lijst plaats deze dan in een voor de respondent zinvolle volgorde (bijv. alfabetisch, geografisch); - Als er geen andere overwegingen zijn, plaats dan de meest voorkomende antwoordmogelijkheid bovenaan; - Laat bij voorkeur aankruisen wat wél van toepassing is (en dus niet doorhalen wat niet van toepassing is) - Laat liever geen codes invullen - Lange lijsten zijn moeilijk te verwerken, gebruik indien mogelijk subkopjes om lange lijsten te ordenen.
<p><i>Antwoordschalen</i></p>	<p>Worden vaak gebruikt als respondent mening over iets moet geven.</p>	<ul style="list-style-type: none"> - Doorgaans worden 4 tot 7 schaalpunten gebruikt; aantal schaalpunten hangt af van analysewensen; - invulgemak neemt af met toenemend aantal schaalpunten; - evenveel negatieve als positieve schaalpunten opnemen in de schaal; - bij schalen tot 7 punten alle schaalpunten labelen met woorden - neutrale optie in schaal midden opnemen; - zo veel mogelijk dezelfde schalen gebruiken binnen één blok en liefst ook binnen één vragenlijst; - lange batterijen schaalvragen zijn vervelend om in te vullen en leiden tot suboptimaal invulgedrag; - gebruik antwoordschalen liever niet voor indirecte beoordelingstaken (zie onder).

We willen hier wel kort ingaan op het gebruik van antwoordschalen voor indirecte beoordelingstaken. Fowler & Cosenza (2008) onderscheiden directe en indirecte

beoordelingstaken bij het gebruik van antwoordschalen en raden het gebruik van indirecte beoordelingstaken sterk af. Er is sprake van een directe taak als de respondenten gevraagd wordt hun oordeel over iets (bijvoorbeeld de eigen gezondheid) te plaatsen op een continuüm dat betrekking heeft op dit aspect (bijvoorbeeld variërend van zeer slecht tot zeer goed). Er is sprake van een indirecte beoordelingstaak als respondenten een stelling wordt voorgelegd (bijvoorbeeld: mijn gezondheid is zeer goed) en zij dan op een continuüm moet aangeven in hoeverre zij het eens zijn met deze stelling. Indirecte beoordelingstaken zijn veel complexer en dus gevoeliger voor meetfouten.

3.5 Gebruik van toelichtingen en instructies

Toelichtingen en instructies zijn belangrijke hulpmiddelen om de respondent duidelijk te maken wat de vraag inhoudt en hoe hij deze dient te beantwoorden. Zeker bij persoonsenquêtes kan men zich afvragen of een vraag wel een goede vraag is indien deze bijzonder veel toelichting en uitleg nodig heeft. Bij bedrijfsenquêtes is dat minder het geval omdat daar vaak nauwkeurige specificaties worden gevraagd die beantwoording niet bemoeilijken. Helaas worden toelichtingen bij zelf in te vullen vragenlijsten vaak slecht gelezen. Dit is vervelend, want als sommige invullers de toelichting wel lezen en andere niet, neemt de vergelijkbaarheid van de antwoorden af. In het algemeen geldt dat de tekst van vragen het beste wordt gelezen. Essentiële informatie die bij alle respondenten bekend moet zijn, dient dus bij voorkeur in de vraagtekst te worden opgenomen. De belangrijkste richtlijnen voor het gebruik van toelichtingen en instructies zijn:

- Overwegingen ten aanzien van de plaats van verklarende uitleg over wat er met een vraag bedoeld wordt of over de betekenis van bepaalde termen:
 - Eerste keus: in de vraagtekst zelf, voorafgaand aan de eigenlijke vraag;
 - Tweede keus: kort en bondig in aparte toelichting, nabij vraag en/of invulvak;
 - Derde keus: in een uitgebreide toelichting onderaan de pagina of aanklikbaar vanuit de vraag als instructie voor de interviewer;
 - Vierde keus: als apart document. Als de verklarende toelichting in een los document staat, dan moet er in de vragenlijst bij de afzonderlijke vragen duidelijk naar verwezen worden.
- Plaats algemene achtergrondinformatie over het onderzoek in een begeleidende brief of in aparte invultips. Ook een folder of een web site zijn mogelijk.
- Invulinstructies die duidelijk maken hoe het antwoord op de vragen genoteerd moet worden (bijvoorbeeld afronden in duizendtallen), moeten worden geïntegreerd in de vraag, direct boven of naast het invulvakje. Indien die instructie voor het merendeel der vragen geldt, dan eventueel aan het begin.
- Houd toelichtingen zo kort mogelijk.
- Gebruik geen lange regels in vragenlijsten die respondenten zelf in moeten vullen.

- Plaats bij CAWI-vragenlijsten foutmeldingen daar waar ze optreden (en liever niet ergens aan het eind van het formulier) en op zo'n manier dat de originele vragen nog zichtbaar zijn voor de respondent. Formuleer deze vriendelijk en helder zodat de respondent begrijpt wat hij 'fout' heeft gedaan.

3.6 Richtlijnen voor de structuur van de vragenlijst

Een vragenlijst goed afnemen of invullen is een complexe taak. Door deze op te delen in kleine, overzichtelijke deeltaken wordt het werk voor respondenten en interviewers een stuk makkelijker. Die deeltaken dienen te worden aangeboden in een logische volgorde en - indien nodig - te worden onderverdeeld in thematische blokken. Voor interviewers is het bijvoorbeeld belangrijk te weten waar ze precies in de vragenlijst zijn. Bij de vroeger gebruikelijke papieren vragenlijsten was dat eenvoudig. Om dit bij computergestuurde vragenlijsten te vergemakkelijken toont het onderste frame van het scherm de voorgaande vragen met gegeven antwoorden en enkele komende vragen.

In het volgende gaat het vooral om door respondenten zelf in te vullen vragenlijsten. Een goede structuur biedt houvast (waar gaat het over, hoe is het formulier opgebouwd, wie moet de informatie verstrekken) en een houvast tijdens het invullen (hoe verhoudt een specifieke vraag zich tot de overige vragen). Dit kan bij computergestuurde vragenlijsten bijvoorbeeld door een index te tonen of verschillende tabbladen te maken.

Het is belangrijk dat de structuur en opmaak van de vragenlijst zo is ontworpen dat de respondent steeds weer gemakkelijk zijn weg kan vinden in de vragenlijst en ziet waar hij is gebleven. Bij een CAWI vragenlijst kan dit bijvoorbeeld door bij het opnieuw openen het scherm te tonen met de laatste vraag die ingevuld is. Bij zelf in te vullen vragenlijsten geldt met name voor bedrijfsenquêtes dat dit typisch iets is wat men tussen de bedrijven door doet en soms even onderbreekt. In die situatie zijn dergelijke hulpmiddelen des te belangrijker.

3.6.1 Naamgeving vragenlijsten en thematische blokken

Geef iedere vragenlijst een duidelijke naam. Door plaats, opmaak en grootte van de naam wordt duidelijk dat dit overkoepelende informatie voor de hele vragenlijst betreft.

- Houd de naam zo kort mogelijk.
- Kies namen die zinvol zijn voor de respondent.
- Bij vragenlijsten die periodiek worden uitgestuurd aan dezelfde respondenten dient de naamgeving ook een vermelding te bevatten van de periode waarop de vragenlijst betrekking heeft.
- Kies voor de titel van een vragenlijst een unieke naam.
- Gebruik bij zelf in te vullen vragenlijsten bij voorkeur ook nummering voor vragen en blokken.

- Een vragenlijstnaam of de naam van een specifiek blok van vragen wordt vaak alleen bij het eerste scannen van de vragenlijst gelezen en bij het daadwerkelijk invullen van de vragen overgeslagen. De vragen moeten dus ook zonder het lezen van de vragenbloknamen correct ingevuld kunnen worden.

Kader 3.11 Voorbeeld van interpretatie van naamgeving vragenlijst

In een concept vragenlijst werd de titel “*Kwartaalenquête Bedrijfstelling*” gebruikt. Dit is voor het CBS een enquête die 4 keer per jaar wordt uitgezonden. Maar voor elke enquête worden nieuwe respondenten geënquêteerd. In dit geval is de toevoeging Kwartaalenquête niet zinvol voor de respondent en roept soms zelfs vragen op: “moet ik dit dan ieder kwartaal invullen als ik nu meedoe?”.

3.6.2 Volgorde van onderwerpen

- De eerste indruk zet de toon. Begin daarom met een gemakkelijk en aansprekend onderwerp, bij voorkeur dus niet met het uitvragen van allerlei achtergrondkenmerken (al is dit soms niet te vermijden als die informatie nodig is voor de routing van een vragenlijst).
- Plaats vragen die bij elkaar horen bij elkaar. Probeer de respondent zo min mogelijk heen en weer te laten springen tussen verschillende onderwerpen, referentieperiodes of antwoordmogelijkheden.
- Plaats gevoelige onderwerpen op het eind van de vragenlijst, de respondent heeft dan al enige band opgebouwd met de interviewer en het onderzoek. Mochten de gevoelige onderwerpen tot afbreken van het interview leiden, dan is in ieder geval een groot deel van de vragen beantwoord.
- Plaats moeilijke vragen die veel inspanning en motivatie vragen niet te laat in de vragenlijst wanneer vermoeidheid kan hebben toegeslagen. Anderzijds is aangetoond dat erg moeilijke vragen leiden tot afbreken. Bij web vragenlijsten zijn duidelijke pieken in ‘drop-out’ te zien bij zulke vragen.
- De onderwerpen in een vragenlijst kunnen elkaar beïnvloeden. Houd hier zoveel mogelijk rekening mee bij de volgorde van de vragen. Vraag bijvoorbeeld niet eerst of iemand werkloos is en direct daarna een algemene mening over de economie.

3.7 Vormgeving van vragenlijsten

Een goede vormgeving van de vragenlijst is zowel van belang voor het gemak waarmee de vragenlijst kan worden ingevuld of afgenomen als voor de kwaliteit van

de antwoorden. Over de vormgeving van vragenlijsten is veel geschreven. Deze paragraaf kan slechts een korte indruk daarvan geven en is grotendeels gebaseerd op hoofdstuk 3.3 uit het *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System* (Brancato et al., 2006). Andere nuttige bronnen met meer informatie over vormgeving van vragenlijsten zijn Dillman, Smyth & Christian (2009) en Couper (2008).

Mensen reageren op drie niveaus op vormgeving: emotioneel, functioneel en reflectief. Reacties op het *emotionele niveau* zijn de eerste, impulsieve gevoelens die worden opgeroepen door hoe de vragenlijst er uitziet. Een dikke vragenlijst, volgedrukt met kleine letters maakt een andere indruk dan een dunne vragenlijst met een verzorgde en leesbare vormgeving. Een prettig ogende vormgeving van een vragenlijst is geen overbodige luxe waarop bezuinigd kan worden. Vragenlijsten die direct al door hun uiterlijk een negatief gevoel oproepen zullen minder vaak en minder goed worden ingevuld. Kenmerken van een prettige vormgeving zijn symmetrische, ronde vormen, terugkerende patronen en voldoende witruimtes.

Op het *functionele niveau* dient de vormgeving respondenten en interviewers te ondersteunen bij hun cognitieve taak. Zo is het nuttig als deeltaken ook visueel duidelijk onderscheiden worden, en als duidelijk is welke antwoordvakjes bij welke vraag horen (duidelijke regelgeleiding) en als lange opsommingen met antwoordmogelijkheden in overzichtelijke brokjes worden ingedeeld. Hierbij is het van belang zoveel mogelijk aan te sluiten bij bestaande praktijken. Enkele bekende aanbevelingen voor goed functioneel ontwerp zijn:

- Mensen lezen in onze cultuur van links boven naar rechts onder: de vragenlijst moet dus links bovenaan beginnen.
- Zaken die bij elkaar horen (vragen binnen een thematisch blok, antwoordmogelijkheden bij een vraag) moeten visueel ook duidelijk bij elkaar gegroepeerd zijn.
- Informatie die niet van belang is voor de interviewer of respondent (maar bijvoorbeeld alleen voor interne verwerking) moet visueel minder opvallen. Deze dus in een kleiner lettertype en bij voorkeur onderaan de pagina plaatsen.
- Het moet visueel direct duidelijk zijn waar de antwoorden moeten worden ingevuld. Dit kan bijvoorbeeld door een licht gekleurde achtergrond te gebruiken in combinatie met witte antwoordvakjes.
- Vormgeving moet steeds functioneel en consequent worden toegepast: gebruik voor dezelfde onderdelen (zoals titels van blokken, antwoordopties of toelichtingen) consequent dezelfde lettergroottes, lettertypes, kleur en grafische symbolen.
- Kleuren mogen alleen als aanvullende opmaak worden gebruikt, nooit als enige vorm van onderscheid. Niet iedereen kan alle kleuren namelijk goed zien en bovendien weet je nooit hoe de vragenlijst getoond wordt bij de respondent of interviewer (denk aan kopiëren van een papieren vragenlijst of de scherminstellingen waardoor kleuren anders over komen). Pas vooral op met de kleuren rood en groen, die worden door veel mensen slecht onderscheiden.

Op het *reflectieve niveau* denkt de respondent bewust na over de responstaak, en kan bijvoorbeeld besluiten wel of niet mee te werken aan het onderzoek of al dan niet waarheidsgetrouw te antwoorden. Voor dit niveau is het van belang dat de vormgeving van de vragenlijst duidelijk uitstraalt dat het onderzoek wordt uitgevoerd door een serieus bureau. Denk daarbij aan het gebruik van het CBS-logo en een sobere, maar verzorgde vormgeving.

4. Methoden voor het ontwikkelen, testen en evalueren van vragenlijsten

4.1 Inleiding

Het is bijna niet mogelijk om in één keer van achter het bureau een goede vragenlijst te ontwikkelen. Met een aantal goed opgezette pre-tests kunnen veel problemen bij het verzamelen en analyseren van de data worden voorkomen. Ook in pre-tests kunnen nooit alle mogelijke problemen met vragenlijsten worden ontdekt en opgelost. Daarom is het zinvol vragenlijsten die in het veld zijn geweest te evalueren: hiermee kan ad hoc worden bijgestuurd, kan nuttige aanvullende informatie worden verkregen voor de analyse van de data en kan de vragenlijst in de toekomst mogelijk worden bijgesteld.

In hoofdstuk 2 is de vragenlijstontwikkeling als proces geschetst en zijn de mogelijke methoden voor het testen en evalueren van vragenlijsten kort weergegeven. In dit hoofdstuk wordt een aantal van de eerder genoemde methoden uitgebreid beschreven. Het betreft hier methoden die binnen het CBS en ook bij andere Statistiekbureaus regelmatig gebruikt worden en als ‘best practices’ of ‘recommended practices’ beschouwd kunnen worden (Brancato et al. 2006, hoofdstuk 6). In dit hoofdstuk behandelen we achtereenvolgens informele tests, expertraadpleging, focusgroepen, cognitieve interviews, vignetten, observaties, technische testen, experimenten en de analyse van proces- en surveydata. Er wordt eerst een schematisch overzicht gegeven van welke testen in welke fasen van het proces van toepassing zijn (zie paragraaf 4.2). Vervolgens worden de genoemde methoden uitvoerig beschreven: het doel van de methode, de fase(n) waarin deze gebruikt wordt en de voor- en nadelen komen aan bod. De belangrijkste bronnen die we bij het opstellen van dit hoofdstuk gebruikt hebben zijn: Brancato et al. (2006), Jansen et al. (1989), Snijkers (2002) en De Leeuw, Hox en Dillman (2008).

De gekozen methode hangt af van het doel van de test en in de praktijk ook van de beschikbare tijd en middelen. Een minimumvereiste van Eurostat is dat een vragenlijst in ieder geval één keer getest moet worden met potentiële respondenten of berichtgevers (Brancato et al., 2006, pagina XI). Het is efficiënt om de methoden in volgorde van de kosten in te zetten: het is bijvoorbeeld zonde om een vragenlijst met evidente fouten al aan proefpersonen voor te leggen.

4.2 Welke methode voor welke fase?

In onderstaande tabel 4.1 staat een schematisch overzicht van de methoden die kunnen worden toegepast in één of meer fasen van het vragenlijstontwikkelproces. Per methode wordt aangegeven in welke fase(n) ze gebruikt kunnen worden.

Tabel 4.1. Overzicht van methoden per fase van vragenlijstontwikkeling

Methoden \ Fasen	Conceptua- lisering & onderzoeks -ontwerp	Vragenlijst- ontwerp	Pre-testen	Herzien	Data- verzameling	Proces- monitoring & evaluatie
Review van documenten	X	X	X			
Informe test	X	X	X	X	X	X
Expert- raadpleging	X	X		X		X
Focusgroep	X	X	X			X
Cognitieve interviews	X	X	X			X
Vignetten	X	X	X			
Observaties			X		X	X
Pilot			X			
Technische testen			X			
Experiment			X		X	
Analyse van procesdata					X	X
Analyse van surveydata					X	X

4.3 Informele tests

Korte beschrijving: Informele tests zijn testmethoden waarmee de vragenlijstmaker zelf de vragenlijst evalueert.

Doel: Om snel een idee te krijgen van de kwaliteit van vraagstellingen of om aanwijzingen te vinden waarop vervolgtests zich kunnen richten is het mogelijk om ‘quick and dirty’-tests te doen. Deze methoden zijn erg effectief voor het snel traceren van moeilijkheden in het vraag-antwoordproces, onduidelijkheid van gebruikte begrippen en termen, en/of de lengte van vraagteksten of vragenlijsten. Hiermee kan men ook vrij eenvoudig slordigheden in formuleringen, moeilijk lopende zinnen, moeilijk taalgebruik, typefouten en layout-technische dingen aan het licht brengen. Deze methode kan ook toegepast worden als er niet veel tijd is voor een grootschalige test of als er weinig tijd is om een test voor te bereiden.

Fase: Deze methode is feitelijk toe te passen in elke fase van de vragenlijstontwikkeling, maar wordt vooral gebruikt in een vroeg stadium van conceptualisering en vragenlijstontwerp.

Uitgebreide beschrijving: Een van de eerste dingen die een vragenlijstmaker zou moeten doen is zelf de vraag en antwoordmogelijkheden hardop voorlezen (Campanelli, 2008). Een vraag die er op papier goed uitziet kan toch heel vreemd klinken. Daarnaast is het zinvol zelf te proberen de vraag serieus te beantwoorden, vanuit de eigen situatie of vanuit een paar denkbeeldige situaties (bijvoorbeeld: “hoe

zou een zelfstandige deze vraag invullen?”, “wat moet iemand die een opleiding heeft afgebroken hier antwoorden?”). Deze methode lijkt triviaal. Als echter alle vragenlijstontwerpers één keer serieus hun eigen vragenlijst voor zichzelf zouden invullen zouden een hoop vervelende vragenlijstfouten tijdig worden opgespoord. Verder is het nuttig om familie, vrienden en collega’s in te schakelen als een soort proefrespondent: verschillende mensen hebben andere perspectieven en vinden ook weer andere fouten (zie kader 4.1). Het ligt immers niet voor de hand dat een vragenlijstontwerper zelf nog ziet dat hij jargon gebruikt in een vraagstelling. Ook het kritisch doorlopen van de eigen vragenlijst aan de hand van een checklist kan gezien worden als een informele testmethode. Er bestaan verschillende checklists voor goede vragenlijsten (zie Bijlage 2 voor een voorbeeld van een checklist).

Kader 4.1 Voorbeeld van een ‘quick and dirty test’

Om al vroeg in het testtraject te achterhalen of de vraag over contacten met de huisarts potentieel problematisch is, is de vraag voorgelegd aan een vriendin.

Vraag: Hoe vaak heeft u gedurende de afgelopen 4 weken, dus sinds {4 weken geleden}, contact gehad voor uzelf met een praktijkondersteuner en/of gespecialiseerd verpleegkundige in de praktijk van uw huisarts?

Reactie: “Ik zou het eigenlijk niet goed weten, wat is een praktijkondersteuner? Is dat de doktersassistente? Dan zou ik zeggen ja, want zij heeft mij het herhaalrecept uitgeschreven.”

Voor- en nadelen: Informele testen zijn relatief snel te organiseren. De bevindingen zijn indicatief voor de problemen die een vragenlijst kan gaan opleveren zelfs als er weinig tijd is om te testen. Het belangrijkste nadeel is uiteraard dat er geen “echte” respondenten of berichtgevers worden geraadpleegd en alleen het perspectief van de vragenlijstmaker en diens directe omgeving wordt betrokken bij de test.

4.4 Expertraadpleging

Korte beschrijving: Een expertraadpleging is een vorm van testen waarbij deskundigen worden geraadpleegd om feedback te geven op een (concept)versie van de vragenlijst.

Doel: Een expertraadpleging heeft als doel de conceptvragenlijst te screenen op potentiële problemen voor interviewers en respondenten. Daarnaast kunnen ook individuele vragen, de operationalisatie van begrippen (vertaling van begrippen naar concrete vraagstellingen) en het design worden gecheckt.

Fase: Een expertraadpleging wordt voornamelijk in de fasen van conceptualisatie en vragenlijstontwerp gehouden. Bevindingen leveren vaak aandachtspunten die als input kunnen dienen voor de testfase in het proces van vragenlijstontwikkeling.

Uitgebreide omschrijving: Er zijn verschillende soorten experts die geraadpleegd kunnen worden en de keuze daarvan hangt uiteraard af van de vragen die leven bij de vragenlijstontwerpers. Als er onzekerheid is of de te meten concepten correct zijn geoperationaliseerd kan het zinvol zijn om beoogde gebruikers van de data te raadplegen. Als er vragen zijn over de begrijpelijkheid van de vraagstelling en de structuur van de vragenlijst is het nuttig om de mening van vragenlijstexperts te vragen. Te denken valt aan de volgende soorten experts:

- Inhoudsdeskundigen / statistici
- Vragenlijstmethodologen
- Vormgevers
- Beoogde verwerkers van de gegevens (zowel technisch als inhoudelijk)
- Helpdeskmedewerkers
- Interviewers en buitendienstmedewerkers

Bij expertraadpleging is er bijvoorbeeld aandacht voor de gebruikte termen en concepten in de vragen, de structuur van de vragen, de antwoordcategorieën, de volgorde van de vragen, de navigatiestructuur van de vragenlijst, interview- en codeerinstructies et cetera.

Er zijn verschillende manieren om de expertraadpleging te organiseren. Een prettige methode is de experts vooraf de vragenlijst systematisch te laten beoordelen en vervolgens met hen in een groeps gesprek (zie paragraaf 4.5) of een persoonlijk gesprek deze beoordeling door te spreken. Een voorbeeld van een zeer gedetailleerd beoordelingsschema is opgenomen in Bijlage 3. Een eenvoudiger manier is de zogenaamde plus-min methode; Daarbij wordt de expert gevraagd bij iedere (eventueel voorafgaand gemarkeerde) vraag een plus of een min te zetten als er daar iets positief of negatief opvalt. Tijdens het gesprek kunnen deze plussen en minnen nader worden toegelicht.

Voor- en nadelen. Een expertraadpleging is in vergelijking met onderzoek met respondenten relatief snel te implementeren, zeker als de experts interne medewerkers zijn. Experts zijn doorgaans goed in staat met zinvolle oplossingen voor vastgestelde problemen te komen. Bij het maken van vragenlijsten moeten vaak compromissen tussen verschillende belangen gesloten worden. Voor het draagvlak van deze compromissen kan het ook zinvol zijn als alle betrokken partijen zoveel mogelijk zijn geraadpleegd bij de totstandkoming van de vragenlijst. Een expertraadpleging is echter slechts zo goed als de geraadpleegde experts. En ook zeer deskundige experts hebben geen volledig inzicht in de problemen en mogelijkheden van respondenten. Voor een voorbeeld van een expertraadpleging, zie kader 4.2.

Kader 4.2 Voorbeeld van een expertraadpleging

In het kader van de evaluatie van de vragenlijsten voor de Productiestatistiek (PS) in 2004 is een groot aantal activiteiten uitgevoerd die tezamen een goed beeld moeten geven van verbeterpunten in de PS-vragenlijsten. Eén van die activiteiten betrof het raadplegen van CBS-medewerkers, die vanuit hun functie (Buitendienst consultants, Helpdesk medewerker, Gaafmakers en Analisten) ervaring hebben opgedaan met de werking van de PS-vragenlijsten. Op hun gebied zijn zij experts.

In focusgroepen zijn de problemen met de PS-vragenlijsten door CBS-medewerkers geïnventariseerd. De gesprekken bestonden steeds uit 3 onderdelen:

1. Een algemeen deel, met algemene problemen van berichtgevers m.b.t. CBS-vragenlijsten;
2. Algemene problemen m.b.t. de PS-vragenlijsten;
3. Een specifiek deel, met problemen gerelateerd aan de specifieke rubrieken in de PS-vragenlijsten.

De focusgroepen hebben een groot aantal opmerkingen opgeleverd, met een conclusie dat de vragenlijsten grondig herzien moesten worden. De opmerkingen hebben richting gegeven aan hoe de vragenlijsten aangepast konden worden.

4.5 Focusgroepen

Korte omschrijving: Focusgroepen zijn gestructureerde groepsgesprekken, met doorgaans tussen de 5 en 7 personen, onder leiding van een gespreksleider (ook wel moderator genoemd).

Doel: Focusgroepen worden gebruikt om kwalitatieve informatie te krijgen over hoe een bepaalde doelgroep denkt over vooraf gedefinieerde onderwerpen, bijvoorbeeld het onderwerp van de vragenlijst, de beoogde modus van waarneming, en/of de gehanteerde concepten en definities in de vraagstellingen.

Fase: Focusgroepen kunnen zowel tijdens het ontwerpen van de vragenlijst als bij de evaluatie van een geïmplementeerde vragenlijst worden toegepast. Een focusgroep is een handzame methode om een vragenlijst te evalueren met bijvoorbeeld een groep interviewers of helpdeskmedewerkers.

Uitgebreide beschrijving: Focusgroepen zijn met name nuttig om inzicht te krijgen in ervaringen, meningen en houdingen. Het voordeel van een groepsgesprek ten opzichte van een één-op-één gesprek is tweeledig. In de eerste plaats is het efficiënter om zes mensen tegelijk te spreken dan achter elkaar. In de tweede plaats kunnen deelnemers door de groepsdynamiek gestimuleerd worden hun meningen en ervaringen meer uitgesproken en gedetailleerd onder woorden te brengen.

Onderwerpen die in een focusgroep met proefrespondenten aan de orde kunnen komen zijn bijvoorbeeld de gevoeligheid van vragen, de interpretatie van vragen, het woordgebruik en in hoeverre het gemakkelijk of moeilijk is de vragen te beantwoorden. Voor een voorbeeld, zie kader 4.3.

Een succesvolle focusgroep vereist een gedegen voorbereiding. Er dient een ‘topic list’ of ‘topic guide’ te zijn – een lijst met aandachtspunten van onderwerpen die in de discussie aan de orde moeten komen. Deze is essentieel voor de richting van de discussie en de gespreksleider dient steeds te bewaken dat de discussie ten dienste staat van de onderwerpen die in de topic list zijn genoemd. De topic list kan ontwikkeld worden op basis van de resultaten van een eerder gehouden review van bestaande documenten of expertraadpleging. Aanbevolen wordt een aantal groepsgesprekken van 5 à 7 deelnemers te organiseren om een zekere spreiding van ervaringen en inzichten te garanderen. Deelnemers moeten niet te veel van elkaar verschillen omdat de discussie dan te weinig gestructureerd verloopt. De rol van gespreksleider is in de eerste plaats aan de deelnemers duidelijk te maken wat er van hen wordt verwacht (in ieder geval hun inzichten en ervaringen over de gevraagde onderwerpen te delen en respectvol om te gaan met de andere groepsleden), ten tweede de topic list bewaken, en tenslotte is het belangrijk om dominante sprekers in de hand te houden en stille mensen een kans te geven toch iets te zeggen.

De duur van een focusgroep is ongeveer 1 tot 1 ½ uur. Het is zeer aan te bevelen om tijdens de focusgroep een observant/notulist aanwezig te hebben die de gespreksleider kan ondersteunen. Bovendien is het nuttig om ten behoeve van de analyse van alles wat gezegd is een video-opname te maken (bij een audio-opname is het heel moeilijk te bepalen wie wat heeft gezegd). De videobeelden kunnen ook gebruikt worden als illustratie bij de rapportage van de resultaten van de focusgroepen. Meer informatie over focusgroepen is te vinden in Krueger (1994).

Een speciale manier van groepsgesprekken organiseren is de interviewerevaluatie zoals die op het CBS wordt gedaan (zie de beschrijving hiervan in Bijlage 4). Hierbij komt een groep van circa 15 interviewers bij elkaar. De interviewers gaan dan in tweetallen uiteen en formuleren stellingen over de vragenlijst die wordt geëvalueerd. De stellingen kunnen gaan over klachten en onduidelijkheden die zij bij de uitvoering van het veldwerk zijn tegengekomen of juist zaken die positief opvielen. Nadat de tweetallen klaar zijn met de inventarisatie worden de genoteerde stellingen opgehangen. De stellingen kunnen wel worden toegelicht, maar niet bediscussieerd. Het resultaat van de bespreking is een lijst met eenduidige stellingen. Aan het eind van deze fase wordt er gekeken of er onderwerpen ontbreken die vooraf van CBS-, of klantzijde zijn ingebracht. Stellingen die worden gemist worden alsnog toegevoegd aan de totale lijst en desgewenst toegelicht. Tenslotte wordt alle interviewers gevraagd op een scoreformulier een cijfer toe te kennen aan elke stelling. Door de interviewers te laten scoren is het mogelijk om de knelpunten gewogen weer te geven. Bij deze methode is er dus weinig ruimte voor discussie en het uitdiepen van ervaringen, maar worden wel snel en efficiënt ervaringen en meningen geïnventariseerd.

Kader 4.3 Voorbeeld uit topiclist van de focusgroep Gezinsmonitor

Deze focusgroep met proefrespondenten volgde op het vertonen van een videoband met een voorbeeldinterview.

Discussie over de videoband. Belangrijk: de discussie richt zich niet op de mening van het paar, maar op het individu!! Als mevrouw iets zegt over 'ons' of praat over 'wij', dan aan meneer vragen wat hij daarvan vindt en omgekeerd.

- Wat vindt u van het afzonderlijk interviewen van gezinsleden?
- Wat vindt u van de aanwezigheid of afwezigheid van de ander terwijl de één geïnterviewd wordt?
- Wat heeft uw voorkeur: de aanwezigheid of de afwezigheid van de ander?
- Is dit bij u thuis haalbaar wat betreft de indeling van het huis
- En wat vindt u van de confrontatie met elkaars antwoorden?
- Zou u zelf meedoen aan een onderzoek waarvoor u en uw partner afzonderlijk geïnterviewd worden?
- En zou u zelf meedoen aan een confrontatie?
- Denkt u dat het haalbaar is op een dergelijke manier onderzoek te doen bij een groot aantal Nederlanders?

Voor- en nadelen: Groepsgesprekken zijn een efficiënte manier om verschillende meningen en ervaringen tegelijk in kaart te brengen. Door de groepsdynamiek kunnen deelnemers gestimuleerd worden hun ideeën duidelijker te formuleren. De sociale interactie kan deelnemers echter ook remmen om een afwijkend geluid te laten horen of überhaupt aan het woord te komen. Focusgroepen met respondenten over conceptvragenlijsten blijven altijd theoretische beschouwingen, tenzij ze gecombineerd worden met het serieus invullen of beantwoorden van de vragenlijst. Vaak blijken problemen bij vragen pas wanneer ze serieus worden beantwoord. Tot slot moet opgemerkt worden dat er twijfels bestaan over de waarde van focusgroepen over hypothetisch gedrag ("wat zou u doen als..."). Focusgroepen zijn wel nuttig om te *begrijpen* welke overwegingen mensen hebben. Maar de voorspellende waarde van discussies over toekomstig gedrag is doorgaans gering.

4.6 Cognitieve interviews

Korte omschrijving: Cognitieve interviews zijn open interviews met als doel inzicht te krijgen in de cognitieve processen die in werking treden bij het beantwoorden van enquêtevragen.

Doel: Het doel van cognitieve interviews is het vergaren van kwalitatieve informatie over hoe vragen door respondenten worden begrepen en beantwoord. Verondersteld

wordt dat de verbale uitingen van proefrespondenten een directe presentatie zijn van het specifieke cognitieve proces dat plaats vindt naar aanleiding van de vragen in de vragenlijst. Het idee is dat het begrijpen van achterliggende cognitieve processen van respondenten een bijdrage kan leveren aan het opzetten van een vragenlijst, de keuze van antwoordcategorieën, vraagvolgordes, visueel ontwerp et cetera.

Fase: Cognitieve interviews kunnen worden gehouden zodra een conceptvraagstelling af is, in de testfase, maar worden ook gebruikt om vragenlijsten die al in het veld zijn te evalueren.

Uitgebreide beschrijving: In cognitieve interviews wordt respondenten gevraagd hun gedachtegang te beschrijven bij het beantwoorden van enquêtevragen. Door het invulgedrag van de proefrespondent te observeren en te bespreken komt men erachter wat de problemen met de vraagstellingen en antwoordmogelijkheden van het formulier zijn. We beschrijven nu eerst de technieken die gebruikt worden bij cognitieve interviews en daarna gaan we in op de praktisch kant van de organisatie van deze interviews.

- Technieken om inzicht te krijgen in cognitieve processen

We behandelen hier twee veelgebruikte technieken om inzicht te krijgen in de cognitieve processen die een respondent doorloopt bij het beantwoorden van enquêtevragen: ‘hardop denken’ en het stellen van metavragen. Ook vignetten kunnen gebruikt worden om inzicht te krijgen in cognitieve processen. Deze methode is echter ook voor andere doeleinden geschikt en wordt daarom apart behandeld in paragraaf 4.7. Zie voor een uitgebreide behandeling van cognitief interviewen Willis (2005).

Hardop denken

Bij de hardop-denktechniek wordt de proefrespondent gevraagd hardop te denken en alle gedachten uit te spreken die boven komen om een antwoord op de betreffende vraag te geven. Hiermee kan men moeilijkheden identificeren met betrekking tot het begrijpen van de vraag, de responstaak, het ophalen van herinneringen en feiten uit het verleden, reacties op gevoelige vragen etc (zie kader 4.4). Voordeel van deze methode is dat de onderzoeker zelf niet stuurt en er dus ook onverwachte problemen naar boven kunnen komen. Nadeel van deze methode is dat het een behoorlijke verbale vaardigheid vraagt van de respondent en het hardop denken juist de cognitieve processen kan verstoren. Hierdoor kan de respondent worden afgeleid van de taak of deze juist veel bewuster gaan vervullen. Dit kan artefacten veroorzaken bij het testen. Enerzijds kunnen problemen ontstaan door het testen die bij een reguliere afname van het interview niet zouden optreden en anderzijds kunnen problemen verborgen blijven, omdat de respondent bij een test veel meer zijn best doet op de vragenlijst dan normaal gesproken het geval zou zijn.

Kader 4.4 Voorbeeld van hardop denken uit de voorbereidingen op het Onderzoek Relatie- en Gezinsvorming onder Jonge Allochtonen in 2004

Reactie op de vraag: "Hoe vaak gaat u in het algemeen naar de kerk of naar een godsdienstige bijeenkomst?":

"Hoe krijg je hier goede antwoorden op? Ik ben moslim dus ik ga niet naar de kerk. Dat is een rare vraag. Maar ik ga wel eens naar de kerk uit nieuwsgierigheid om te zien waarom mijn christelijke vrienden daar naartoe gaan. Ik wil dat weten. Maar hoe moet ik die vraag nu beantwoorden? Moet ik mijn kerkbezoek of mijn moskeebezoek opgeven?"

Metavragen

Metavragen zijn aanvullende vragen die worden ingevoegd tussen de enquêtevragen (zie kader 4.5). Het doel van metavragen is inzicht te verschaffen in de denkprocessen van respondenten. Voorbeelden van metavragen zijn "waar dacht u aan bij het beantwoorden van deze vraag?" of "U heeft ja geantwoord, hoe bent u tot dit antwoord gekomen?". Metavragen kunnen zowel voorafgaand aan de interviews worden bedacht als spontaan worden gesteld ("U moest zuchten bij die vraag, waarom was dat?"). Speciale vormen van metavragen zijn:

Probing ofwel *doorvragen*. Met deze techniek wordt het antwoord van de proefrespondent op de enquêtevraag met behulp van aanvullende vragen verduidelijkt. Deze metavragen worden gebruikt als de informatie die een proefrespondent geeft incompleet is of als vermoed wordt dat er problemen zijn met een vraag. De interviewer kan bijvoorbeeld vragen hoe de proefrespondent tot zijn antwoord is gekomen of waarom hij heeft gekozen voor een bepaalde antwoordmogelijkheid.

Kader 4.5 Voorbeeld enquêtevraag en bijbehorende metavragen

Inleiding: De volgende vragen gaan over ziekenhuisopnamen. Het betreft alle soorten ziekenhuizen. Bezoeken aan de spoedeisende hulp afdeling/eerste hulp of poliklinische bezoeken moeten niet worden meegerekend.

Vraag: Bent u in de afgelopen 12 maanden, dus sinds {datum 1 jaar geleden} in het ziekenhuis opgenomen voor een dagopname, d.w.z. een opname waarbij u wel een bed kreeg toegewezen, maar niet hoefde te overnachten?

Metavraag: Deze vraag ging over opname in het ziekenhuis. Aan welke ziekenhuizen dacht u bij het beantwoorden van deze vraag? Wat verstaat u onder "ziekenhuisopname" en wat onder "dagopname"?

Parafraseren: bij deze techniek wordt proefrespondenten gevraagd om de vraag te herhalen in eigen woorden. Met behulp van deze metavraag kan worden achterhaald of de proefrespondent de vraag begrijpt en op de juiste manier interpreteert. Parafraseren kan ook betere woorden voor definities en concepten opleveren indien verschillende proefrespondenten consistent dezelfde terminologie gebruiken. Deze techniek is bijzonder nuttig bij complexe vraagstellingen en het achterhalen van jargon.

Voorbeelden van parafraseren zijn: “Kunt u in uw eigen woorden deze vraag herhalen?”, “Stel u moet deze vraag in uw eigen woorden aan uw buurman voorleggen, hoe zou u hem dan formuleren?”.

Bij het vooraf formuleren van de metavragen is het van belang dat er al een idee of hypothese is over mogelijke problemen met de enquêtevragen. Deze informatie kan naar voren zijn gekomen uit eerder gehouden focusgroepen of reviews. Idealiter wordt per metavraag ook een testdoel geformuleerd, zodat de onderzoeker steeds duidelijk voor ogen heeft wat de bedoeling is van een bepaalde metavraag (zie kader 4.6).

Kader 4.6 Voorbeelden van metavragen en testdoelen

Energieverbruik: Let u meer dan vroeger op het energieverbruik omdat de energieprijzen zo hoog zijn?

Testdoel:

Wat betekent het als mensen zeggen dat ze wel of niet meer letten op hun energieverbruik?

Metavragen:

Indien men ja zegt: voorbeelden vragen van hoe, sinds wanneer en waarom men dat doet;

Indien men nee zegt vragen of dat is "nee, ik let niet op mijn energieverbruik" of "nee ik let om andere redenen / altijd al op mijn energieverbruik".

Kringlooppapier Hoe vaak maakt u gebruik van kringlooppapier?

Testdoel:

In het veld komt bij deze vraag hier relatief vaak (4%) 'weet niet' voor. Wat betekent dat: dat mensen niet weten wat kringlooppapier is, of dat zij niet altijd weten of het papier dat zij gebruiken kringlooppapier is, of iets anders?

Metavragen:

Indien men een antwoord geeft: kunt u deze vraag in uw eigen woorden herhalen? Hoe heeft u deze vraag opgevat?

Indien men geen antwoord geeft: wat vindt u onduidelijk in deze vraag?

Metavragen kunnen op verschillende momenten gesteld worden: direct na een enquêtevraag, na een blok van vragen of helemaal retrospectief, na afloop van het invullen van de gehele vragenlijst. Het voordeel van het direct stellen van metavragen is dat de respondent de enquêtevraag en het gegeven antwoord nog vers in het geheugen heeft. Het nadeel ervan is dat de metavragen het normale responsproces verstoren en er zo artefacten kunnen worden gecreëerd (zie ook hardop denken). Het kan vanuit dit oogpunt nuttig zijn om bij een deel van de proefrespondenten direct metavragen te stellen en bij een ander deel retrospectief.

Ook bij metavragen geldt dat de effectiviteit van deze cognitieve techniek afhangt van de vaardigheden en het vermogen van de proefrespondent om zijn gedachtegang te verbaliseren en om bijvoorbeeld alternatieve begrippen te kiezen.

- Werving en selectie van proefrespondenten

Cognitieve interviews worden doorgaans met speciaal geworven proefrespondenten uitgevoerd. Bij de werving van de proefrespondenten moet gestreefd worden naar voldoende diversiteit, zodat zij representatief zijn voor de mogelijke problemen die kunnen optreden in een vragenlijst. Dit kan betekenen dat bijvoorbeeld onevenredig veel zelfstandigen, laagopgeleiden of zieke mensen worden geworven, wanneer juist bij deze doelgroepen problemen met de vragenlijst worden verwacht. Bij het testen van persoonsenquêtes is het raadzaam om - afgezien van onderwerpgebonden eigenschappen - te zorgen voor een spreiding op geslacht, leeftijdsgroepen, opleidingsniveaus en al dan niet werkzaam zijn. Bij het testen van bedrijfsenquêtes is spreiding op grootteklasse en branche wenselijk.

Het gewenste aantal proefrespondenten hangt af van een inschatting van de risico's op problemen met de conceptvragenlijst en het aantal verschillende routes dat getest moet worden. Doorgaans worden tussen de 20 en 30 proefrespondenten geworven.

- Rol van de cognitief interviewer

Voor interviewers van cognitieve interviews is het van belang dat zij zich neutraal opstellen. Zeker als men de eigen vragen test is de neiging vaak groot om respondenten die fouten maken te gaan corrigeren of de vraagstelling te verdedigen. Dat is niet de bedoeling, aangezien dit de uitkomsten van het cognitief interview vertekent. Bij bedrijfsenquêtes, waar er vaak een langdurige relatie bestaat tussen de respondent en het CBS kan het wel zinvol zijn om de respondent uitleg te geven over de betekenis van vragen. Dit dient dan altijd na afloop van het feitelijke cognitieve interview te gebeuren.

Bij de cognitieve interviews moet de proefrespondent goed worden geïnstrueerd over zijn taak en dient benadrukt te worden dat de vragenlijst getest wordt en niet de respondent. Respondenten vinden het vaak vervelend als blijkt dat zij een begrip niet kennen of een vraag verkeerd invullen. Indien mogelijk moeten zij er niet eens op gewezen worden wanneer zij fouten maken.

- **Uitvoering**

De cognitieve interviews worden vaak gehouden in afgesloten ruimtes waarbij van de gesprekken video- of geluidsopnamen gemaakt worden. Dit materiaal wordt gebruikt voor verdere analyses en eventueel als illustratie bij een presentatie van de testresultaten. De ervaring leert dat één minuut film van een proefrespondent die een volkomen verkeerde interpretatie van een vraag geeft overtuigender kan werken om een vraagstelling te herzien dan een nota van twintig pagina's.

Voor bedrijfsenquêtes worden cognitieve interviews meestal gehouden bij de berichtgevers/bedrijven op locatie, omdat men daar toegang heeft tot de benodigde administratie. Bij elektronische vragenlijsten kan het ook nuttig zijn om opnames te maken van het beeldscherm. (Zie ook paragraaf 4.8 Observeren van respondenten en interviewers).

Cognitieve interviews worden uitgevoerd aan de hand van protocollen. Hierin staan de doelen van de test omschreven, de vragen waarbij moeilijkheden worden verwacht, en instructies voor de interviewers. Het is raadzaam om ook gebruik te maken van een antwoord- of verslagsjabloon waarin de interviewer per vraag de relevante reacties van de proefrespondent weergeeft.

Voor- en nadelen: Cognitieve interviews leveren kwalitatieve informatie op die bij een kleine groep respondenten verzameld is. Hierdoor is de interpretatie en generalisatie van de resultaten soms moeilijk. Daar komt nog bij dat het werven van respondenten een arbeidsintensieve klus is. Verder worden de lange doorlooptijd en vereiste verbale vaardigheden van de proefrespondenten als nadelig beschouwd. Daartegenover levert elk cognitief interview rijke informatie op over het vraag-antwoordproces. Daarmee kunnen fouten tijdig worden opgespoord en fouten bij de implementatie van de vragenlijst voorkomen worden. In onze ervaring geeft geen andere testmethode zoveel inzicht in de oorzaken van meetfouten als het cognitieve interview. Wij onderschrijven dan ook de aanbeveling van Eurostat om in ieder geval iedere vragenlijst bij een aantal proefrespondenten uit de beoogde doelgroep uit te testen.

4.7 Vignetten en invulling van een vragenlijst aan de hand van een praktijkvoorbeeld

Korte omschrijving: Vignetten zijn situatieschetsen waarop proefrespondenten reageren. Op deze wijze komt aan het licht wat zij onder een begrip precies verstaan. Ook kan proefrespondenten worden gevraagd aan de hand van een voorbeeld uit de praktijk, een zogenaamde casus, een hele vragenlijst in te vullen.

a) Vignetten

Doel: In vragenlijstonderzoek wil de onderzoeker soms ingewikkelde begrippen gebruiken waarvan niet duidelijk is wat respondenten daaronder verstaan en waarvan evenmin duidelijk is of alle respondenten er hetzelfde onder verstaan. In

zo'n geval kan het voorleggen van een fors aantal vignetten aan een serie proefrespondenten uitkomst bieden.

Fase: Doorgaans in een vroeg stadium van vragenlijstontwikkeling, vaak in combinatie met een focusgroep.

Uitgebreide beschrijving: Een vignettenonderzoek gaat uit van een lastig begrip waarvan men niet weet welk scala aan betekenissen eraan worden toegekend door respondenten en/of welke emotionele ladingen het begrip heeft. Dergelijke begrippen dienen, in principe, te worden gemeden in vragenlijstonderzoek, maar soms ontkomt men er niet aan. Een voorbeeld is het begrip 'van werkring veranderen' waarnaar in de vervolgpeilingen van de EBB onderzoek wordt gedaan. Om erachter te komen wat respondenten een werkring vinden en wat een verandering daarvan, kunnen (maximaal) enkele tientallen situaties worden uitgewerkt. Deze dienen alle een realistisch voorbeeld te zijn van iets wat net wel of net niet een verandering van werkring genoemd kan worden. Deze test die indertijd is uitgevoerd met proefrespondenten, was al eens veel eerder gedaan met CBS-medewerkers. Bij beide groepen traden veel definitieverschillen op. Nadat proefpersonen deze vragen hebben beantwoord, bijvoorbeeld met 'wel' of 'geen' verandering, kan een focusgroep worden gehouden waarbij men de uitkomsten vergelijkt. Ook kunnen individuele nagesprekken worden gehouden, of kan respondenten worden gevraagd hardop te denken bij het beoordelen van de vignetten.

Voor- en nadelen: Een vignettenonderzoek biedt uitstekende mogelijkheden begrippen en de aspecten daarvan zoals die leven onder de doelpopulatie te beschrijven. Op basis daarvan kan een omschrijving worden gezocht die in de vragenlijst kan worden gebruikt, of kan besloten worden het begrip uit te splitsen in verscheidene vragen of zelfs besloten worden het überhaupt niet te gebruiken. Een nadeel is dat de kwaliteit van de uitkomsten in sterke mate afhankelijk is van de kwaliteit van de vignetten en dus van de kennis van de onderzoeker over wat binnen de doelpopulatie leeft.

b) Invulling op basis van casus

Doel: Soms is een te testen vragenlijst alleen van toepassing in een bepaalde situatie of voor een beperkte doelpopulatie. Het is niet altijd even gemakkelijk om proefrespondenten of berichtgevers te vinden die precies voldoen aan de specifieke eisen. Om toch bepaalde vragen of vragenlijsten te kunnen testen is het mogelijk om een (denkbeeldige) casus te gebruiken en deze te laten invullen door proefrespondenten die niet precies voldoen aan de gestelde eisen maar wel als een soort 'proxi' kunnen optreden. Buitendienstmedewerkers met veel ervaring in het veld zijn een voorbeeld. Daarnaast kunnen dergelijke casussen of praktijkvoorbeelden gebruikt worden om goed vergelijkbare resultaten te krijgen over het functioneren van een vragenlijst. Zo kan het zinvol zijn twee alternatieve vragenlijsten te laten invullen aan de hand van dezelfde vignetten.

Fase: Doorgaans in combinatie met focusgroepen in de testfase.

Uitgebreide beschrijving: In deze methode wordt aan proefrespondenten een korte beschrijving, casus of situatie (vignet) voorgelegd (voor een voorbeeld, zie kader 4.7). Aan de hand van die informatie wordt gevraagd om een vraag of vragenlijst te beantwoorden. Bij het invullen kan de onderzoeker vragen of de proefrespondent hardop wil denken, om zo het invulgedrag te verduidelijken. Dit kan ook in een focusgroep gebeuren, dan wordt aan de groep gevraagd om de vragenlijst in te vullen aan de hand van één of meerdere voorgeschreven situaties. Dan kunnen in de groep bepaalde beslissingen besproken worden.

Het resultaat van deze methode is een indicatie over mogelijke problemen bij op specifieke doelgroepen gerichte vragen.

Een goede casus is een realistische situatieschets die alle informatie bevat die nodig is om een invul- taak uit te voeren. Plaats de informatie bij voorkeur in een natuurlijke volgorde (en dus niet per se in de volgorde waarin de informatie wordt gevraagd in de vragenlijst). Als de informatie moet worden opgezocht in bestaande stukken (zoals verzekeringspolis of loonadministratie) neem dan de benodigde stukken op bij het vignet. Hoe realistischer de invul- (en opzoek)taak, hoe zinvoller de test.

Kader 4.7 Voorbeeld van een casus

In de workshop Begrijpelijke Formulieren door het CBS gegeven samen met het ministerie voor Binnenlandse Zaken, wordt een van de deelnemers gevraagd om aan de hand van onderstaand vignet het Formulier 'Aangifte van vestiging/vertrek/verhuizing' van de gemeente Heerlen in te vullen.

Vignet: Vul het formulier in als vader van Maurice van der Wal (19 jaar):

Hij gaat studeren en verlaat het ouderlijke huis in Heerlen om op kamers te gaan wonen in een driekamerappartement in Eindhoven samen met zijn beste vriend Tim. Hij heeft zijn vader, Jan van der Wal, gemachtigd om het formulier in te vullen.

Door het invulgedrag van deze deelnemer vervolgens te observeren en te bespreken komt men erachter wat de problemen met de vraagstellingen en antwoordmogelijkheden van het formulier zijn.

Voor- en nadelen: Een test met gebruikmaking van praktijkvoorbeelden maken het mogelijk om in weinig tijd en met weinig middelen een vragenlijst te testen bedoeld voor een specifieke doelgroep. Zo kunnen vragen waarvoor het bijna onmogelijk is proefrespondenten te vinden getest worden. Bovendien is het aantrekkelijke van deze praktijkvoorbeelden dat de input wordt gestandaardiseerd waardoor er verschillende versies van vragenlijsten systematisch kunnen worden vergeleken. Er kleven twee grote nadelen aan het gebruik van een casus-benadering. In de eerste plaats vereist deze techniek van de proefrespondent/berichtgever de nodige cognitieve vaardigheden en creativiteit om de informatie uit het gegeven voorbeeld te verwerken en de taak te begrijpen. Het tweede nadeel is dat deze respondenten als proxy optreden en derhalve nooit echt representatief zijn voor de doelpopulatie.

4.8 Observeren van respondenten en interviewers

Korte omschrijving: Bij het observeren van respondenten of interviewers worden bepaalde vormen van gedrag systematisch geregistreerd. Zo kan bijvoorbeeld bij het naluisteren van een aantal afgenomen interviews worden bijgehouden bij welke vragen problemen ontstaan zoals het niet letterlijk oplezen van de vraagteksten door de interviewer, het onderbreken van de interviewer door de respondent, of het niet direct kunnen kiezen van een antwoordmogelijkheid. Enquêtevragen waarbij dit soort gedrag vaak voorkomt moeten waarschijnlijk herzien worden.

Doel: Door middel van het observeren en daarna analyseren van het vraag-antwoordproces kan men identificeren waar respondenten en eventueel interviewers problemen ondervinden bij het stellen en/of beantwoorden van de vragen. Het is een manier om de effectiviteit van de vragenlijst te onderzoeken.

Fase: Observaties kunnen plaatsvinden nadat er een conceptvragenlijst af is, doorgaans na de fase van vragenlijstontwerp en in de testfase. Bij het herzien van een bestaand onderzoek kan de methode worden gebruikt om te kijken hoe een lopende vragenlijst werkt in het veld.

Uitgebreide omschrijving: Observatie kan op verschillende manieren worden uitgevoerd. Geheel onopvallend, bijvoorbeeld wanneer opnames worden gemaakt van CATI interviews of een audittrail wordt gemaakt, of heel opvallend, met als extreem voorbeeld het vastleggen van de oogbewegingen van een respondent met speciale apparatuur die op het hoofd wordt bevestigd.

Hieronder worden kort een aantal manieren van observeren beschreven:

- Gedragsobservaties. Dit houdt in dat het gedrag van een respondent en/of een interviewer geobserveerd wordt. Hierbij wordt zowel op lichaamstaal als op gesproken taal gelet, beide zijn belangrijk.
- Conversatieanalyse. Dit houdt in dat alleen het gesprek tussen de respondent en de interviewer geanalyseerd wordt. Lichaamstaal wordt buiten beschouwing gelaten, maar paralinguïstische uitingen (zoals zuchten of stemverheffing) kunnen wel worden gecodeerd. Er wordt gelet op het vragen naar verduidelijking van een vraag door de respondent, of het laten vallen van een stilte voordat een antwoord wordt gegeven, of het al dan niet veranderen van een voorgeschreven vraagtekst door een interviewer. In kader 4.8 staat een voorbeeld van een uitgeschreven conversatie uit een telefonisch interview. Voor meer informatie zie: Mazeland (1992) of Psathas (1995).
- Screen capture. Dit houdt in dat een opname wordt gemaakt van het beeldscherm dat een invuller voor zich heeft. Elke beweging van de muis, of alles wat ingevuld wordt, wordt zo opgenomen. Hiermee kan de bruikbaarheid van een webvragenlijst en ook het invulgedrag tijdens het afnemen van een webvragenlijst beoordeeld worden. Bijvoorbeeld: kan een invuller de juiste knop vinden om door de vragenlijst te bladeren, of zien de invullers de knop waaronder extra informatie te vinden is?

Kader 4.8 Voorbeeld van conversatieanalyse van een vraag-antwoordsequentie uit het Consumenten Conjunctuuronderzoek, 2002

Transcriptie:	Coderingsschema:
96. I Wat verwacht u van de komende twaalf maanden, zal dan de economische situatie in Nederland beter worden, slechter worden of hetzelfde blijven?	Alle zinnen worden genummerd. I = teksten van de interviewer. R = teksten van de respondent.
97. R (.) Mm: ik denk beter worden.	[] = De rechte haken geven aan waar interviewer en respondent tegelijk spreken.
98. I En wat denkt u, enigszins beter of duidelijk beter?	(0.0) = Tussen haakjes wordt de duur van een stilte aangegeven.
99. R Ik denk enigszins beter.	Een stilte die te kort is om in seconden aan te geven, maar die wel merkbaar aanwezig is wordt aangegeven met alleen een punt.
100. I De prijzen. vindt u dat de prijzen in de afgelopen twaalf maanden gestegen zijn, gedaald, of vrijwel gelijk gebleven?	: = Een dubbele punt achter een klank geeft aan dat deze langer dan normaal wordt aangehouden.
101. R U:h (.) ik denk dat ze enigszins gestegen zijn.	
102. I (2.0) En kan ik dan matig invullen of zwak?	
103. R Gestegen bedoelt [u?]	
104. I [Ja.]	
105. R Ja ik zou zeggen zwak.	

- Observatie in het kader van een cognitief interview. Hak en Muller (2001) beschrijven de mix van observatie en interview als het Drie-stappen Test-interview. Bij stap 1 voert de proefpersoon hardop denkend een taak uit (zoals het invullen van een vragenlijst). Bij stap 2 reconstrueren onderzoeker en proefpersoon vervolgens zoveel mogelijk wat er is gebeurd, met als doel een zo volledig mogelijk beeld te krijgen als aanvulling op de observatie bij stap 1. Bij stap 3 wordt ten slotte een kwalitatief interview gehouden waarbij aanvullende informatie en meningen kunnen worden gepeild.
- Audittrails. Deze worden ook wel logfiles genoemd. In de audittrails worden alle handelingen die de respondent op de computer uitvoert bijgehouden: elke muisklik en elke toetsaanslag in de tijd. Dit geeft net als bij de screen capture een beeld van hoe een respondent door een webvragenlijst is gegaan. Dit soort gegevens over het verloop van de enquête worden paradata genoemd. Zie voor meer informatie kader 4.9.

Voor het coderen en analyseren van gedrag dat is vastgelegd op beeld- of geluidsmateriaal, is op het CBS software beschikbaar: het programma Interact.

Kader 4.9 Toelichting audittrails

Als een respondent een vragenlijst invult, kunnen zijn acties worden vastgelegd in een audittrail. Hierin worden combinaties van actie, datum en tijdstip vastgelegd.

Aan de hand van de audittrails kan worden afgeleid welke acties men uitvoert en hoeveel tijd daaraan besteed wordt. Zo kan onder meer worden bepaald welke route men volgt in de vragenlijst, welke vragen zijn gewijzigd bij het invullen, hoeveel tijd nodig is voor het beantwoorden van een vraag en wanneer de hulpfunctie wordt geraadpleegd.

De toepassingen van audittrails zijn nog in ontwikkeling. De mogelijkheden zijn zeer divers zoals het identificeren van moeilijke vragen, vragen die als een blok worden ingevuld, problemen met het vragenlijstprogramma en het identificeren van type gebruikers (Morren & Snijkers, 2008).

Voor- en nadelen: Indien de observatie onopvallend gebeurt, ontstaat een realistisch beeld van het vraag-antwoordproces. Het kan gegevens opleveren van echte situaties in het veld. Observatie -zowel in het veld als in het laboratorium- levert rijke en realistische gegevens op over o.a. het invulgedrag, de bruikbaarheid van een vragenlijst en de interacties tussen interviewer en respondent. Het grootste nadeel van gedragsobservaties is dat zowel de dataverzameling als de analyse veel tijd kost.

4.9 Analyse van kwalitatief onderzoeksmateriaal

De informatie die wordt verzameld bij focusgroepen, cognitieve interviews en observaties is doorgaans kwalitatief van aard. Het is gemakkelijk om in dergelijk rijk materiaal het overzicht te verliezen en het is daarom raadzaam de gegevens op een systematische wijze te verwerken.

Een handzame werkwijze is de volgende: Maak aan de hand van het testprotocol of de topic list een tabel met in de kolomkoppen de onderzochte enquêtevragen en eventuele aandachtspunten en metavragen en in de rijkoppen de nummers van de proefpersonen of focusgroepen. Kijk of luister vervolgens naar de bandopnames en noteer per proefpersoon of focusgroep de bevindingen in de betreffende kolommen. Waar problemen met de vragenlijst naar voren komen dienen die uitvoerig te worden beschreven, liefst met behulp van letterlijke citaten en eventueel samen met een eerste idee over de oorzaak en oplossing van het probleem. Nadat zo alle interviews of gesprekken zijn beschreven is het zinvol om met meerdere onderzoekers het ruwe analyseschema te bespreken. Hierbij worden voorlopige hypothesen geformuleerd over de vragenlijst. Bijvoorbeeld: “Er zijn te veel antwoordcategorieën bij vraag x” of “Vraag y is niet te beantwoorden voor mensen die uitwonende kinderen hebben”. Sommige problemen en oplossingen die naar

voren komen zijn evident, bijvoorbeeld als een relevante antwoordmogelijkheid ontbreekt. Andere problemen vergen een zorgvuldige analyse van het materiaal, waarbij de voorlopige hypothesen worden getoetst aan het ruwe analyseschema of aan de originele bandopnames.

Bovenstaande methode is vooral geschikt om de vermoede risico's die aanleiding waren voor het formuleren van de metavragen systematisch te onderzoeken waarbij enig zicht op de aard en omvang van deze problemen ontstaat. Op basis daarvan worden dan dringende of minder urgente wijzigingen voorgesteld.

Anderzijds is het belangrijk bij het af luisteren of bekijken een open oor te hebben voor onvermoede problemen met de vragenlijst. Hoe ervaren een vragenlijst-ontwerper of -tester ook is, er kunnen geheel onverwachte dingen opduiken. Daarom zijn opnamen zo noodzakelijk en is de aanwezigheid van een onbevangen observator - althans bij enkele interviews - ook gewenst. Ook bovengenoemde groepsdiscussies bij de analyse kunnen helpen onvermoede problemen boven water te krijgen. Het kan dan aanbeveling verdienen de bandjes nogmaals te bestuderen om te kijken of zulke problemen incidenteel of structureel opgetreden zijn.

Kwalitatief onderzoek verdient – anders dan wel eens wordt gesuggereerd – een even systematische en gedegen analyse als kwantitatief onderzoek. Ook vereist het een even systematische en gedegen gegevensverzameling. Voor een overzicht, zie Denzin en Lincoln (2003).

4.10 Technisch testen

Korte omschrijving: Bij technische testen (ook wel functionele testen genoemd) wordt gecontroleerd of de computergestuurde vragenlijst volgens de specificaties is gebouwd.

Doel: Doel van het technische testen is het opsporen van fouten bij het bouwen van de vragenlijst. In deze testen wordt onder andere gecontroleerd of alle routes in de vragenlijst goed lopen en/of met alle voorkomende situaties rekening wordt gehouden.

Fase: Deze testen worden altijd uitgevoerd vóór de dataverzamelingsfase.

Uitgebreide omschrijving: De vragenlijst wordt getest met behulp van de "Testspecificatie techniek van de proces cyclus" (Pol, Teunissen & Veenendaal 2000). Met deze techniek worden eerst de routings (beslispunten) in de vragenlijst geïdentificeerd. Dit gebeurt door na elk beslispunt in de vragenlijst een nummer te geven aan elk uitgaand pad. Vervolgens worden de testsituaties gedefinieerd. Dat betekent dat alle combinaties van paden rond een beslispunt geïnventariseerd worden. Tot slot worden al deze testsituaties weer gecombineerd tot een compleet pad van het begin tot het eind van de vragenlijst. Er worden net zoveel complete paden gedefinieerd als nodig is om alle testsituaties in de test aan de orde te laten komen. Deze paden kunnen in protocollen vertaald worden, waarin elk in te vullen antwoord bij elke vraag uitgeschreven wordt. Dergelijke technische tests worden uitgevoerd vooraf aan de fase van dataverzameling. Het gaat hier met name om

testen van het proces of systeem van waaruit de data worden verzameld. Hierbij moet ook gedacht worden aan het testen van de technische testspecificaties van websurveys. Voorbeelden hiervan zijn het maken van automatische berekeningen, foutmeldingen, het versturen van de webvragenlijst et cetera.

Voor- en nadelen: Deze techniek levert ondubbelzinnige testresultaten: er is wel of geen fout gevonden. Het herstellen van fouten in de vragenlijst en deze vervolgens aanpassen in de testopzet vereist een goede coördinatie van de verschillende betrokkenen zoals data proces experts, onderzoekers en vragenlijstontwerpers.

4.11 Experimenten

Korte omschrijving: Bij experimenteel onderzoek naar vragenlijsten worden doorgaans alternatieve vraagstellingen of vragenlijsten aan een random steekproef uit de onderzoekspopulatie voorgelegd. Bij experimenten worden condities gemanipuleerd (bijvoorbeeld de versies van een vraag). Deze condities worden op basis van toeval voorgelegd aan de respondenten.

Doel: In experimenten kunnen verschillende varianten van een vragenlijst of verschillende modes worden getest en gecontroleerd om meetfouten vast te stellen. Aanleiding voor het houden van experimenten zijn bijvoorbeeld twee alternatieve vraagstellingen die in een eerdere discussie of pretest niet tot een definitieve keuze hebben geleid. Ook worden experimenten vaak gebruikt om mogelijke trendbreuken te onderzoeken indien een nieuwe vraagstelling wordt gebruikt of het uitvragen in een andere modus wordt toegepast.

Fase: Experimenten kunnen worden uitgevoerd als onderdeel van een pilot in de testfase of meedraaien in lopende surveys, de fase van dataverzameling.

Uitgebreide omschrijving: Experimenten kunnen worden gehouden met losse vragen, met een set van vragen of met een gehele vragenlijst. Voorwaarde is dat de steekproefomvang in een experiment groot genoeg is zodat significante verschillen gemeten kunnen worden. Het is ook noodzakelijk dat het design zodanig is dat het toewijzen van varianten van de vragenlijstversies op een gerandomiseerde manier gebeurt zodat verschillen kunnen worden toegeschreven aan de vragenlijst en niet aan de samenstelling van de steekproef. Zie ook Van den Brakel (2008) over het houden van experimenten in de context van een reguliere dataverzameling. Voor een checklist voor het opzetten van een experiment zie kader 4.10.

Voor- en nadelen: Het opzetten en uitvoeren van een experiment is een intensieve en relatief dure manier van testen en kent een lange doorlooptijd voordat resultaten bekend zijn. Resultaten leveren wel betrouwbare en kwantificeerbare conclusies op over vraagvarianten of modes.

Kader 4.10 Checklist opzetten van een experiment

- Definieer en beschrijf het probleem
- Bepaal de varianten van de vragen(lijsten)
- Formuleer de hypothesen die getoetst kunnen worden
- Bepaal ontwerp en grootte van het experiment
- Voer het experiment uit
- Toets je hypothesen en interpreteer de resultaten
- Neem besluiten en implementeer

(Statistics Sweden, 2004, pag. 65)

4.12 Analyses van procesgegevens

Korte omschrijving: Bij de implementatie van een vragenlijst kunnen tal van gegevens worden vastgelegd over hoe het proces van dataverzameling is verlopen. Dit worden ook wel paradata genoemd. Het analyseren van deze gegevens kan inzicht geven in het functioneren van de vragenlijst.

Doel: Naast diverse testmethoden voorafgaand aan implementatie van de vragenlijst, kan men ook na implementatie van de vragenlijst, als deze in het veld is voor dataverzameling, onderzoeken of het waarneeminstrument aan de gestelde kwaliteitsdoelstellingen voldoet.

Fase: De procesdata moeten logischerwijs in de dataverzamelingsfase worden verzameld en kunnen zowel tijdens de dataverzameling (bijvoorbeeld na de eerste tien dagen van een veldwerkperiode of de eerste 100 respondenten) als na afloop daarvan worden geanalyseerd.

Uitgebreide omschrijving: Met procesgegevens of paradata bedoelen we de gegevens die over het proces van dataverzameling worden vastgelegd. Denk hierbij aan informatie over hoe vaak een respondent gerappelleerd is, op welk tijdstip een interview is afgenomen of bijvoorbeeld een audit trail waarin precies wordt vastgelegd hoe een elektronische vragenlijst is gebruikt. Deze gegevens kunnen als onbedoeld bijproduct van de dataverzameling ontstaan, maar ook expliciet worden verzameld voor evaluatiedoeleinden. Hieronder noemen we een aantal veelgebruikte methoden om procesdata te gebruiken voor het evalueren van vragenlijsten:

- Analyse van de hoogte en duur van respons: Cijfers over de unitrespons en de tijd tussen het uitsturen van een vragenlijst en het ontvangen van ingevulde of deels ingevulde vragenlijsten kunnen inzicht geven in welke vragenlijsten of bij welke doelgroepen er problemen zijn.
- Kwantitatieve of kwalitatieve analyse van opmerkingen en vragen van respondenten die worden vastgehouden in het Customer Relationship Management (CRM) systeem kunnen inzicht geven in welke vragenlijsten de meeste problemen oproepen en welk soort problemen dat zijn. Zie kader 4.11

voor een voorbeeld hiervan. Dit geldt voor bedrijfsvragenlijsten. Voor persoonsvragenlijsten gaat het om opmerkingen van respondenten die door interviewers worden vastgelegd of vragen van interviewers die bij het CBS binnenkomen.

- Metavragen in de uitgezette vragenlijst: met deze methode worden in de vragenlijst metavragen opgenomen die aan de respondenten gesteld worden bij bepaalde vragen of aan het einde van de vragenlijst. Men kan vragen stellen over de responslast en de perceptie van de vragenlijst. Hiermee kan men de kwaliteit van het vraag-antwoordproces meten en lastige of gevoelige vragen opsporen. Het CBS maakt hiervan weinig gebruik maar vele marktonderzoeksbureaus stellen stelselmatig enkele vragen aan het einde van het gesprek, o.a. om te achterhalen welke onderwerpen goed of slecht in de markt liggen. Indien deze vragen in een afzonderlijk interviewtje worden gesteld, is sprake van RAS, zoals hieronder besproken.
- RAS (Response Analysis Survey) is een gestandaardiseerd nagesprek met (non)respondenten. RAS wordt vooral gebruikt voor a) controle op de interviewers en b) analyse van de nonrespons. In dat laatste geval worden aan nonrespondenten enkele kernvragen gesteld waarmee inzicht kan worden verkregen in kenmerken van nonrespondenten. Het CBS houdt jaarlijks een klanttevredenheidsonderzoek onder (non-)respondenten van bepaalde enquêtes. Dit kan ook gezien worden als een vorm van RAS.
- Kort vragenlijstje over procesdata gesteld aan interviewers. Ook interviewers kunnen informatie verstrekken over het proces, bijvoorbeeld of anderen dan de respondent aanwezig waren bij het interview, of vragen als erg indringend werden ervaren etc. Dergelijke vragen worden in een afzonderlijke lijst aan de interviewers voorgelegd.

Voor- en nadelen: Deze methoden leveren gegevens op over het daadwerkelijke dataverzamelingsproces die doorgaans kwantitatief geanalyseerd kunnen worden. Hiermee kunnen betrouwbare uitspraken worden gedaan over waar kwaliteitsproblemen optreden en hoe groot deze problemen zijn. Er zijn twee grote nadelen aan de analyse van procesdata. In de eerste plaats kun je er alleen problemen mee vast stellen die al in het veld zijn. Beter is het natuurlijk om deze problemen te voorkomen. In de tweede plaats geven deze methoden meestal weinig inzicht in de oorzaken en mogelijke oplossingen van de problemen. Bij een revisieronde kunnen inzichten op basis van procesdata daarom het beste worden gebruikt om nader kwalitatief onderzoek te sturen

Kader 4.11 Voorbeeld analyse CRM gegevens

Resultaten codering opmerkingen van respondenten over de Productiestatistiek 2003 (Giesen & van Putten, 2004).

	Hoofdcode	Aantal	Percentage
1	Opmerkingen over specifieke vragen / blokken vragen	291	13,0
2	Opmerkingen over onduidelijkheid van vragenlijst in het algemeen	26	1,2
3	Opmerking over aansluiting vragenlijst / administratie	610	27,3
4	Opmerkingen over responslast	665	29,8
5	Algemene opmerking over kwaliteit van de antwoorden	179	8,0
6	Overige opmerkingen	461	20,7
	Totaal	2232	100

4.13 Analyses van surveydata

Korte omschrijving: Vragenlijsten bepalen mede de kwaliteit van de verzamelde gegevens. Omgekeerd kunnen dus problemen met de kwaliteit van de verzamelde gegevens duiden op problemen met de vragenlijst.

Doel: Het doel van het analyseren van surveydata als evaluatiemethode voor vragenlijsten is doorgaans de kwaliteit van de verzamelde gegevens beoordelen en onderzoeken in hoeverre problemen in de kwaliteit veroorzaakt worden door fouten in de vragenlijst.

Uitgebreide omschrijving: Er zijn verschillende methoden om de kwaliteit van surveydata te beoordelen. In het kader van de evaluatie van vragenlijsten wordt veelal gekeken naar item-nonrespons, de plausibiliteit van de gegevens, invulpatronen van antwoordschalen en de spreiding over de antwoordopties. Het valt buiten het bestek van dit deel van de Methodenreeks om hier uitgebreid op in te gaan. Een nuttig overzicht van kwaliteitsmaten voor surveydata is onder andere te vinden in Lyberg et al. (1997).

Voor- en nadelen: De voor- en nadelen van het analyseren van surveydata zijn vergelijkbaar met die van het analyseren van procesdata. Een extra nadeel is dat er naast de vragenlijst zelf nog vele andere aspecten in het dataverzamelingsproces van invloed zijn op de datakwaliteit (denk bijvoorbeeld aan rappels, ondersteuning van de buitendienst en de modus van dataverzameling). Het is niet altijd eenduidig vast te stellen in hoeverre kwaliteitsproblemen in de surveydata direct toe te schrijven zijn aan problemen met de vragenlijst.

5. Afsluiting

Bij het CBS worden de meeste nieuwe vragenlijsten en vernieuwde vragenlijsten getest. Dit gebeurt op verschillende manieren en op verschillende afdelingen, met name door de afdelingen Ontwikkeling en Ondersteuning van de Bedrijfseconomische Statistieken (BOO) en de Sociale en Ruimtelijke Statistieken (SOO), Dataverzameling (PDV) en Methodologie en Kwaliteit (DMK). Bij DMK is er een groep waarnemingsmethodologen die zich in het testen van vragenlijsten gespecialiseerd heeft: het Vragenlab.

Het Vragenlab geeft adviezen aan vragenlijstontwikkelaars door het testen van de nieuwe of vernieuwde vragenlijsten. Het Vragenlab maakt daarbij gebruik van alle testmethoden die in hoofdstuk 4 besproken zijn. Een vragenlijstontwikkelaar kan bij het Vragenlab terecht met vragen over aanschrijfbrieven, benaderingswijzen, vraagstellingen, bruikbaarheid en toepasbaarheid van vragenlijsten of voor pilot-onderzoeken en experimenten.

Het zal zelden voorkomen dat binnen één testprogramma alle methoden voor het evalueren van vragenlijsten worden gebruikt. Daarvoor ontbreken doorgaans de tijd en middelen. Voor het testen van een vragenlijst blijft vaak weinig tijd over omdat eerdere fases van de ontwikkeling uitlopen. Er zal dan op basis van doelen en middelen een selectie plaats moeten vinden. Naar Dehue (1997) en Snijkers (2002) kan het volgende minimum pakket voor een test- en evaluatieprogramma worden opgesteld:

Tabel 5.1. Minimum pakket voor test- en evaluatieprogramma

Fase	Vragenlijstontwikkeling	Methoden
1 & 2	Conceptualisering en onderzoeksontwerp & vragenlijstontwerp	<ul style="list-style-type: none">▪ Review van documenten en literatuur▪ Informele test▪ Expert raadpleging▪ Indien bovenstaande onvoldoende overtuigt: demonstratie met focusgroep en/of cognitieve interviews
3	Testen	<ul style="list-style-type: none">▪ Observatie van een aantal standaardinterviews met proefrespondenten, gecombineerd met globale evaluatievragen over inhoud en procedure vragenlijst
5 & 6	Dataverzameling en proces- monitoring en evaluatie	<ul style="list-style-type: none">▪ Observatie van een aantal interviews in het veld, gecombineerd met evaluatievragen voor respondent en/of interviewer▪ Nagesprekken met respondenten▪ Evaluatieve focusgroepen met interviewers▪ (Item) non-responsanalyse▪ Data analyse

De ervaring leert dat ook dit minimale pakket niet altijd haalbaar is. We adviseren om in ieder geval iedere vragenlijst uit te testen met een aantal proefpersonen uit de doelgroep en daarnaast zoveel mogelijk verschillende methoden te gebruiken. Indien verschillende methoden leiden tot dezelfde conclusies geeft dat reden tot vertrouwen in deze conclusies. Daarnaast is het van groot belang dat de resultaten uit de gebruikte test- en evaluatiemethoden teruggekoppeld worden naar onderzoekers en vragenlijstontwikkelaars zodat aanpassingen en verbeteringen kunnen worden doorgevoerd. Zoals al aangegeven in hoofdstuk 2 (zie ook figuur 2.1), is vragenlijstontwikkeling een iteratief proces, waarbij informatie uit elke fase in het proces als input kan dienen voor een verbetering van de conceptvragenlijst of update van de bestaande vragenlijst.

6. Literatuur

- Beukenhorst, D., Dam, G. van, Göttgens, R., Sniijkers, G. en Verbruggen, M. (2005), *CBS brede waarneemstrategie*. CBS Heerlen, internetnotitie. BPA-nummer: BOO-2005-H09, Datum: 30-11-2005.
- Brakel, J.A. van den (2008), Design-based analysis of embedded experiments with applications in the Dutch Labour Force Survey. *Journal of the Royal Statistical Society (Series A)* 171 (3), 581-613 .
- Bureau Taal (2006), *Schrijfwijzer Begrijpelijke Formulieren*. SDU. www.begrijpelijkeformulieren.nl
- Biemer, P. en Lyberg, L. (2003), *Introduction to Survey Quality*. Wiley, New York.
- Billiet, J. (2006), De gestandaardiseerde vragenlijsten. In: Billiet, J. en Waage, H. (Red.). *Een samenleving onderzocht*. De Boeck, Antwerpen, pp. 223-283.
- Brancato, G., Macchia, S., Murgia, M., Siognore, M., Simeoni, G., Blanke, K., Körner, T., Nimmergut, A., Lima, P., Paulino, R. en Hoffmeyer-Zlotnik, J. (2006), *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System*. European Commission Grant Agreement 2004103000002.
- Campanelli, P. (2008), Testing survey questions. In: De Leeuw, E., Hox, J. and Dillman D. (Eds.). *International Handbook of Survey Methodology*. Lawrence Erlbaum, New York, pp 176-200.
- Commissie voor de Europese Gemeenschappen (2005), *Europese Praktijkcode, Over de onafhankelijkheid, integriteit en verantwoordingsplicht van de nationale en communautaire statistische instanties*. Brussel, COM(2005) 217 definitief. http://circa.europa.eu/Public/irc/dsis/coded/library?l=/2005_final_nlpdf/NL_1.0_&a=d
- Couper, M. (2008), *Designing effective web surveys*. Cambridge University Press, Cambridge.
- Dehue, F. (1997), *Standaardprotocollen voor de ontwikkeling en de test van dataverzamelingsprocedures*. CBS Heerlen, interne nota: H7247-97-GWM.
- De Leeuw, Hox, J. en Dillman, D. (Eds.) (2008), *International Handbook of Survey Methodology*. Lawrence Erlbaum, New York, .
- Denzin, N., Lincoln, Y. (Eds.) (2003), *Collecting and interpreting qualitative materials*. Sage, Thousands Oaks.
- Dillman, D., Smyth, D. & Christian, J. (2009), *Internet, Mail and Mixed Mode Surveys. The Tailored Design Method*. New York: Wiley.

- Dillman, D. en Redline, C. (2004), Testing paper self-administered questionnaires: cognitive interview and field test compared. In: Presser, S., Rotgheb, J., Couper, M., Lessler, J., Martin, E., Martin, J. and Singer, E. (Eds). *Methods for Testing and Evaluating Survey Questionnaires*. Wiley, New York, pp 299-317.
- Eurostat (2005), *European statistics code of practice for the national and community statistical authorities*. European Union,
http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/VERSIONE_INGLESE_WEB.PDF
- Fowler, F. (1995), Improving survey questions: design and evaluation. *Applied Social Research Methods Series*, vol. 38, Sage Publications.
- Fowler, F., Roman, A. & Di, Z. (1998), Mode effects in a survey of Medicare prostate surgery patients. *Public Opinion Quarterly* 62, 29-46.
- Fowler, F. en Cosenza, C. (2008), Writing effective survey questions. In: De Leeuw, E., Hox, J. and Dillman D. (Eds.), *International Handbook of Survey Methodology*. Lawrence Erlbaum, New York, pp 136-160.
- Gemeente Den Haag. Website geraadpleegd op 27 maart 2009. *Helder Haags Woordenboek*. <http://www.denhaag.nl/smartsite.html?id=55844>.
- Giesen, D. (2004), *Evaluatie vragenlijsten productiestatistieken: Resultaten en aanbevelingen*. CBS, Heerlen: Interne nota, BPA H4415-04-TMO.
- Giesen, D. en Putten, J. van (2004), Analyse van opmerkingen van berichtgevers in Logiquest. In Projectgroep Evaluatie PS *Resultaten Inventarisatiefase Evaluatie Vragenlijsten Productiestatistiek*. CBS, Heerlen, interne nota: H2058-04-TMO/intern, pp.12-19.
- Grice, H.P. (1975), Logic and Conversation. In Cole, P. & Morgan, J.L. (Eds.), *Speech Acts*. Academic Press, New York, pp 41-58.
- Hak, T. en Muller, J. (2001), De toolkit van de kwalitatieve marktonderzoeker. Het Drie-Stappen Test-Interview. *Onderzoek*, 18 (12), pp 8-9.
- Houtkoop-Steenstra, H. (2000), *Interaction and the standardized survey interview: the living questionnaire*. Cambridge University Press.
- Jansen, C., Steehouder, M., Edens, K., Mulder, J., Pander Maat, H., Slot, P. (1989), *Formulierenwijzer. Handboek formulieren redigeren*. SDU, Den Haag.
- Krosnick, J. (1991), Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5, 213-236.
- Krosnick, J., Holbrook, A., Berent, M., Carson, R., Hanemann, W., Kopp, R., Mitchel, R., Presser, S., Ruud, P., Smith, V., Moody, W., Green, M. en Conaway, M. (2002), The impact of “no opinion” response options on data quality. *Public Opinion Quarterly* 66, 371-404.
- Krueger, R. A. (1994), *Focus groups: A practical guide for applied research*. Sage Publications, London.

- Lyberg, L., Biemer, P., Collins, M. de Leeuw, E., Dippo, C., Schwarz, N. en Dennis, T. (Eds.) (1997), *Survey Measurement and Process Quality*. Wiley, New York.
- Maynard, D., Houtkoop-Steenstra, H., Cate Schaeffer, N. en Zouwen, J. van der (Eds.) (2002), *Standardization and Tacit Knowledge*. Wiley, New York.
- Mazeland, H. (1992), *Vraag/antwoord sequenties*. Stichting Neerlandistiek VU, Amsterdam.
- Morren, M. en Snijkers, G. (2008), *Utility of audit trails in the annual structural business survey*. CBS, Heerlen: interne nota DMH-2008-09-26 MMRN.
- Pol, M., Teunissen, R. en Veenendaal, E. van (2000), *Testen volgens TMAP*. Tutein Nilthenius, 's Hertogenbosch (par. 15.11 Procescyclustest).
- Psathas, G. (1995), *Conversation Analysis: the study of talk-in-interaction*. Sage Publications Inc., Thousand Oaks.
- Redline, C. en Dillman, D. (2002), The influence of alternative visual designs on respondents' performance with branching instructions in self-administered questionnaires. In: R. Groves, Dillman, D., Eltinge, J., and R. Little (Eds.), *Survey Nonresponse*. Wiley, New York.
- Schwarz, N., Knäuper, B., Hippler, H-J., Noelle-Neuman, E. & Clark, F. (1991), Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly* 55, 618-630.
- Snijkers, G. (2002), *Cognitive laboratory experiences: on pre-testing computerised questionnaires and data quality*. Proefschrift: CBS/UU Utrecht.
- Statistics Sweden. (2004), *Design your questions right*. Sweden.
- Sudman, S en Bradburn, N M (1983). *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco, CA: Jossey-Bass
- Toepoel, V., Vis, C., Das, M. en Soest, A. van (2009), Design of Web Questionnaires. An information-processing perspective for the effect of response categories. *Sociological Methods & Research* 37, 371-392.
- Tourangeau, R., Rips, L.J. en Rasinski, K. (2000), *The Psychology of Survey Response*. Cambridge University Press, New York.
- Vis-Visschers, R. (2002), *Conversatie Analyse en het testen van survey-vragen*. CBS, Heerlen, interne nota H3330-02-TMO/intern.
- Vis-Visschers R. (2007), *Dutch Testing Phase Feedback of the Questionnaire on International Sourcing*. CBS Heerlen: Interne nota.
- Vis-Vischers, R., Arends-Tóth, J., Giesen, D. en Meertens, V. (2008), *Het aanbieden van weet niet en toelichtingen in een webvragenlijst*. CBS Heerlen: interne nota.
- Vis-Visschers, R., Morren, M., Meertens, V. en Arends-Tóth, J. (2008), *CBS cognitive test of the health care module and the health determinants module of*

the European health interview survey. Heerlen: interne nota. DMH-2008-06-30-RVCS.

- Vosmer, M. en Engelen, F. (2008), Het herzien van de vragenlijst van het NKO 2006. In H. Schmeets (Ed.) . *Het Nationaal Kiezersonderzoek 2006. Opzet, uitvoering en resultaten*. Centraal Bureau voor de Statistiek, Voorburg/Heerlen, 55-61.
- Willeboordse, A. (2008), *Methodenreeks: Inleiding in het statistisch proces en de Methodenreeks*. CBS Voorburg.
- Willimack, D. en Nichols, E. (2001), Building an Alternative Response Process Model for Business Surveys. *U.S. Proceedings of the Annual Meeting of the American Statistical Association*, Alexandria, V.A.: American Statistical Association
- Willis, G. (2005), *Cognitive interviewing: a tool for improving questionnaire design*. Sage Publications: London.

Bijlage 1. Taalniveaus

Taalniveaus volgens Common European Framework of Reference

	Begrijpen		Spreken		Schrijven
	Luisteren	Lezen	Productie	Interactie	
C2	Ik kan moeiteloos gesproken taal begrijpen, in welke vorm dan ook, hetzij in direct contact, hetzij via radio of tv, zelfs wanneer in een snel moedertaaltempo gesproken wordt als ik tenminste enige tijd heb om vertrouwd te raken met het accent.	Ik kan moeiteloos vrijwel alle vormen van de geschreven taal lezen, inclusief abstracte, structureel of linguïstisch complexe teksten, zoals handleidingen, specialistische artikelen en literaire werken.	Ik kan een duidelijke, goed lopende beschrijving of redenering presenteren in een stijl die past bij de context en in een doeltreffende logische structuur, zodat de toehoorder in staat is de belangrijke punten op te merken en te onthouden.	Ik kan zonder moeite deelnemen aan welk gesprek of discussie dan ook en ben zeer vertrouwd met idiomatische uitdrukkingen en spreektaal. Ik kan mezelf vloeiend uitdrukken en de fijnere betekenissen precies weergeven. Als ik een probleem tegenkom, kan ik mezelf hernemen en mijn betoog zo herstructureren dat andere mensen het nauwelijks merken.	Ik kan een duidelijke en vloeiende lopende tekst in een gepaste stijl schrijven. Ik kan complexe brieven, verslagen of artikelen schrijven waarin ik een zaak weergeef in een doeltreffende, logische structuur, zodat de lezer de belangrijke punten kan opmerken en onthouden. Ik kan samenvattingen van en kritieken op professionele of literaire werken schrijven.
C1	Ik kan een langer betoog begrijpen, zelfs wanneer dit niet duidelijk gestructureerd is en wanneer relaties slechts impliciet zijn en niet expliciet worden aangegeven. Ik kan zonder al te veel inspanning tv-programma's en films begrijpen.	Ik kan lange en complexe feitelijke en literaire teksten begrijpen, en het gebruik van verschillende stijlen waarderen. Ik kan gespecialiseerde artikelen en lange technische instructies begrijpen, zelfs wanneer deze geen betrekking hebben op mijn terrein.	Ik kan duidelijke, gedetailleerde beschrijvingen geven over complexe onderwerpen en daarbij sub-thema's integreren, specifieke standpunten ontwikkelen en het geheel afronden met een passende conclusie.	Ik kan mezelf vloeiend en spontaan uitdrukken zonder merkbaar naar uitdrukkingen te hoeven zoeken. Ik kan de taal flexibel en effectief gebruiken voor sociale en professionele doeleinden. Ik kan ideeën en meningen met precisie formuleren en mijn bijdrage vaardig aan die van andere sprekers relateren.	Ik kan me in duidelijke, goed gestructureerde tekst uitdrukken en daarbij redelijk uitgebreid standpunten uiteenzetten. Ik kan in een brief, een opstel of een verslag schrijven over complexe onderwerpen en daarbij de voor mij belangrijke punten benadrukken. Ik kan schrijven in een stijl die is aangepast aan de lezer die ik in gedachten heb.
B2	Ik kan een langer betoog en lezingen begrijpen en zelfs complexe redeneringen volgen, wanneer het onderwerp redelijk vertrouwd is. Ik kan de meeste nieuws- en actualiteitenprogramma's op de tv begrijpen. Ik kan het grootste deel van films in standaarddialect begrijpen.	Ik kan artikelen en verslagen lezen die betrekking hebben op eigentijdse problemen, waarbij de schrijvers een bepaalde houding of standpunt innemen. Ik kan eigentijds literair proza begrijpen.	Ik kan duidelijke, gedetailleerde beschrijvingen presenteren over een breed scala van onderwerpen die betrekking hebben op mijn interessegebied. Ik kan een standpunt over een actueel onderwerp verklaren en de voordelen en nadelen van diverse opties uiteenzetten.	Ik kan zodanig deelnemen aan een vloeiend en spontaan gesprek, dat normale uitwisseling met moedertaalsprekers redelijk mogelijk is. Ik kan binnen een vertrouwde context actief deelnemen aan een discussie en hierin mijn standpunten uitleggen en ondersteunen.	Ik kan een duidelijke, gedetailleerde tekst schrijven over een breed scala van onderwerpen die betrekking hebben op mijn interesses. Ik kan een opstel of verslag schrijven, informatie doorgeven of redenen aanvoeren ter ondersteuning vóór of tégen een specifiek standpunt. Ik kan brieven schrijven waarin ik het persoonlijk belang van gebeurtenissen en ervaringen aangeef.
B1	Ik kan de hoofdpunten begrijpen wanneer in duidelijk uitgesproken standaarddialect wordt gesproken over vertrouwde zaken die ik regelmatig tegenkom op mijn werk, school, vrije tijd enz. Ik kan de hoofdpunten van veel radio- of tv-programma's over actuele zaken of over onderwerpen van persoonlijk of beroepsmatig belang begrijpen, wanneer er betrekkelijk langzaam en duidelijk gesproken wordt.	Ik kan teksten begrijpen die hoofdzakelijk bestaan uit hoogfrequente, alledaagse of aan mijn werk gerelateerde taal. Ik kan de beschrijving van gebeurtenissen, gevoelens en wensen in persoonlijke brieven begrijpen.	Ik kan uitingen op een simpele manier aan elkaar verbinden, zodat ik ervaringen en gebeurtenissen, mijn dromen, verwachtingen en ambities kan beschrijven. Ik kan in het kort redenen en verklaringen geven voor mijn meningen en plannen. Ik kan een verhaal vertellen, of de plot van een boek of film weergeven en mijn reacties beschrijven.	Ik kan de meeste situaties aan die zich kunnen voordoen tijdens een reis in een gebied waar de betreffende taal wordt gesproken. Ik kan onvoorbereid deelnemen aan een gesprek over onderwerpen die vertrouwd zijn, of mijn persoonlijke belangstelling hebben of die betrekking hebben op het dagelijks leven (bijvoorbeeld familie, hobby's, werk, reizen en actuele gebeurtenissen).	Ik kan eenvoudige samenhangende tekst schrijven over onderwerpen die vertrouwd of van persoonlijk belang zijn. Ik kan persoonlijke brieven schrijven waarin ik mijn ervaringen en indrukken beschrijf.

		Begrijpen		Spreeken		Schrijven
		Luisteren	Lezen	Productie	Interactie	
A2	Ik kan zinnen en de meest frequente woorden begrijpen die betrekking hebben op gebieden die van direct persoonlijk belang zijn (bijvoorbeeld basisinformatie over mezelf en mijn familie, winkelen, plaatselijke omgeving, werk). Ik kan de belangrijkste punten in korte, duidelijke eenvoudige boodschappen en aankondigingen volgen.	Ik kan zeer korte eenvoudige teksten lezen. Ik kan specifieke voorspelbare informatie vinden in eenvoudige, alledaagse teksten zoals advertenties, folders, menu's en dienstregelingen en ik kan korte, eenvoudige, persoonlijke brieven begrijpen.	Ik kan een reeks uitdrukkingen en zinnen gebruiken om in eenvoudige bewoordingen mijn familie en andere mensen, leefomstandigheden, mijn opleiding en mijn huidige of meest recente baan te beschrijven.	Ik kan communiceren over eenvoudige en alledaagse taken die een eenvoudige en directe uitwisseling van informatie over verwante onderwerpen en activiteiten betreffen. Ik kan zeer korte sociale gesprekken aan, alhoewel ik gewoonlijk niet voldoende begrip om het gesprek zelfstandig gaande te houden.	Ik kan korte, eenvoudige notities en boodschappen opschrijven. Ik kan een zeer eenvoudige persoonlijke brief schrijven, bijvoorbeeld om iemand voor iets te bedanken.	
A1	Ik kan vertrouwde woorden en basiszinnen begrijpen die mezelf, mijn familie en directe concrete omgeving betreffen, wanneer de mensen langzaam en duidelijk spreken.	Ik kan vertrouwde namen, woorden en zeer eenvoudige zinnen begrijpen, bijvoorbeeld in mededelingen, op posters en in catalogi.	Ik kan eenvoudige uitdrukkingen en zinnen gebruiken om mijn woonomgeving en de mensen die ik ken, te beschrijven.	Ik kan deelnemen aan een eenvoudig gesprek, wanneer de gesprekspartner bereid is om zaken in een langzamer spreektempo te herhalen of opnieuw te formuleren en mij helpt bij het formuleren van wat ik probeer te zeggen. Ik kan eenvoudige vragen stellen en beantwoorden die een directe behoefte of zeer vertrouwde onderwerpen betreffen.	Ik kan een korte, eenvoudige ansichtkaart schrijven, bijvoorbeeld voor het zenden van vakantiegroeten. Ik kan op formulieren persoonlijke details invullen, bijvoorbeeld mijn naam, nationaliteit en adres noteren op een hotelinschrijvingsformulier.	

© Council of Europe / Conseil de l'Europe

http://www.coe.int/T/DG4/Portfolio/documents/assessment_grid_dutch.doc

Bijlage 2. Checklist Vragenlijstontwikkeling

Versie maart 2009

Gebruik van de checklist

Deze checklist is bedoeld als hulpmiddel om na te gaan of er niets is vergeten. Er kunnen goede redenen zijn om op een onderdeel 'nee' of 'min of meer' te scoren. Zo is het bijvoorbeeld voor het overzicht op sommige bedrijfsenquêtes veel beter om de vragen in korte steekwoorden te formuleren in plaats van als vraagteksten (zie punt B1 uit de checklist). Als er echter geen redenen zijn om af te wijken van de checklist, adviseren wij de vragenlijst aan te passen.

De checklist is o.a. gebaseerd op de handboeken *Formulierenwijzer* (Jansen, et al., 1989; <http://waaier.begrijpelijkeformulieren.nl/upload/files/formulierenwijzer.pdf>), en de checklist Begrijpelijke formulieren zoals deze is gebruikt bij de workshop Begrijpelijke Formulieren.

Aanvullende checklists en tips voor het maken van vragenlijsten zijn te vinden op www.begrijpelijkeformulieren.nl -> dertig tips / schrijfwijzer begrijpelijke formulieren

www.formulierenwaaier.nl

www.wetsmatiging.be -> Hoe maak ik een goed formulier

www.bizztools.info/downloads/vragenlijst.pdf

Checklist Vragenlijstontwikkeling Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling	Ja Min of meer Nee Niet van toepassing
--	---

A. Algemene zaken

1. Geeft de ingevulde vragenlijst antwoord op de onderzoeksvragen?	□ □ □ □
2. Wordt alleen informatie gevraagd die noodzakelijk is? ➤ Vraag geen gegevens die al bekend zijn. ➤ Neem nooit een vraag op “omdat het misschien interessant is”.	□ □ □ □
3. Wordt er voldoende algemene informatie over het doel van de vragenlijst verstrekt?	□ □ □ □
4. Wordt in een begeleidende brief of introductie bij de vragenlijst voldoende duidelijk gemaakt wat van de respondent wordt verwacht?	□ □ □ □
5. Heeft de vragenlijst een duidelijke en aansprekende titel?	□ □ □ □
6. Kan de respondent de gevraagde gegevens gemakkelijk verstrekken?	□ □ □ □
7. Is de vragenlijst getest met proefrespondenten uit de doelgroep?	□ □ □ □

B. Formulering vraagstelling

1. Zijn alle vragen in de vorm van een vraag gesteld? ➤ Sluit elke vraag af met een vraagteken.	□ □ □ □
2. Zijn CAPI en CATI vragen makkelijk op te lezen?	□ □ □ □
3. Worden er alleen enkelvoudige vragen gesteld? ➤ Stel slechts één vraag tegelijk. Woorden als ‘en’, ‘of’, ‘omdat’ en ‘vanwege’ wijzen vaak op een meervoudige vraagstelling.	□ □ □ □
4. Zijn de vragen ondubbelzinnig interpreteerbaar? ➤ Geef definities van begrippen die verschillend kunnen worden opgevat of onduidelijk zijn.	□ □ □ □
5. Is steeds duidelijk op wie, op wat en op welke periode de vraag betrekking heeft? ➤ Geef indien nodig referentieperiodes aan. ➤ Kies de referentieperiode in relatie tot het belang van de gebeurtenis - hoe onbelangrijker de gebeurtenis, hoe korter de referentieperiode.	□ □ □ □
6. Zijn de vragen vrij van veronderstellingen? ➤ Dus niet: “Zoals u weet..”	□ □ □ □
7. Zijn dubbele ontkenningen vermeden?	□ □ □ □
8. Zijn de vragen en zinnen kort?	□ □ □ □

Checklist Vragenlijstontwikkeling	Ja Min of meer Nee Niet van toepassing
Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling	

<p>9. Is eenvoudig Nederlands gebruikt?</p> <ul style="list-style-type: none"> ➤ Kies korte en alledaagse woorden. ➤ Wees consistent in de woordkeuze. ➤ Gebruik geen jargon of abstracte termen. ➤ Formuleer teksten actief en positief 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
10. Zijn spelling en interpunctie correct?	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

C. Antwoordmogelijkheden

<p>1. Is steeds duidelijk wat voor antwoord er verwacht wordt?</p> <ul style="list-style-type: none"> ➤ De antwoordmogelijkheden moeten aansluiten bij de vraag. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>2. Worden alle mogelijke antwoordcategorieën vermeld?</p> <ul style="list-style-type: none"> ➤ Neem eventueel een restcategorie op: Anders, te weten: ____ ➤ Noem de restcategorie steeds als laatste. ➤ Noem de meest voorkomende antwoordcategorieën als eerste. ➤ Voeg bij vragen waar “weet niet “ of “weigert” relevante antwoordmogelijkheden zijn expliciet deze mogelijkheden aan. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>3. Sluiten de antwoordcategorieën elkaar uit?</p> <ul style="list-style-type: none"> ➤ Zorg dat er geen overlap is tussen de mogelijkheden. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>4. Is de betekenis van de antwoordmogelijkheden duidelijk?</p> <ul style="list-style-type: none"> ➤ Zijn de antwoordmogelijkheden duidelijk interpreteerbaar? 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

D. Toelichtingen en instructies

<p>1. Zijn alleen waar het nodig is toelichtingen en instructies gegeven?</p> <ul style="list-style-type: none"> ➤ Hoe meer tekst, hoe minder men geneigd is het te lezen. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>2. Zijn de toelichtingen en instructies zo dicht mogelijk bij de vraag geplaatst?</p>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>3. Zijn de toelichtingen en instructies helder en duidelijk?</p> <ul style="list-style-type: none"> ➤ Maak korte zinnen: “Denkt u aan ...?”, “Noem hier: ..” ➤ Pas op met woorden als ‘exclusief’ en ‘niet’, daar lezen respondenten vaak overheen. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<p>4. Als gegevens moeten worden opgezocht voor het beantwoorden van de vraag, worden dan tips gegeven waar men deze informatie kan vinden?</p> <ul style="list-style-type: none"> ➤ Opzoektaken zijn lastig, maak het de respondent zo makkelijk mogelijk. 	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Checklist Vragenlijstontwikkeling Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling	Ja Min of meer Nee Niet van toepassing
--	---

E. Structuur van de vragenlijst

1. Kan de respondent of interviewer zich gemakkelijk in de vragenlijst oriënteren? ➤ Geef duidelijk aan waar de vragenlijst en ieder blok begint. ➤ Begin links boven.	□ □ □ □
2. Zijn de vragen in herkenbare thema's (blokken) gegroepeerd? ➤ Elk blok handelt over een afgerond thema. ➤ Geef elk blok een duidelijke aparte titel.	□ □ □ □
3. Is elk blok een afgeronde taak? ➤ Mensen werken van taak naar taak: structureer het werk in duidelijk afgeronde deeltaakjes. ➤ Neem de respondent zoveel mogelijk bij de hand.	□ □ □ □
4. Is de volgorde van de blokken logisch voor de respondent? ➤ Achtergrondkenmerken bij voorkeur als laatste vragen.	□ □ □ □
5. Is de route door de vragenlijst duidelijk? ➤ Zorg dat steeds duidelijk is op welke plaats in de vragenlijst men zich bevindt en hoe men verder moet gaan.	□ □ □ □
6. Zijn gevoelige vragen aan het eind van de vragenlijst geplaatst?	□ □ □ □
7. Zijn moeilijke vragen in het midden van de vragenlijst geplaatst?	□ □ □ □
8. Is de nummering van vragen, blokken en pagina's inzichtelijk en correct?	□ □ □ □

F. Vormgeving

1. Is de gekozen vormgeving functioneel? ➤ Iedere variatie in opmaak, ieder symbool et cetera moet een functie hebben. ➤ Gebruikte symbolen moeten duidelijk zijn en eventueel worden toegelicht.	□ □ □ □
2. Is de gekozen vormgeving consequent doorgevoerd? ➤ Opmaak, symbolen, regelafstand e.d. moet overal in de vragenlijsten dezelfde betekenis hebben.	□ □ □ □
3. Is de bladindeling van de vragenlijst inzichtelijk? ➤ Gebruik bij zelf in te vullen vragenlijsten bij voorkeur een 3-kolommen structuur: vraag-toelichting-antwoordvak ➤ Wees consequent: plaats dezelfde soort informatie steeds op dezelfde plek en dezelfde wijze	□ □ □ □

Checklist Vragenlijstontwikkeling Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling	Ja Min of meer Nee Niet van toepassing
--	---

4. Is kleurgebruik welovervogen toegepast? ➤ Kleuren alleen als aanvullende vormgeving gebruiken, nooit als enige vorm van onderscheid. ➤ Pas op met rood en groen, is voor veel mensen niet goed te onderscheiden.	□ □ □ □
5. Blijkt uit de vormgeving duidelijk welke onderdelen van de vragenlijst bij elkaar horen? ➤ Zorg dat steeds duidelijk is welk antwoordvakje bij welke vraag hoort. Breng waar nodig regelgeving aan. ➤ Zorg dat vragenblokken duidelijk herkenbaar zijn, bijvoorbeeld door witruimte tussen vraagblokken aan te brengen.	□ □ □ □
6. Is steeds duidelijk waar het antwoord vermeld moet worden? ➤ Gebruik steeds dezelfde invulvakjes of kaders voor het antwoord, bijvoorbeeld omkaderde, witte vakjes tegen een grijze achtergrond	□ □ □ □
7. Is er steeds voldoende ruimte om het antwoord in te vullen?	□ □ □ □
8. Is informatie die niet voor de respondent bestemd is op een aparte plaats vermeld of minder opvallend vormgegeven? ➤ Wat de respondent niet hoeft te weten, hoeft hij niet te zien.	□ □ □ □
9. Is de tekst 'scanbaar'? ➤ Gebruik bij voorkeur kolommen van maximaal 50 posities breed. ➤ Verdeel lange teksten in kleine alinea's.	□ □ □ □

G. Specifieke punten voor internet vragenlijsten

1. Is duidelijk waar het formulier gedownload of geopend kan worden? ➤ Zorg voor een kort en duidelijk web-adres. Mensen maken snel typfouten. ➤ Zorg voor een duidelijke en overzichtelijke webpagina	□ □ □ □
2. Is rekening gehouden met de scherminstellingen van de doelgroep? ➤ Men werkt niet altijd op de modernste apparatuur, ontwerp voor 800x600 pixels.	□ □ □ □
3. Is rekening gehouden met verschillende (versies van) browsers bij de doelgroep? ➤ Niet iedereen werkt met de nieuwste versie van internet	□ □ □ □

Checklist Vragenlijstontwikkeling	Ja Min of meer Nee Niet van toepassing
Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling	

explorer. Bepaal voor welke (versies) van browsers de vragenlijst bedoeld is en test de vragenlijst ook daarmee.	
4. Gebeuren er voor de respondent geen onverwachte dingen? ➤ De respondent moet het gevoel hebben 'in control' te zijn over het invullen van het formulier.	□ □ □ □
5. Ontvangt de respondent feedback bij iedere handeling? ➤ Iedere handeling moet een zichtbaar resultaat hebben.	□ □ □ □
6. Heeft de respondent steeds inzicht in de plaats waar hij is? ➤ Zorg dat duidelijk zichtbaar is welk invoerveld actief is. ➤ Zorg dat de respondent gemakkelijk kan navigeren. ➤ Zorg dat de respondent een indruk heeft van hoe ver hij is in de vragenlijst. ➤ Maak duidelijk welke schermen en vragen nog moeten worden ingevuld.	□ □ □ □
7. Is vastlopen in de vragenlijst voorkomen? ➤ Is er een terug-knop voor als mensen vastlopen of terug willen?	□ □ □ □
8. Is scrollen vermeden? ➤ Met name horizontaal scrollen is zeer ongewenst. ➤ Indien scrollen nodig is: geef aan dát er gescrold moet worden. Een scrollbar alleen is niet altijd voldoende.	□ □ □ □
9. Is bij antwoordvelden waar meer invulruimte is dan getoond wordt, aangegeven hoeveel posities het antwoord kan innemen?	□ □ □ □
10. Is waar nodig toelichting beschikbaar die tegelijk met vraag en antwoordmogelijkheden op het scherm getoond kan worden? ➤ Let op dat toelichting niet over de vraag heen valt.	□ □ □ □
11. Zijn foutmeldingen helder en duidelijk? ➤ Geef zo concreet mogelijk aan wat het probleem is ➤ Zorg dat de foutmelding niet over de desbetreffende vraag op het scherm verschijnt.	□ □ □ □
12. Indien de route door het formulier vastligt, is deze dan geautomatiseerd?	□ □ □ □
13. Geeft het programma waarschuwingen als de respondent een relevante vraag vergeet te beantwoorden of onwaarschijnlijke antwoorden geeft?	□ □ □ □
14. Is het rekenwerk geautomatiseerd?	□ □ □ □
15. Kan de respondent overzicht krijgen van de ingevulde gegevens?	□ □ □ □

<p>Checklist Vragenlijstontwikkeling</p> <p>Bijlage bij CBS-Methodenreeks Vragenlijstontwikkeling</p>	<p>Ja Min of meer Nee Niet van toepassing</p>
---	---

<p>16. Worden de gegevens tijdens het invullen opgeslagen? ➤ Maak het mogelijk de gegevens tussentijds op te slaan en later verder te gaan met invullen.</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>17. Kan de respondent het ingevulde formulier elektronisch archiveren?</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>18. Kan de respondent het ingevulde formulier uitprinten?</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>19. Wordt de respondent bewust gemaakt van het moment waarop hij de ingevulde gegevens verstuurt? ➤ Voorkom dat de respondent per ongeluk de gegevens verstuurt.</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>20. Krijgt de respondent na het versturen van een elektronische vragenlijst een ontvangstbevestiging?</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>21. Is rekening gehouden met gebruiksvriendelijkheid voor gehandicapten? Zie www.webrichtlijnen.nl voor richtlijnen voor de overheid op dit gebied.</p>	<p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>

Bijlage 3. Scoreformulier expertraadpleging vraagstelling

Francine Dehue

1. Interpretatie

- 1.1. TV: Toepasselijkheid vraag: De vraag is niet realistisch genoeg:
 - 1.1.1. Er wordt gevraagd naar hypothetische of gefingeerde gegevens (TVH)
 - 1.1.2. De respondent is de verkeerde informant (TVI)
 - 1.1.3. Er wordt gevraagd naar toekomstige gegevens (TVT)
 - 1.1.4. De gevraagde gegevens zijn onbekend (TVO)

- 1.2. SB: Strekking sleutelbegrippen: Het sleutelbegrip in de vraag is niet duidelijk:
 - 1.2.1. Het sleutelbegrip is onbekend (SBO)
 - 1.2.2. Het sleutelbegrip is vaag (SBV)
 - 1.2.3. Het sleutelbegrip heeft een dubbele betekenis (SBD)

- 1.3. SV: Strekking vraag: De bedoeling van de vraag is niet duidelijk:
 - 1.3.1. De betekenis van de vraag is onvoldoende concreet (SVO)
 - 1.3.2. Het referentie kader is onvoldoende gespecificeerd (SVR)
 - 1.3.3. De instructie bij de vraag is onduidelijk of misleidend (SVI)
 - 1.3.4. De vraag wordt als een dubbele vraag beschouwd (SVD)
 - 1.3.5. Er wordt een andere dan de bedoelde betekenis aan de vraag of de antwoordcategorieën toegekend (SVB)

2. Taakzwaarte

- 2.1. LC: Logische consistentie: De aansluiting van een vraag en een antwoord is niet consistent
 - 2.1.1. De vraag is in strijd met eerdere antwoorden (LCR)
 - 2.1.2. De vraag lokt een vraag-antwoord mismatch uit (LCA)
 - 2.1.3. Er is meer dan een enkel toegestaan antwoord mogelijk (LCM)
 - 2.1.4. Antwoordcategorieën ontbreken, overlappen, zijn irrelevant of van een andere dimensie dan de vraag suggereert (LCC)

- 2.1.5. De vraag lokt een onderbreking van het voorlezen uit (LCO)
 - 2.1.6. De formulering van de vraag is in strijd met de regels van de taal (LCT)
- 2.2. CV: Complexiteit vraag: De vraag is te ingewikkeld
 - 2.2.1. De vraag bevat teveel sleutelbegrippen (CVS)
 - 2.2.2. De vraag bevat teveel instructies (CVI)
 - 2.2.3. De vraag bevat een dubbele ontkenning (CVO)
 - 2.2.4. De vraag bevat impliciete veronderstellingen (CVV)
 - 2.3. CB: Cognitieve belasting: Het beantwoorden van de vraag vereist teveel informatieverwerking
 - 2.3.1. De antwoordcategorieën zijn te uitgebreid of te moeilijk (CBC)
 - 2.3.2. Het antwoordformaat is te gecompliceerd (CBA)
 - 2.3.3. Het is te moeilijk om tot een oordeel te komen (CBO)
 - 2.3.4. Het is te moeilijk om de gewenste informatie te herinneren (CBH)
 - 2.3.5. Het antwoord vereist gecompliceerde afleidingen (CBF)
- 3. Beleving**
- 3.1. AM: Affectie / motivatie: De vraag is een ongewenste vraag
 - 3.1.1. De relevantie is onduidelijk (AMR)
 - 3.1.2. De vraag is te persoonlijk (AMP)
 - 3.1.3. De vraag is niet compleet (AMI)
 - 3.2. NG: Normgevoeligheid: De vraag is niet neutraal en sturend
 - 3.2.1. De vraag bevat beladen begrippen (NGB)
 - 3.2.2. De vraag roept sociaal wenselijke antwoorden op (NGS)
 - 3.2.3. De antwoordcategorieën zijn normerend (NGA)

Bijlage 4. Methode Interviewerevaluatie

Wim Burgers

Doel

Het is gebruikelijk en ook wenselijk om persoons- en gezinsenquêtes te evalueren met interviewers. Dit levert informatie op over het verloop van het onderzoek, over (mogelijke) problemen die worden ervaren door interviewers en/of respondenten. Deze informatie vormt mede de basis op grond waarvan besloten kan worden tot aanpassingen, verbeteringen van de opzet, de benaderingswijze en/of de vragenlijst.

In de afgelopen jaren is een methode ontwikkeld die op een efficiënte manier in een kort tijdsbestek met de inzet van een beperkt aantal interviewers relevante en gewenste informatie oplevert over een onderzoek. Het resultaat van een op deze wijze uitgevoerde evaluatie is een overzicht van knelpunten gewogen naar het belang dat door de interviewers aan de afzonderlijke punten wordt toegekend.

Werkwijze

Vorbereiding

Uitgangspunt van de evaluatiemethode is enerzijds dat de deelnemende interviewers in principe zo weinig mogelijk in een bepaalde (denk) richting worden gestuurd en anderzijds dat iedere deelnemende interviewer een bijdrage levert aan de evaluatie. Dit betekent onder meer dat de onderwerpen die CBS-medewerkers ter evaluatie willen voorleggen aan de interviewers niet eerder aan de orde worden gesteld – voor zover ze niet door de interviewers naar voren zijn gebracht – dan nadat de interviewers hun zegje hebben gedaan. De inventarisatie van de onderwerpen die volgens de CBS-betrokkenen (onderzoeker, projectleider, enquêtebegeleider) in ieder geval tijdens de evaluatie aan bod moeten komen vindt vooraf plaats.

Voor een evaluatie worden circa 15 interviewers uitgenodigd. De ervaring heeft geleerd dat, mits de groep interviewers representatief is naar regio en stedelijkheid, de resultaten van de evaluatie met circa 15 interviewers representatief zijn voor het enquêtecorps.

Inventariseren

Aan het begin van de bijeenkomst wordt aan de interviewers uitgelegd wat het doel is van de bijeenkomst en wat en hoe er geëvalueerd gaat worden. Afhankelijk van het onderzoek kan de evaluatie meer of minder aspecten van het veldwerk betreffen: benaderingsstrategie, vragenlijst, enquêtemateriaal, logistiek et cetera. Vervolgens wordt de groep verdeeld in tweetallen. Ieder tweetal inventariseert welke problemen, klachten, onduidelijkheden, zij bij de uitvoering van het veldwerk zijn tegengekomen. Ook

worden zij uitgenodigd om zaken die positief opvielen te vermelden. Ieder tweetal noteert alle punten die worden genoemd in de vorm van een stelling op een (groot) vel papier. Nadat de tweetallen klaar zijn met de inventarisatie worden de genoteerde stellingen opgehangen.

Toelichten

De stellingen worden een voor een met de gehele groep besproken. Alle aanwezigen mogen om toelichting vragen aan de interviewer(s) die een stelling heeft (hebben) opgeschreven. Belangrijk daarbij is dat er geen discussie over de stellingen plaatsvindt om te voorkomen dat de (juistheid van een) stelling door de interviewer(s) van wie de stelling afkomstig is moet worden verdedigd. Daarnaast kan iedere interviewer later via een score aangeven of zij/hij het met een stelling eens is of niet. De bespreking van de stellingen heeft ten doel om eventuele onduidelijkheden te verhelderen zodat voor alle deelnemers duidelijk is wat wordt bedoeld met de stelling. De stelling wordt zo nodig anders geformuleerd zodat er geen misverstanden kunnen ontstaan. Daarnaast biedt de bespreking de mogelijkheid om het aantal stellingen te verkleinen door (vrijwel) eensluidende stellingen te bundelen.

Zo nodig kan een aanwezige onderzoeker of projectleider achtergrondinformatie geven over het waarom van een bepaalde vraag of blok vragen in de vragenlijst. Maar ook hiervoor geldt dat dit niet mag tot een discussie leiden. Resultaat van de bespreking is een lijst met eenduidige stellingen. Deze stellingen worden van een volgnummer voorzien. Aan het eind van deze fase wordt er gekeken of de onderwerpen die van CBS-, of klantzijde vooraf zijn geïnventariseerd door de interviewers naar voren zijn gebracht. Stellingen die worden gemist worden alsnog toegevoegd aan de totale lijst. Uiteraard worden deze stellingen desgewenst toegelicht.

Scoren

Tenslotte worden alle interviewers gevraagd op een scoreformulier een cijfer toe te kennen aan elke stelling. Door de interviewers te laten scoren is het mogelijk om de knelpunten gewogen weer te geven. Interviewers kennen op de volgende manier hun scores toe:

Er wordt gescoord met een cijfer van 1 tot en met 10, waarbij een 1 betekent “helemaal oneens met de stelling” en een 10 “helemaal eens met de stelling”. Interviewers worden vooraf gevraagd om goed te discrimineren in hun score, dus niet alleen enen en tien toe te kennen, maar juist voor zichzelf na te gaan in welke mate zij het eens, of oneens zijn met een bepaalde stelling. Indien zij de in een stelling beschreven situatie nooit hebben meegemaakt, of er geen mening over hebben, scoren zijn een nul. (Sommige problemen zijn gerelateerd aan een bepaald gebied b.v. de grote stad en niet alle interviewers zullen zich dan in dat probleem kunnen herkennen.)

Verwerking

Als er op alle stellingen is gescoord is de sessie met de interviewers klaar. Nu kan er met het verwerken van de scores worden begonnen.

Alle stellingen worden op volgnummer in een tabel gezet en vervolgens wordt per interviewer de scores op de stellingen ingevoerd. Als alle scores zijn ingevoerd wordt de totaalscore en de gemiddelde score per stelling berekend.

Om de scores vergelijkbaar te laten zijn moeten de scores op de stellingen allen dezelfde kant opwijzen. Hiervoor is het soms noodzakelijk bij het verwerken van de scores 'om te waarden'. Normaal gesproken betekent dit dat de scores van de positief geformuleerde stellingen (b.v. "Dit is een prettige vragenlijst om mee te werken") 'negatief' worden omgewaardeerd. Hierbij wordt de score van '1' een score van '10', een score '2' wordt '9' enz

Tenslotte wordt een op- of aflopende sortering op de gemiddelde score doorgevoerd, zodat er een overzicht wordt verkregen dat inzicht geeft in de door de interviewers aangegeven belangrijkheid van een stelling.



Scoreformulier interviewerevaluatie

Stelling	Score		Stelling	Score
1.			26.	
2.			27.	
3.			28.	
4.			29.	
5.			30.	
6.			31.	
7.			32.	
8.			33.	
9.			34.	
10.			35.	
11.			36.	
12.			37.	
13.			38.	
14.			39.	
15.			40.	
16.			41.	
17.			42.	
18.			43.	
19.			44.	
20.			45.	
21.			46.	
22.			47.	
23.			48.	
24.			49.	
25.			50.	

Bijlage 5a. Voorbeeld observatieformulier CAPI / CATI interview

Interviewer..... Datum..... Tijd uur Respondentnummer #..... Versie															
Opmerkingen															
ENQ	Vraag loopt niet														
	Herhalen antwoord														
	Herhalen vraag														
RESPONDENT	Antwoord	Reden weigering													
		Noemt meerdere antwcat													
		Te veel antw. cats.													
		Te lange antw. cats													
		Begrijpt begrip niet													
		Kan ant niet kwijt													
	Vraag	Mist toelichting													
		Te lang (R raakt draad kwijt)													
		Begrijpt begrip niet													
		Begrijpt vraag niet													
EHIS 2008	VRAAG														

Bijlage 5b. Voorbeeld observatieformulier CAWI vragenlijst

Observatieformulier Vragenlabtest Mixed-Mode 2007

Interviewer	
Datum	
Tijd	
Respondent	
Vragenlijstversie	

Blok	Opmerkingen
Opening	
Hoe is eerste reactie van R op het scherm?	
Leest R de inleiding?	
Kan R door naar de volgende vraag?	
Overige opmerkingen	
Achtergrondvragen	
Kan R de vragen invullen?	
Wat noemt R spontaan als duidelijk of onduidelijk?	
Mist R toelichtingen?	
Mist R antwoordcategorieën?	
Overige opmerkingen	
Werk en opleiding	
Kan R de vragen invullen? Kan R zijn bedrijf en beroep/functie kwijt in de categorieën?	
Wat noemt R spontaan als duidelijk of onduidelijk?	
Mist R toelichtingen?	
Mist R antwoordcategorieën?	
Hoe zeker is R over de antwoorden over het verleden?	
Hoe zeker is R over de antwoorden over de partner?	
Overige opmerkingen	

Bijlage 6. Standaardprotocol bedrijfsbezoeken Vragenlijstevaluatie

Doel van het protocol en het bezoek

Het protocol is een leidraad van wat er idealiter aan bod zou moeten komen bij een vragenlijstevaluatie bij een bedrijf. De praktijk is - zeker bij tests op locatie - weerbarstig en niet alles zal altijd aan bod kunnen komen. Het belangrijkste doel van de evaluatiebezoek is inzicht te krijgen in de volgende zaken:

- Waar worden fouten gemaakt bij de invullen?
- Waar ontstaat hoge responslast?
- Waarom ontstaan deze problemen en wat is er eventueel aan te doen?

Dit standaardprotocol dient nader aangevuld te worden met specifieke observatiepunten en vragen voor de te testen vragenlijst.

Twee soorten bezoeken: observatiebezoeken en nagesprekken

We onderscheiden twee soorten testbezoeken: observatiebezoeken en nagesprekken. Bij observatiebezoeken vult de respondent de vragenlijst in tijdens het bezoek en wordt daarbij geobserveerd. Het doel is dat de respondent zoveel mogelijk te werk gaat zoals hij dat zou doen als er niemand van het CBS aanwezig zou zijn - zo moet men de vragenlijst immers straks na implementatie ook kunnen invullen. Na afloop wordt het invulproces nabesproken en kunnen nog aanvullende vragen worden gesteld. Bij nagesprekken heeft de respondent al voorafgaand aan het bezoek de vragenlijst ingevuld en wordt het invulproces gereconstrueerd in het nagesprek.

Het voordeel van een observatiebezoek is dat je heel precies kan zien waar problemen ontstaan. Gedetailleerde informatie zoals waar precies een foutmelding komt in een elektronische vragenlijst of dat een vraag aanvankelijk wordt overgeslagen is na afloop van het invullen haast niet meer te achterhalen. Daarom is zeker voor het onderzoeken van de gebruiksvriendelijkheid van het ontwerp van een vragenlijst observatie een goede methode. Het nadeel van observeren is echter dat de aanwezigheid van de onderzoekers het normale invulproces kan verstoren. Respondenten kunnen zich opgejaagd voelen door de observatie en daardoor zaken gaan afraffelen, juist heel erg hun best gaan doen of niet meer zelf op zoek gaan naar hulpinformatie maar de onderzoekers als hulp gaan gebruiken. Bij nagesprekken heb je daarom een meer realistisch invulproces, maar dan wel minder gedetailleerd in beeld gebracht. Meestal is het goed om binnen een test zowel nagesprekken als observaties te doen.

Eerst kijken en luisteren, dan pas helpen en uitleggen

Bij de tests van bedrijfsenquêtes geldt, anders dan bij persoonsenquêtes, dat de proefrespondent meestal al ervaring heeft met het invullen van CBS enquêtes en in de toekomst de testvragenlijst en andere vragenlijsten van het CBS zal invullen. Er moet bij het testbezoek daarom ook ruimte zijn om vragen van de respondent te beantwoorden en eventuele invulfouten die bij de test worden vastgesteld te corrigeren. Echter: het hoofddoel van de test is te kijken hoe de vragenlijst werkt. Daarvoor is het van belang dat de respondent de vragenlijst bij de test zoveel mogelijk gewoon invult zoals hij dat normaal gesproken zou doen. Na afloop daarvan is er gelegenheid om vragen die opkomen tijdens het invullen te beantwoorden. Alleen als iemand echt helemaal vastloopt en bijvoorbeeld de helpdesk zou gaan bellen bij een reguliere enquête, mogen de onderzoekers helpen. Dit verloop van de test moet voorafgaand aan de test goed worden uitgelegd aan de respondent.

Standaard Protocol Bedrijfsbezoek evaluatie Testvragenlijst Versie januari 2010	
Wat	Doel & aandachtspunten
1. Introductie	Voorstellen bezoekers en uitleggen doel en opbouw van bezoek <ul style="list-style-type: none">▪ Doel bezoek: 1. evalueren van het gebruiksgemak vernieuwde vragenlijst 2. indien van toepassing: helpen /corrigeren invulling▪ Structuur bezoek:<ul style="list-style-type: none">○ algemene vragen vooraf○ invullen vragenlijst (bij observatiebezoek) OF (bij nagesprek) zo precies mogelijk in kaart brengen hoe het invullen is verlopen○ vragen over het invulproces○ eventueel uitleg van inhoudelijke zaken en ondersteuning bij de invulling▪ Toestemming voor opname vastleggen
2. Algemene vragen vooraf	Algemene omgang met vragenlijsten van het CBS en de Testvragenlijst in het bijzonder <ul style="list-style-type: none">▪ Wat voor soort bedrijf is dit?▪ Wat is de functie van de contact persoon (hierna: cp)?▪ Is de Testvragenlijst al vaker ingevuld? (uiteraard alleen indien deze al in het veld is)▪ Welke vragenlijsten vult het bedrijf nog meer in voor het CBS, wie doet dat?

	<ul style="list-style-type: none"> ▪ Hoe is deze vragenlijst in het takenpakket van cp gekomen? ▪ Hoe wordt in het bedrijf bepaald óf en wanneer de testvragenlijst wordt ingevuld? ▪ Wie is verantwoordelijk voor het invullen van de testvragenlijst? Wie zijn er bij betrokken? Indien meerdere personen: wat is de taakverdeling? ▪ Worden speciaal met het oog op de vragenlijsten specifieke gegevens bijgehouden? Zo ja, wat dan en hoeveel tijd kost dat? ▪ Wat vindt cp van begeleidend materiaal (brief, e-mail, eventuele folder et cetera) ▪ Heeft cp contact gehad met het CBS over de Testvragenlijst? Zo ja, waarover en hoe is dat gegaan?
<p>3. Observeren of reconstrueren invulproces</p>	<p>Vastleggen hoe cp de vragenlijst invult of heeft ingevuld</p> <p>1. Beginnen vragenlijst</p> <ul style="list-style-type: none"> ▪ Hoe ging het beginnen met de vragenlijst (kon cp bijvoorbeeld makkelijk inloggen / downloaden / beginnen op juiste pagina e.d.) ▪ Worden aanvullende materialen zoals brief, folder geraadpleegd? ▪ Worden introducerende teksten / schermen gelezen? ▪ Wordt eerst gebladerd door de vragenlijst of direct begonnen met invullen? <p>2. Invullen van de vragenlijst</p> <ul style="list-style-type: none"> ▪ Bekijk per vragenlijstitem hoe cp tot het antwoord komt: Noteer bij welke items er fouten worden gemaakt (en indien mogelijk ook: wat er misgaat en waarom) en welke items een hoge responslast veroorzaken (ook hier graag aangeven waarom). ▪ Zijn er tijdens het invullen waarschuwingen en foutmeldingen gekomen van het programma? Wat dan? Hoe reageerde cp hierop? ▪ Specifieke aandachtspunten voor invulproces Testvragenlijst hier toevoegen. <p>3. Verzenden en afsluiten</p> <ul style="list-style-type: none"> ▪ <i>NB: alleen elektronische vragenlijsten laten verzenden indien de ingevulde vragenlijst nog kan worden geopend. Anders eerst nabespreken en fouten corrigeren bij observatiebezoek.</i> ▪ Hoe ging het verzenden? ▪ Hoe lang heeft het invullen van de vragenlijst geduurd? ▪ Is cp bij het invullen zodanig vastgelopen dat hij geholpen moest worden? Zo ja: wanneer en waarom is dat gebeurd?
<p>4. Mening cp over de vragenlijst</p>	<p>Mening van cp over vragenlijst in kaart brengen en navragen specifieke aandachtspunten</p> <ul style="list-style-type: none"> ▪ Wat vindt cp van de inhoud en het gebruiksgemak van deze vragenlijst, zowel positief als negatief? ▪ Als er iets veranderd moet worden aan de testvragenlijst wat zou cp het liefste anders zien?

	<ul style="list-style-type: none"> ▪ Voor zover nog niet aan de orde gekomen, wat vindt cp specifiek van: <ul style="list-style-type: none"> ○ benaderingswijze (de brief, de folder) ○ downloaden en installeren? ○ navigeren (index, bladertoetsen en akkoordknop). ○ Layout algemeen (bijv. schermindeling, kleurgebruik, leesbaarheid letters) ○ toelichtingen: plaats en inhoud ○ menufuncties ○ het verzenden? <p>Dit lijstje aanvullen met specifieke aandachtspunten voor de testvragenlijsten</p> ▪ Vraag om een print / kopie / screendumps van de ingevulde vragenlijst (indien mogelijk ook een elektronisch exemplaar).
5. Uitleggen en corrigeren formulier	<p>Hier kan ingegaan worden op inhoudelijke problemen en kan geholpen worden met het invullen</p> <ul style="list-style-type: none"> ▪ Waar nodig posten corrigeren ▪ CP antwoord geven op eerder gestelde vragen en waar nodig uitleg geven.
6. Afsluitende vragen	<p>Openstaande punten navragen en gesprek afronden.</p> <ul style="list-style-type: none"> ▪ Maakt cp zelf wel eens gebruik van publicaties van het CBS? Zo ja, welke dan? En maken andere mensen in het bedrijf wel eens gebruik van CBS cijfers? ▪ Heeft cp nog vragen of opmerkingen die nog niet aan de orde zijn gekomen?