

De kwaliteit van schriftelijke vragen over alcoholconsumptie

Vergelijking van een bureauonderzoek met een veldonderzoek

Harrie Jansen en Tony Hak

1 Inleiding

Surveyvragenlijsten stellen vaak hoge eisen aan de motivatie en aan de cognitieve en verbale vermogens van respondenten, hetzij door de aard van de gevraagde informatie (normativiteit, complexiteit), hetzij door de manier waarop die informatie wordt opgevraagd (formulering van de vraag en/of antwoordcategorieën). Dit leidt tot problemen met betrekking tot de datakwaliteit: geldigheid, volledigheid (item nonrespons), betrouwbaarheid, nauwkeurigheid.

Er zijn diverse methoden in gebruik om dergelijke problemen in vragenlijsten te onderzoeken. Traditioneel gebruikelijk is de beoordeling van de (concept)vragenlijst door collega's en externe deskundigen, het *expert review*. Voor zulk bureauonderzoek zijn wel handleidingen en beoordelingsformulieren (bijv. Sudman, Bradburn en Schwarz 1996) ontwikkeld, maar er bestaat nauwelijks publieke documentatie van de praktijk van die beoordelingen, laat staan onderzoek naar de kwaliteit ervan. Datzelfde gold lange tijd voor het pretesten van vragenlijsten in proefinterviews; het is al heel lang gebruikelijk om dit te doen, maar het is tot voor kort weinig publiek gedocumenteerd en onderzocht. Er werd op vertrouwd dat je na een aantal goed gespreide proefinterviews voldoende informatie hebt om een vragenlijst aan te passen en dat er dan eigenlijk geen problemen meer zijn. Het stellen van een vraag werd toch vaak gezien als het aanwijzen van een laatje, waar de respondent dan alleen maar de tekst uit hoeft te pakken en voor te lezen. Daar is echter de afgelopen decennia verandering in gekomen; de datakwaliteit van surveys is fundamenteel ter discussie gesteld (vgl. De Heer, De Leeuw en Van der Zouwen 1999 voor een geschiedenis van het survey) en tegelijk zijn er meer sophisticated methoden ontwikkeld voor het pretesten van vragenlijsten. Daarbij speelt de toepassing van inzichten uit de cognitieve psychologie een belangrijke rol. Sinds kort wordt ook de kwaliteit (geldigheid en betrouwbaarheid) van de verschillende pretestmethoden zelf onderzocht (bijv. Willis en Schechter 1997; Campanelli 1997). Hieraan willen wij in dit artikel een bijdrage leveren door een vergelijking van twee testmethoden toegepast op dezelfde set surveyvragen, te weten een bureauonderzoek in de vorm van een expert review en een veldonderzoek in de vorm van het Drie Stappen Test Interview

(DSTI). Bovendien stellen wij met die driestapsprocedure een uitbreiding voor van de gebruikelijke cognitieve pretestmethode. De vragen die getest worden, hebben betrekking op alcoholconsumptie en zijn afkomstig uit de survey 'Risikante leefgewoonten onder de Rotterdamse bevolking' (Van de Goor e.a. 1996; Bongers 1998).

2 Het probleem van onderrapportage in alcohol surveys

Doordat de verkoop van alcoholhoudende dranken op landelijk niveau geregistreerd wordt in verband met de heffing van accijns, is er bij surveys naar alcoholconsumptie een redelijk 'hard' extern validiteitscriterium beschikbaar voor de totale hoeveelheid drank die door de respondenten wordt opgegeven, namelijk het verkoopcijfer. Gemiddeld blijkt dan dat bij surveys bijna de helft van de verkochte alcohol niet wordt gerapporteerd. Deze bevinding heeft ertoe geleid dat onderrapportage sinds lange tijd het centrale thema is in de methodologische discussie over surveyonderzoek naar alcoholconsumptie (vgl. Midanik 1982; Garretsen 1983; Lemmens 1987). Onderrapportage is overigens niet specifiek voor alcoholgebruik; het is een algemeen verschijnsel bij onderzoek naar min of meer routinematig gedrag (vgl. Mingay e.a. 1994). Wel is de onderschatting bij alcohol groter dan bij bijvoorbeeld frisdrank. Daar is de onderrapportage ongeveer 25% volgens TNO-onderzoeker Hendriks (geciteerd in Reinhard en Rood-Bakker 1998).

Leugens en vergissingen. Aangezien er op individu-niveau geen extern geldigheids criterium beschikbaar is zoals de verkoopcijfers op landelijk populatieniveau, is het de vraag of en hoe de kwaliteit van zelfrapportages in surveys kan worden vastgesteld (en verbeterd). Dat hangt mede af van het soort rapportagefouten. Veelal worden twee soorten onderscheiden: 'opzettelijke misrapportage' (uit schaamte, opschepperij of wat ook) en 'onbedoelde herinneringsfouten' (door vergeten van drinkgebeurtenissen of door misrekening van tijdsspannen waarover het gedrag wordt gevraagd) (Lemmens 1991; Bongers 1998). Empirisch zijn deze twee soorten fouten moeilijk te onderscheiden. Maar het onderscheid is methodologisch wel van belang, omdat de bestrijding deels op verschillende momenten in het surveyproces moet plaatsvinden en met verschillende middelen. Met onder meer Midanik (1988) beschouwen wij de dataproductie niet als een eenzijdige activiteit van de respondent, maar als een interactiviteit tussen onderzoeker en respondent. In dat proces spelen diverse factoren een rol, zoals de afnamesituatie, de biografische context van de respondent, de eisen die aan verbale en herinneringsvermogens worden gesteld en de interpretatie van de vragen. De onderzoeker (in ruime zin) heeft tot taak de respondent in staat te stellen tot valide rapportage van 'de' feiten; dat is de maatstaf voor de kwaliteit van het onderzoek c.q. de vragenlijst in kwestie. Dit vereist twee dingen:

- a De respondent moet zich identificeren met de doelstellingen van de onderzoeker en dus gemotiveerd worden om de vragen zo goed mogelijk te beantwoorden. Om deze houding te bevorderen, is vooral ideologische activiteit en retorische vaardigheid van de onderzoeker nodig. Dit geldt bijvoorbeeld voor de modellering van de onderzoeksrelatie in de aanbiedingsbrief en in de formulering van vragen. Onder verwijzing naar Goffman spreekt Harkness (1996) in dit verband van 'the (re)presentation of self in everyday questionnaires'. Empirisch onderzoek naar 'interpretatie en waardering van aanschrijfbrieven' is in Nederland gedaan door Luppés (1993) bij het CBS. Overigens moet de motivatie – als deze er eenmaal is – ook onderhouden worden tot en met (in het geval van de schriftelijke enquête) het posten van de enveloppe met de ingevulde vragenlijst. Dat stelt eisen aan aantal, levendigheid en vooral relevantie van vragen.
- b De respondent moet begrijpen wat de onderzoeker wil weten en in staat gesteld worden om zijn/haar waarheid op een voor de onderzoeker ondubbelzinnige wijze kenbaar te maken. Hier gaat het vooral om de cognitieve eisen die door vraagwijze (volgorde, formulering) en antwoordcategorisering aan de respondent worden gesteld.

3 Onderzoek naar de kwaliteit van vragen(lijsten)

De kwaliteit van een vraag (of vragenlijst) kan uiteindelijk alleen worden afgemeten aan de kwaliteit (validiteit, volledigheid en nauwkeurigheid) van de antwoorden. Idealiter zou men daarom de antwoorden moeten kunnen vergelijken met 'harde' informatie uit andere bronnen.

Zoals hierboven vermeld, is dat in het alcoholonderzoek op populatieniveau gedaan door vergelijking van enquêteresultaten met verkoopcijfers. Ook is wel gewerkt met rapportage door partners als vergelijkingsmateriaal voor zelfrapportages (Garretsen en Knibbe 1983; Knibbe 1982; Garretsen 1983; Lemmens 1991; Bongers e.a. 1999). Het probleem daarbij is, dat partners natuurlijk (net als de respondenten zelf) hun eigen motieven en herinneringsproblemen hebben, waardoor moeilijk hard te maken is dat de partnerrapportage beter is dan de zelfrapportage (vgl. Mingay e.a. 1994 en Sudman e.a. 1994).

In feite is men vrijwel altijd aangewezen op indirecte toetsing. Globaal genomen zijn hiervoor twee benaderingen in zwang: experimentele vergelijking van verschillende vraagwijzen en cognitief onderzoek.

Experimentele vergelijking van vraagwijzen. Er zijn diverse onderzoeken uitgevoerd waarin verschillende vraagwijzen naar alcoholconsumptie experimenteel met elkaar vergeleken zijn: Blair e.a. (1977), Lemmens (1987), Waterson en Murray-Lyon (1989), Lemmens, Tan en Knibbe (1992), Midanik (1994), Single en

Wortley (1994). In al deze gevallen wordt ervan uit gegaan dat misrapportage vooral een gevolg is van vergeten en dat bijgevolg een hogere gerapporteerde consumptie meer valide is dan een lagere. De resultaten van deze onderzoeken samenvattend, kan worden gezegd dat hogere rapportages van alcoholconsumptie bereikt worden door: a) toepassing van *open* vragen en b) *specificatie* van de vragen naar: dranksoort, drinksituatie, type drinkdagen (Midanik 1994) enzo-voorts. Hierbij moeten we echter opmerken dat genoemd onderzoek voor een deel betrekking heeft op interviews. Voor schriftelijke ondervraging lijken de cognitieve voordelen van specificatie voorbij een bepaald punt te zullen worden opgeheven door motivationele nadelen. Vergaande specificatie leidt tot lange reeksen vrijwel identieke vragen, waardoor motivatieproblemen ontstaan die in een schriftelijke enquête moeilijk te hanteren zijn, met als gevolg hoge item nonrespons en onnauwkeurigheid.

Naast de inhoudelijke vergelijking van antwoorden worden ook vergelijking van de item (non)respons en van antwoordpatronen gehanteerd voor de beoordeling van de relatieve kwaliteit van vragen (Esposito en Rothgeb 1997; Campanelli 1997).

Cognitief onderzoek. Vanaf de jaren zeventig zijn kwalitatieve onderzoeksmethoden uit de cognitieve psychologie toegepast voor de verzameling van kennis over de werkwijze van respondenten bij herinneringsvragen (Schwarz en Sudman 1994). De belangrijkste zijn a) het cognitieve interview (Loftus, Fienberg en Tanur 1985; Koehnken e.a. 1995), waarbij respondenten door middel van gedetailleerd doorvragen worden geholpen om nauwkeurige informatie te produceren en b) de protocolanalyse van teksten die verkregen worden door respondenten hardop te laten denken (*think aloud*) bij het invullen van vragen. Beide methoden worden veel toegepast bij het testen en ontwikkelen van vragenlijsten in vragenlaboratoria van onderzoeksinstituten (vgl. Bergman 1995; Blair en Presser 1994; Lindström 1996; Snijders 1998). Bij pretests in het veld wordt cognitieve informatie verzameld door middel van observatie tijdens het invullen, specifiek doorvragen en interviews achteraf. Daarnaast kan ook *expert analysis*, dat wil zeggen kritische beoordeling van vragen door ervaren onderzoekers, worden gezien als een vorm van cognitief onderzoek. Dit soort onderzoek wordt heel vaak als vanzelfsprekend en op basis van 'geschoolde intuïtie' uitgevoerd, maar slechts zelden en detail gerapporteerd.

Campanelli (1997) vergeleek een achttal technieken voor de toetsing van interviewvragen, die grotendeels ook van toepassing zijn op schriftelijke vragen. Haar voorzichtige, want op slechts drie studies gebaseerde, conclusie luidt dat cognitieve interviews en expert analyse het meest productief zijn (Campanelli 1997: 10). Aangezien verschillende methoden tot verschillende resultaten leiden zonder dat bij voorbaat vaststaat welke methode beter is, bevelen Esposito en Rothgeb

(1997) aan om altijd meer methodes tegelijk te gebruiken voor de identificatie van problemen. Deze aanbeveling en Campanelli's conclusie ondersteunen de relevantie van een vergelijking van resultaten uit een expert analyse met die uit cognitieve interviews.

4 Identificatie en classificatie van problemen met behulp van cognitieve interviews

Voor de identificatie en classificatie van problemen kunnen verschillende theoretische ingangen worden gekozen. Midanik en Hines (1991) baseerden zich op de theorie van de *cognitieve strategieën*. Deze houdt in dat bepaalde personen specifieke manieren (strategieën) hanteren om een cognitieve opdracht (het beantwoorden van een vraag) uit te voeren en dat anderzijds bepaalde vragen specifieke strategieën vereisen. Dat betekent dat er kwaliteitsproblemen ontstaan in gevallen waarin een respondent een andere strategie hanteert dan die welke nodig is voor een valide beantwoording van de vraag. Om deze theorie te toetsen pasten zij 'hardop denken' toe op vier vragen naar alcoholconsumptie: drinkfrequentie, gemiddelde consumptie per drinkgelegenheid, frequentie van vijf of meer consumpties per gelegenheid en frequentie van dronkenschap. Zij onderscheidden een vijftal cognitieve strategieën voor de classificatie van de antwoordprocessen. Het bleek dat de strategie van *anchoring and adjustment* (initiële herinnering gevolgd door toets op redelijkheid en aanpassing) het meest werd toegepast. Maar verder bleek dat er zowel per respondent als per vraag veel variatie was in toegepaste (combinaties van) strategieën. De keuze van een cognitieve strategie lijkt niet systematisch te worden bepaald door de formulering van de vraag. Deze benadering lijkt daarmee niet erg productief te zijn voor de identificatie en interpretatie van problemen in een vraag.

Conrad en Blair (1998) ontwierpen een schema van *respondent problems* dat nauwer aansluit bij de gangbare praktijk van het cognitieve interviewen. Hierbij wordt uitgegaan van de veronderstelling dat de respondent de vraag louter opvat als een zoek- en rapportageopdracht en deze zo nauwkeurig mogelijk tracht uit te voeren. Zij onderscheiden twee dimensies in de respondentproblemen:

- a de drie fasen in het responsproces: het begrijpen van de zoekopdracht, het uitvoeren van de zoekopdracht ('graven' in het geheugen) en de weergave van het zoekresultaat in het aangeboden antwoordformat;
- b typen problemen die in elk van die drie fasen kunnen optreden: lexicaal (woordbegrip), temporeel (tijdsperiode waarop de vraag betrekking heeft), logisch (betekenis van verbindingswoorden: 'behalve', 'en' 'of'), rekenkundig (bij vragen naar frequenties en hoeveelheden) en afgrenzing (reikwijdte van begrippen).

Een voordeel van dit schema is, dat het een vrij concreet overzicht geeft van de meest voorkomende soorten cognitieve moeilijkheden van respondenten. Daardoor is het gemakkelijker hanteerbaar als attenderingslijst bij de beoordeling van vragenlijsten en het opzetten van itemlijsten voor cognitieve interviews. Wat ontbreekt zijn normatieve en motivationele problemen die – zeker bij moreel beladen onderwerpen – een belangrijke bron van vertekening kunnen zijn.

5 Probleemstelling en onderzoeksoptzet

In het kader van het project ter optimalisering van de schriftelijke alcoholenquête hebben wij twee onderzoeken uitgevoerd om problemen op te sporen in de vragenlijst: eerst een bureauonderzoek (expert analysis) en daarna een cognitief veldonderzoek. De vraag waar het hier om gaat is nu: brengt het cognitieve veldonderzoek andere (en meer) problemen aan het licht dan de expert analyse op het bureau en zo ja, wat is dan de ‘winst’ van het relatief tijdrovende cognitieve veldonderzoek?

Naar ons weten is er niet eerder onderzoek gepubliceerd met deze vraagstelling. Er is veel literatuur over cognitief onderzoek (dat overigens vrijwel altijd in het laboratorium wordt uitgevoerd), maar de enige ons bekende vergelijking met expert analysis is het geciteerde artikel van Campanelli (1997) waarin de methodes kwantitatief met elkaar vergeleken worden op een schaal voor bruikbaarheid. Er is niet onderzocht hoeveel van de problemen die met de ene methode worden ‘gevonden’, met de andere worden gemist en omgekeerd, en ook niet of er een systematisch verschil is in de aard van de gevonden problemen.

De onderzochte vragen. De kern van de set vragen die wij voor ons onderzoek selecteerden, wordt gevormd door de zogenaamde *quantity-frequency* vragen, namelijk: hoe vaak drinkt u en hoeveel drinkt u per keer? Daaraan zijn toegevoegd:

- a Een inleidende vraag naar de *soort drank* die men meestal drinkt, met als antwoordcategorieën een reeks alcoholische dranken. Deze vraag heeft vooral de functie om het begrip ‘alcoholische drank’ te operationaliseren voor de respondent en om niet-drinkers te identificeren.
- b Een specifieke vraag naar de frequentie van *zes of meer glazen per dag*; deze dient om te bepalen of volgens onder onderzoekers gangbare criteria de respondent wel of niet een ‘excessieve drinker’ is.
- c Twee vragen naar de *variatie* in het drankgebruik over de afgelopen tien jaar (periodes van meer of minder drinken dan op dit moment).

Het bureauonderzoek. Het eerste deel van het onderzoek bestond uit een nauwkeurige lezing van de vragen met als aandachtspunten: welke ‘denkstappen’ moe-

ten de respondenten doen om de verlangde informatie te produceren en in hoeverre worden zij daarbij geholpen of juist gehinderd door de formulering van de vraag en de antwoordcategorieën? Daarbij lieten we ons mede informeren door vergelijking van deze vragen met soortgelijke vragen uit andere surveys (Garretsen en Knibbe 1983; CBS 1989; IMGZ 1991; GGD-en Groningen 1994) en interviews met de auteurs van de vragenlijst (vgl. Hak en Jansen 1997).

Het veldonderzoek. De identificatie van 'problemen' beruiste bij het bureau-onderzoek vooral op de door ons ingeschatte moeilijkheidsgraad van cognitieve bewerkingen die door respondenten zouden moeten worden uitgevoerd bij de beantwoording van de vragen. De 'papieren' identificatie van zulke 'logische' of 'cognitieve' problemen betekent echter nog niet automatisch,

- 1 dat er voor respondenten ook (deze of andere) 'problemen' zijn en, zo ja,
- 2 dat die van praktisch belang zijn voor de onderzoeksbevindingen.

Om dit vast te stellen werd een empirisch veldonderzoek opgezet, bestaande uit drie stappen, het Drie Stappen Test Interview (DSTI, zie bijlage 2). Eerst werd de respondenten verzocht om de vragenlijst hardop denkend in te vullen (*think aloud*); daarna werd een zogenaamd *cognitief interview* afgenomen, waarin mondeling nagevraagd werd hoe de respondent te werk was gegaan bij het invullen van de vragenlijst en welke moeilijkheden zich daarbij hadden voorgedaan. Als derde stap werd een intensief *inhoudelijk interview* gehouden over de drinkgewoonten van de respondent. Dit interview diende om vergelijkingsmateriaal te genereren voor de vaststelling van de praktische betekenis van de schriftelijke antwoordproblemen – in termen van misrapportage.

De verzadigingsprocedure. Dit onderzoek is niet bedoeld om een frequentieverdeling van invulproblemen in een bepaalde populatie te produceren. Het heeft een kwalitatieve doelstelling: vaststellen welke verschillende problemen er optreden bij de beantwoording van deze vragen uit de alcoholenquête. Dat betekent dat de omvang van de onderzoeksgroep op zichzelf niet ter zake doet, maar *wel* moet de onderzoeksgroep zodanig divers zijn samengesteld (op relevant geachte kenmerken) dat *alle* problemen die zich voordoen in de doelpopulatie worden gesignaleerd. In de *grounded theory* (Glaser en Strauss 1967) methodologie is hiervoor de procedure van de 'theoretische verzadiging' ontwikkeld. Deze houdt in dat na elk interview (of observatie) wordt geanalyseerd of dit relevante nieuwe informatie heeft opgeleverd. Als dat een aantal keren niet meer het geval is geweest, kan worden gestopt met interviewen of observeren. Wij hebben een variant van deze methode gevolgd. De doelpopulatie was die van in Rotterdam wonende Nederlandstalige volwassenen. Wij startten met zeven gemakkelijk bereikbare bekenden (nummers 1-7 in tabel 1). Die zeven interviews leverden in totaal al

voldoende verschillende problemen op om op basis daarvan een eerste rapport te schrijven. Maar vooral omdat er nauwelijks laagopgeleiden bij waren en het opleidingsniveau een relevant kenmerk moet worden geacht, werd daarna nog een aantal interviews gehouden. Eerst drie interviews met laagopgeleiden uit een onderzoeksbestand (nummers 8-10). Dit leverde twee nieuwe problemen op. Daarna werden nog eens zes nieuwe interviews uitgevoerd met personen die uit een ander bestand geselecteerd waren op basis van opleidingsniveau, sekse en niveau van alcoholconsumptie (nummers 11-16). Deze laatste serie interviews is geanalyseerd in de volgorde van afname. Daarbij werd één geheel nieuw normatief probleem ontdekt (vraag 3 in interview 14) en twee nieuwe versies van het 'glasmaat' probleem (vraag 2 e.v. in de interviews 15 en 16). Zo blijkt dat het niet eenvoudig is om te definiëren, wanneer een verzadigingspunt is bereikt. Dat hangt mede af van de definities van categorieën. In dit geval: glasmaatproblemen zijn alle problemen die ontstaan doordat de respondent zijn drank niet uit een voor die drank standaardglas drinkt (met vaste hoeveelheid alcohol), maar bijvoorbeeld uit een veel groter glas of uit een fles. Als je afsprekt dat het glasmaatprobleem één probleem is, dan zijn in de interviews 15 en 16 geen nieuwe problemen ontdekt. Hoe dan ook, na respondent 10 zijn er geen nieuwe *plaatsen* ontdekt waar problemen ontstonden en ook inhoudelijk kwamen er in de laatste zes interviews zo weinig 'echt' nieuwe problemen naar voren, dat wij het verantwoord vonden om te stoppen.

Tabel 1: Overzicht sociale kenmerken van de respondenten

1.	man	49 jaar	alleenstaand	HBO	leidinggevende functie
2.	vrouw	74 jaar	weduwe	BO	huisvrouw
3.	vrouw	44 jaar	gehuwd + kinderen	WO	lerares MBO en VWO
4.	man	31 jaar	alleenstaand	WO	werkloos
5.	vrouw	27 jaar	samenwonend	HBO	inrichtingswerker
6.	man	27 jaar	samenwonend	WO	onderzoeker
7.	vrouw	53 jaar	alleenstaand	HBO	werkloos
8.	man	29 jaar	alleenstaand	MAVO	administratieve functie
9.	man	47 jaar	onbekend	BO	handarbeid volcontinu
10.	vrouw	49 jaar	gehuwd + kinderen	MBO	verzorging
11.	vrouw	48 jaar	gehuwd, geen k	WO	leraar wiskunde
12.	man	61 jaar	gehuwd + kinderen	LBO	onderhoudsmonteur
13.	man	18 jaar	ongehuwd bij ouders	LBO	loodgieter
14.	man	67 jaar	gehuwd	LBO	(ex-) meubelmaker
15.	man	34 jaar	samenwonend	HBO-niveau	IT consultant
16.	vrouw	56 jaar	gehuwd + kinderen	MBO	docent creatieve vaardigheden

Validering: het inhoudelijke interview. In de analyse werd eerst nagegaan welke problemen naar voren kwamen uit het hardop denkend invullen en het daarop volgende *cognitieve interview* over de wijze waarop de vragen werden beantwoord. Vervolgens gingen wij na in hoeverre de antwoorden van de respondenten afwaken van de gegevens over drinkpatronen die wij verkregen in het derde onderdeel, het inhoudelijke interview over drinkgewoonten.

Al tijdens dit laatste interview werden de gegevens die hierin naar voren kwamen, zowel door de onderzoeker als de respondenten behandeld als *geldiger* (dat wil zeggen dichter bij de 'werkelijke' waarde) dan de antwoorden die de respondenten eerder hadden gegeven bij de schriftelijke beantwoording met 'hardop denken'. De reden daarvan is dat tijdens dit inhoudelijke interview vaak drankgebruik herinnerd werd dat eerder vergeten was, of dat de hoeveelheden en frequenties nader werden gespecificeerd, zodat het oorspronkelijke antwoord kon worden gecontroleerd. In een aantal gevallen kon dan worden vastgesteld dat het antwoord op de vraag anders zou zijn geweest als die 'nieuwe' gegevens wel bij de beantwoording waren betrokken. Wanneer respondenten werden geconfronteerd met verschillende uitkomsten die het resultaat waren van verschillende meetmethodes (schriftelijke vragenlijst en interview), gaven zij zelf onveranderlijk aan dat de uitkomst van het interview de meest juiste was, omdat daarbij *naar het oordeel van de respondent* fouten van de andere procedure (de schriftelijke vragenlijst) werden gecorrigeerd (en nooit andersom). Ook het leereffect van de voorafgaande bevraging (hardop denken en cognitief interview) heeft waarschijnlijk sterk bijgedragen aan de ervaren validiteit van de interviewgegevens. Onze bevindingen komen wat dit betreft overeen met die van Means e.a. (1994) naar rapportage over roken. Daarbij bleek dat het hardop denken, bedoeld om informatie te krijgen over de manier waarop respondenten te werk gaan, ook een stimulerend effect heeft op die herinnering zelf. De instructie om hardop te denken werkte zelfs beter dan een training in een specifieke cognitieve strategie.

We kunnen nooit zeker zijn van de perfecte validiteit van welke zelfrapportage dan ook, maar op basis van onze ervaringen lijkt ons wel aannemelijk dat na de 'think aloud', het cognitieve interview en het open inhoudelijke interview, de hoogst haalbare validiteit bereikt is en dat het verantwoord is om de laatste rapportage te gebruiken als valideringscriterium voor de eerste schriftelijke rapportage.

6 Een casus: vraag 2

In de bijlagen geven we schematische overzichten van alle problemen die in de zes onderzochte vragen werden gevonden bij het bureauonderzoek en het veldonderzoek. Om het ruimtebeslag te beperken geven wij hier alleen voor één van de vragen een meer gedetailleerd verslag van de bevindingen. Wij kiezen hierbij

voor vraag 2, omdat dit de meest complexe vraag is en omdat deze vraag de belangrijkste informatiebron is voor de opsporing van 'excessief gebruik', dat wil zeggen een mate van gebruik die medisch riskant wordt geacht.

Uit onderzoek is gebleken, dat een belangrijk deel van de bevolking min of meer geregeld zes of meer glazen alcoholhoudende drank op een dag gebruikt.¹

Heeft u het afgelopen halfjaar² wel eens zes of meer³
glazen⁴ alcoholhoudende drank⁵ op een dag⁶ gedronken?

- 1 () ja, elke dag
- 2 () ja, 5 à 6 keer per week
- 3 () ja, 3 à 4 keer per week
- 4 () ja, 1 à 2 keer per week
- 5 () ja, 1 à 3 keer per maand
- 6 () ja, 3 à 5 keer in dit halfjaar
- 7 () ja, 1 à 2 keer in dit halfjaar
- 8 () nee, geen enkele keer dit halfjaar
- 9 () weet niet

Voor een overzichtelijke weergave van de resultaten hebben we zeven plaatsen in de vraag onderscheiden waar in het bureauonderzoek en/of het veldonderzoek problemen werden gesignaleerd.

1 De inleiding

Deze is bedoeld om potentiële schaamte bij excessieve drinkers te verlichten. Dit werd op het bureau niet als probleembron aangemerkt, maar bleek dat in het veld soms wel te zijn. De inleiding kan juist stigmatiserend werken. Respondent 7 zei bij lezing tijdens de think aloud: 'dat zijn alcoholisten'. Bij het cognitieve interview zei zij verder:

R: 'Dat brengt je op een dwaalspoor ... dat verhaal ... uit onderzoek is gebleken dat een belangrijk deel van de bevolking min of meer geregeld zes of meer glazen alcohol ... dat vind ik een soort draak ... dat leidt me af.'

2 Het afgelopen half jaar

Op het bureau werd betwijfeld of respondenten de termijn 'het afgelopen half jaar' nauwkeurig (kunnen) inschatten. Dat blijkt inderdaad niet het geval te zijn, maar wel wordt steeds aan periodes gedacht van langer dan een of twee maanden, zodat wel ongeveer het langetermijnpatroon wordt gerapporteerd, zoals ook de bedoeling van de onderzoekers is.

Respondenten 4 en 7 meten wel vrij nauwkeurig het afgelopen half jaar af. Respondent 7 doet dat vooral om onderscheid te maken tussen zomer en winter:

I: 'Hoe heb je bij deze vraag geredeneerd?'

R: 'Het afgelopen halfjaar was belangrijk ... als je deze vraag bijvoorbeeld in juni stelt, dan denk ik aan de winter en als je hem nu stelt, dan denk ik aan de zomer en het drankgedrag in zomer en winter bij *mij* ... is anders.'

Op het bureau werd ook als probleem opgemerkt dat het berekenen van week- of maandgemiddelden over een half jaar een erg complexe opdracht is, vooral bij hoge drinkfrequenties.

Dit bleek in het veld inderdaad een probleem te zijn; echter niet zozeer bij hoge drinkfrequenties, maar vooral bij patronen die niet in een week- of maandtelling passen. Bijvoorbeeld respondent 4 die eens in de zes weken veel drinkt en respondent 13 die net twee weken naar Salau was geweest en daar dagelijks grote hoeveelheden had gedronken, maar de rest van het jaar veel minder drinkt. Ook respondent 8 had een probleem.

I: 'Er wordt gevraagd over een half jaar; ... is dat een periode die je kunt overzien?'

R: 'Ja, ik denk het wel ... Maar dit is meer het begin van het half jaar, want nu drink ik weinig zeg maar.'

I: 'je hebt nu een periode dat je weinig drinkt, en voor de zomer heb je meer gedronken?'

R: 'Ja, duidelijk.'

I: 'okee, daar komen we zo dadelijk nog op terug.'

Uit het vervolg bleek dat hij sinds twee maanden veel minder is gaan drinken; bij het invullen van de vraag had hij het 'oude' drinkpatroon aangehouden.

3 Zes of meer

Op het bureau werd opgemerkt dat het de vraag is of mensen hun glazen tellen bij veel drinken en of zij, afgezien daarvan, deze vraag niet zullen opvatten als een vraag naar *voor hun eigen doen* veel drinken.

Dit laatste bleek in het veld inderdaad voor te komen; respondent 7 telt de dagen dat zij *meer dan drie* glazen heeft gedronken:

I: 'Heb je gedacht hoe vaak heb ik nou meer dan zes glazen gedronken?'

R: 'Ja ... en dat weet ik heel goed want ik word er doodziek van ... dus ik weet perfect wanneer ik boven de drie glazen zit bijvoorbeeld.'

I: 'Okee ... maar de vraag gaat over zes of meer.'

R: 'Ja dan ben ik heel dronken.'

4 Glazen

De volumemaat 'glas' is bedoeld als aanduiding van het standaardglas voor de betreffende dranksoorten. Op het bureau werd, op basis van de literatuur, voor-

zien dat niet alle respondenten uit zulke standaardglazen drinken; daarbij werd met name gedacht aan het drinken van bier uit flesjes of blikjes en grote glazen. In het veld bleken twee respondenten inderdaad in flesjes te rekenen; respondent 8 drinkt bier uit flesjes, respondent 15 drinkt uit grote glazen waar een heel flesje in gaat. Respondent 16 drinkt Riesling (8% alcohol, dacht zij) uit een limonadeglas.

5 Alcoholhoudende drank

Gevreesd werd dat er een selectie-effect uit zou kunnen gaan van de voorafgaande vraag, waarbij de respondent één dranksoort moest kiezen als de drank die hij/zij *meestal* drinkt.

Dit effect bleek inderdaad op te treden: een respondent vroeg: 'gaat het nu alleen om bier?' (zijn keuze bij vraag 1).

6 Op een dag

Op het bureau werd voorzien dat de respondent niet alle drank van een *dag* meetelt; maar *drinkgelegenheden* telt waarbij zes of meer glazen zijn geconsumeerd, met als gevolg mogelijke onderrapportage van het aantal dagen waarop zes of meer glazen werd gedronken. Dit probleem werd in het veld niet opgemerkt.

7 De antwoordcategorieën

Hierover werd in het bureauonderzoek geen opmerking gemaakt, maar in het veld bleken er verschillende problemen aan te kleven. Op de eerste plaats was er een respondent die eigenlijk 2 à 3 maal zou willen invullen, maar moest kiezen tussen 1 à 2 en 3 à 5 keer. Daarnaast zorgden deze antwoordcategorieën voor de nodige rekenproblemen, die hierboven onder rubriek 2 al zijn genoemd. En tot slot waren er vergissingen en onbegrip:

Respondent 9 had erg veel moeite met deze vraag. Hij begreep niet dat hij een gemiddelde moest uitrekenen en kruiste twee antwoorden aan:

I: 'En het is dus ongeveer drie of vier keer per week dat u zes of meer glazen drinkt?'

R: 'Ken ook een week dat ik niets drink, ken ook een week nemen dat zes eh ...'

I: 'En nu heeft u hier ook een kruisje ingevuld, bij 'weet niet'.'

R: 'Ja, ik weet niet.'

I: 'Waarom heeft u dat kruisje daar neer gezet?'

R: 'Nou, de ene keer drie en de volgende keer niet.'

Uit het inhoudelijke interview bleek dat hij in de volcontinu ploegendienst werkt en alleen drinkt in de vrije periode.

Respondent 13 zette een kruisje bij '3 à 5 keer dit halfjaar', terwijl hij bedoelde: 3 à 5 keer *per week*, dit halfjaar.

7 Samenvatting en conclusie

In het bureauonderzoek werden in vraag 2 zes passages als probleembron aangewezen (2, 3, 4, 5, 6 en 7). Behalve passage 6 bleken al deze passages in het veld ook inderdaad tot problemen te leiden. Dat bleek bovendien het geval te zijn met passage 1, die in het bureau niet was opgemerkt. Er is dus bij deze vraag een hoge mate van overeenstemming tussen beide onderzoeken wat betreft de identificatie van *probleemplaatsen*. Dat blijkt ook bij de andere vragen het geval te zijn zoals kan worden nagegaan in bijlage 1, waar wij alle onderzoeksresultaten in kort bestek hebben weergegeven. Van de in totaal 17 *probleemplaatsen* die in het bureauonderzoek werden aangewezen, werden er 14 in het veldonderzoek bevestigd; daarenboven leverde het veldonderzoek nog 10 nieuwe *probleemplaatsen* op. Dus het veldonderzoek is wel duidelijk productiever in de identificatie van *probleemplaatsen*, nog afgezien van het feit dat bij een aantal *probleemplaatsen* in het veldonderzoek verschillende problemen werden gesignaleerd.

Behalve dit verschil in productie zijn er ook belangrijke verschillen in *probleeminhouden*. Het veldonderzoek bracht een aantal specifieke problemen aan het licht die niet gemakkelijk uit de formulering van de vraag (grammatica, interne logica) zijn af te leiden, omdat zij sterk gebonden zijn aan de specifieke biografische inhoud (in dit geval het verloop van het drinkgedrag) bij de betreffende respondenten. Bij het bureauonderzoek was niet gedacht aan:

- mensen in ploegendienst met een daaraan verbonden drinkritme;
- mensen die na drie glazen al dronken zijn;
- mensen die gedurende het afgelopen half jaar heel verschillende drinkpatronen hebben gehad;
- mensen die net terug zijn van een zuipvakantie.

Het bureauonderzoek detecteert vooral cognitieve problemen die ontstaan door de complexiteit van de logische operaties die respondenten moeten uitvoeren om tot een antwoord te komen en door inconsistenties binnen en tussen vragen. Het veldonderzoek detecteert vooral problemen die ontstaan doordat 'de werkelijkheid' (feitelijke drinkpatronen) afwijkt van het 'ideale', regelmatige, in maanden of weken telbare patroon dat impliciet in de vragenlijst verondersteld is. Dit alles geldt niet alleen bij de hier uitgewerkte casus (vraag 2), maar ook bij de andere vragen (zie bijlage 1).

Uit het feit dat al de genoemde onvoorziene drinkgewoonten aangetroffen zijn in de kleine groep van 16 respondenten die toch vrij willekeurig gekozen werd, kan bovendien worden opgemaakt dat het niet om zeldzame 'uitzonderingen' gaat die je nu eenmaal altijd wel eens hebt. Ook in deze zin blijkt het cognitieve veldonderzoek uitermate informatief te zijn; zelfs zo'n kleine 'theoretische' steekproef geeft een indicatie van de omvang van problemen die het bureau-

onderzoek niet kan leveren, omdat de daar gesignaleerde problemen uit de aard der zaak *hypothetisch* blijven. Bovendien is gebleken dat de uitbreiding van het bij survey-instituten inmiddels min of meer gebruikelijke ‘hardop denken’ en ‘cognitieve interview’ met een derde stap, het inhoudelijke interview, duidelijk bijdraagt aan inzicht in de problemen die bij respondenten ontstaan bij het beantwoorden van deze vragen.

Een andere conclusie kan zijn dat het veldonderzoek laat zien, dat er naast cognitieve ook normatieve problemen zijn; in vraag 2 was dat het geval bij de inleiding. Bij vraag 3, ‘Hoeveel dagen per maand drinkt u gemiddeld genomen?’, was het probleem dat ‘drinken’ zonder specificatie de connotatie heeft van ‘alcoholisme’. Respondent 14 bij het hardop denken: ‘het lijkt wel of ik elke dag dronken ben’. De vraag ‘Drinkt u?’ is een andere dan de vraag ‘Wat drinkt u?’ Twee respondenten gaven op vraag 3 dan ook een *normatief* antwoord. Zij vulden in wat zij voor zichzelf acceptabel vonden; zij vonden het raar om in te vullen dat zij bijna elke dag drinken, hoewel zij elke dag wel een of twee glazen alcoholische drank drinken.

In hoeverre de hier gevonden verschillen tussen bureauonderzoek en cognitief veldonderzoek generaliseerbaar zijn, valt niet gemakkelijk te beoordelen en ligt ook niet vast. Op de eerste plaats ging het hier om vragen die een beroep doen op het geheugen en die rekenwerk vereisen. Wellicht dat bij andere soorten vragen, bijvoorbeeld naar attitudes of feitenkennis, andere verschillen gevonden worden. Verder onderzoek zal dat moeten leren.

Op de tweede plaats zal het bureauonderzoek aan productiviteit winnen naarmate de experts zich meer laten informeren door resultaten van cognitief onderzoek. Maar dat neemt niet weg dat cognitief onderzoek onder de doelpopulatie – bij voorkeur in het ‘veld’ – onontbeerlijk blijft voor de bepaling van de kwaliteit van surveyvragen; bureauonderzoek alleen is nuttig, maar uit de aard der zaak onvolledig.

Tot slot. Wat betekent een en ander nu voor de praktijk van het surveyonderzoek? Hoewel vaak goede alternatieven voorhanden zijn, geldt voor de dagelijkse praktijk van het sociologische onderzoek nog veelal dat het survey zonder verdere reflectie als de waarnemingsmethode bij uitstek wordt gezien (Ganzenboom 1998; Hak en Jansen 1998). Het is daarom bijzonder belangrijk dat bij ieder survey opnieuw wordt gecontroleerd of de gewonnen data geldig en voldoende nauwkeurig zijn. Wij bevelen aan dat voorafgaande en tijdens ieder survey een kwalitatief nevenonderzoek wordt verricht, waarin de kwaliteit van de via de vragenlijst verkregen data wordt getest.

Daarbij kan gebruik worden gemaakt van het hier gepresenteerde Drie-Stappen Test-Interview (DSTI) of een goed alternatief daarvoor.

Geraadpleegde literatuur

- Bergman, L.R. (1995) Pre-testing procedures at Statistics Sweden's Measurement, Evaluation and Development Laboratory. *Journal of Official Statistics*, 11, 309-323.
- Blair, E., Sudman, S., Bradburn, N.M., en Stocking, C. (1977) How to ask questions about drinking and sex: Response effects in measuring consumer behavior. *Journal of Methodological Research*, 14, 316-321.
- Blair, J., & Presser, S. (1994) Survey pretesting: Do different methods produce different results? In: P. Marsden (red.) *Sociological Methodology* (pp. 73-114). New York: Blackwell.
- Bongers, I., Goor, L. van de, Garretsen, H. en Oers, H. van (1999) Aggregate comparison of self-reported versus non-self-reported drinking in a general population survey. *Substance use and misuse*, 34, 421- 441.
- Bongers, I. (1998) Problem drinking among the general population: A public health issue? Ph.D. dissertation. Rotterdam: Addiction Research Institute. IVO series nr 18.
- Campanelli, P. (1997), Testing survey questions: New directions in cognitive interviewing. *Bulletin de Méthodologie Sociologique*, 55, 5-17.
- CBS (1989) Vragenlijst Gezondheidsenquête van het Centraal Bureau voor de Statistiek. Heerlen: CBS, Hoofdafdeling gezondheidsstatistiek.
- Conrad, F. en Blair, J. (1998) From impressions to data: Increasing the objectivity of cognitive interviews. *The 1996 proceedings of the section on survey research methods: Annual meeting of the American Statistical Association*, 1-10.
- Espósito, J.L. en Rothgeb, J.M. (1997) Evaluating survey data: Making the transition from pretesting to quality assessment. In: L.Lyberg e.a. (red.) *Survey measurement and process quality*, pp. 73-114. New York: Wiley & Sons.
- Ganzenboom, H.B.G. (1998) Enige observaties over sociologen en biologen. *FACTA*, 6, nr.4 (juni 1998), 28-29.
- Garretsen, H.F.L. (1983) *Probleemdrinken: Prevalentiebepaling, beïnvloedende factoren en preventiemogelijkheden*. Proefschrift, Katholieke Universiteit Nijmegen.
- Garretsen, H.F.L., en Knibbe, R.A. (1983) *Alkohol prevalentie onderzoek Rotterdam/Limburg*. Ministerie WVC.
- GGD-en Groningen (1994) Vragenlijst over de gezondheid en leefstijl van bewoners van de provincie Groningen.
- Glaser, B. en Strauss, A. (1967) *The discovery of Grounded Theory*. Chicago: Aldine.
- Goor, I.A.M. v. d., Oers, H.A.M. v., Bongers, I.M.B., en Garretsen, H.F.L. (1996) *Risante leefgewoonten in Rotterdam anno 1994: Een onderzoek naar het gebruik van alcohol, slaap- en kalmerende middelen, tabak, hasj en marihuana en de deelname aan kansspelen onder de Rotterdamse bevolking* (Vol. 10). Rotterdam: Instituut voor Verslavingsonderzoek.
- Hak, T., en Jansen, H. (1997) *Het meten van alcoholconsumptie per postenquête: Een onderzoek naar cognitieve problemen bij de beantwoording van vragen naar frequenties en hoeveelheden*. Rotterdam: Instituut voor Verslavingsonderzoek.
- Hak, T., en Jansen, H. (1997) Over sociologisch waarnemen. *FACTA*, 6, nr.8 (december 1998), 24.
- Hak, T., en Veer, C. v.d. (1999) *Validation of the Illegal Alien (AI) Scale through cognitive interviewing – Interviewing guidelines*. Amsterdam: Vrije Universiteit, Afdeling Methoden en Technieken SCW.
- Harkness, J. (1996) *The (re)presentation of self in everyday questionnaires*. Paper presented at the Fourth International Social Science Methodology Conference, University of

- Essex, Colchester UK.
- Heer, W. de, Leeuw, E. de, en Zouwen, J. van der (1999) Methodological issues in survey research: A historical review. *Bulletin de Méthodologie Sociologique*, 64, 25-48.
- iMGZ (1991) Vragenlijst Gezondheid en Levensomstandigheden Bevolging Eindhoven en omstreken. (GLOBE). Rotterdam: Erasmus Universiteit, Instituut Maatschappelijke Gezondheidszorg.
- Knibbe, R.A. (1982) *Probleemdrinken in Limburg*. Proefschrift, Rijks Universiteit Limburg.
- Koehnken, G., Schimossek, E., Asschermann, E., en Hoefler, E. (1995) The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology*, 80, 671-684.
- Lemmens, P.H.H.M. (1987) *Het Ledermann-model nader bezien: Gevolgen van onderschatting voor de verdeling van alcoholconsumptie*. Vakgroep Medische Sociologie. Rijksuniversiteit Limburg.
- Lemmens, P.H.H.M. (1991) *Measurement and distribution of alcohol consumption*. Proefschrift, Rijksuniversiteit Limburg.
- Lemmens, P., Tan, E.S., en Knibbe, R.A. (1992) Measuring quantity and frequency of drinking in a general population survey: A comparison of five indices. *Journal of Studies on Alcohol*, 53, 476-486.
- Lindström, H.L. (1996) *Measurement research development at Statistics Sweden*. Paper presented at the Fourth International Social Science Methodology Conference, University of Essex, Colchester UK.
- Loftus, E.F., Fienberg, S.E., en Tanur, J.M. (1985) Cognitive psychology meets the national survey. *The American Psychologist*, 40, 175-180.
- Luppes, M. (1993) *Interpretatie en waardering van aanschrijfbrieven* (Intern rapport). CBS Heerlen.
- Means, B., Swan, G.E., Jobe, J.B., en Esposito, J.L. (1994) The effects of estimation strategies on the accuracy of respondents' reports on cigarette smoking. In: N. Schwarz en S. Sudman (red.) *Autobiographical memory and the validity of retrospective reports*. New York: Springer, 107-120.
- Midanik, L. (1982) The validity of self-reported alcohol consumption and alcohol problems: A literature review. *British Journal of Addiction*, 77, 357-382.
- Midanik, L.T. (1988) Validity of self-reported alcohol use: A literature review and assessment. *British Journal of Addiction*, 83, 1019-1029.
- Midanik, L. en Hines, A. (1991) 'Unstandard' ways of answering standard questions: Protocol-analysis in alcohol survey research. *Drug and Alcohol Dependence*, 27, 245-252.
- Midanik, L.T. (1994) Comparing usual quantity/frequency and graduated frequency scales to assess yearly alcohol consumption: Results from the 1990 US National Alcohol Survey. *Addiction*, 89, 407-412.
- Mingay, D.J., Shevell, S.K., Bradburn, N.M., en Ramirez, C. (1994) Self and proxy reports of everyday events. In: N. Schwarz en S. Sudman (red.) *Autobiographical memory and the validity of retrospective reports* (pp. 235-250). New York: Springer.
- Reinhard, O.P.M., en Rood-Bakker, D.S. (1998) *Alcoholgebruik in beeld: Standaardmeetlat*. Rotterdam: Nederlands Economisch Instituut.
- Schwarz, N., en Sudman, S. (red.) (1994) *Autobiographical memory and the validity of retrospective reports*. New York: Springer.
- Single, E. en Wortley, S. (1994) A comparison of alternative measures of alcohol consumption in the Canadian National Survey of alcohol and drug use. *Addiction*, 89, 395-399.

- Snijkers, G (1998) Dataverzameling via vragenlijsten: Cognitief pretesten van vragenlijsten. Handout lezing Amsterdam 28 oktober 1998. Heerlen: Centraal Bureau voor de Statistiek.
- Sudman, S., Bickart, B., Blair, J., en Menon, G. (1994) The effect of participation level on reports of behavior and attitudes by proxy reporters. In: N. Schwarz en S. Sudman (red.) *Autobiographical memory and the validity of retrospective reports*, 251-266. New York: Springer.
- Sudman, S., Bradburn, N.M., en Schwarz, N. (1996) *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
- Waterson, E.J., en Murray-Lyon, I.M. (1989) Screening for alcohol related problems in the antenatal clinic. *Alcohol and Alcoholism*, 24, 21-30
- Willis, G.B. en Schechter, S. (1997) Evaluation of cognitive interviewing techniques: do the results generalize to the field? *Bulletin de Méthodologie Sociologique*, 55, 40-66.

Bijlage 1: Samenvatting en vergelijking resultaten bureauonderzoek en veldonderzoek*

Vraag 1	bureauonderzoek	veldonderzoek
Welke van de onderstaande alcoholische drank gebruikt u meestal ¹ als u drinkt? S.v.p. één antwoord aankruisen. ²	– ² beperking tot één antwoord leidt mogelijk tot onterechte selectie van deze dranksoort bij verdere vragen + ⁵ iemand die meestal alcoholarm of alcoholvrij bier drinkt maar ook wel eens iets anders, wordt ten onrechte doorgestuurd naar 5 – ⁶ alcoholvrij bier wordt als alcoholhoudende drank voorgesteld; kan leiden tot overrapportage	– ¹ R5 vat <i>meestal</i> op als hoogste frequentie, R2 en R6 als grootste hoeveelheid + ² R8 vraagt bij volgende vraag of het nu alleen gaat om bier, de hier door hem gekozen dranksoort – ² R8: één antwoord kan niet – ² R3 leest over de instructie heen R9 begrijpt het niet; beiden kruisen twee antwoorden aan – ³ R4: vreemde categorie, zou bier en wijn samen nemen – ⁴ R1 vermijdt deze categorie, hoewel toepasselijk; R7 en R13 kiezen dit bij gebrek aan beter + ⁵ R9 drinkt soms alcoholarm, gaat door naar vraag 5 terwijl hij ook gewoon bier drinkt
1 ☐ bier		
2 ☐ wijn, sherry, port of vermout ³		
3 ☐ likeur, advocaat, bessen-jenever of citroenjenever		
4 ☐ jenever, brandewijn, vieux, rum, cognac, whisky, wodka of ander gedestilleerd		
5 ☐ frisdrank gemengd met alcoholhoudende drank		
6 ☐ verschilt erg per keer ⁴		
7 ☐ alcoholarm ⁵ of alcoholvrij ⁶ bier —> door naar vraag 5		
8 ☐ ik drink nooit alcohol —> door naar vraag 5		

* Een signaleerd probleem is gemerkt met '+' als dit ook bij het andere onderzoek naar voren kwam, en met '-' als het in het andere onderzoek niet naar voren is gekomen.

Vraag 2	bureauonderzoek	veldonderzoek
<i>Uit onderzoek is gebleken, dat een belangrijk deel van de bevolking min of meer geregeld zes of meer glazen alcoholhoudende drank op een dag gebruikt.</i>¹ <i>Heeft u het afgelopen ²halfjaar wel eens zes of ³meer glazen ⁴alcoholhoudende ⁵drank op een dag ⁶gedronken?</i> 1 ☐ ja, elke dag 2 ☐ ja, 5 à 6 keer per week 3 ☐ ja, 3 à 4 keer per week 4 ☐ ja, 1 à 2 keer per week ⁷ 5 ☐ ja, 1 à 3 keer per maand ⁷ 6 ☐ ja, 3 à 5 keer in dit halfjaar 7 ☐ ja, 1 à 2 keer in dit halfjaar 8 ☐ nee, geen enkele keer dit halfjaar 9 ☐ weet niet	+ ² waarschijnlijk zullen R-en de periode 'half jaar' niet nauwkeurig afmeten + ² vooral R-en met hoge frequenties zullen extrapoleren van afgelopen week naar half jaar + ³ niet elke R houdt aantal glazen bij als hij/zij drinkt + ³ zes glazen wordt wellicht opgevat als <i>veel</i> gedronken + ⁴ uit de literatuur is bekend dat mensen thuis vaak bier uit flesjes of blikjes drinken en andere drank vaak uit grotere glazen dan de standaard maat + ⁵ als gevolg van selectie-effect van vraag 1 zal niet alle bedoelde drank worden meegeteld – ⁶ R rekent wellicht per drinkgelegenheid i.p.v. per dag	– ¹ de inleiding werkt bij R7 averechts + ² R4 en R7 meten het afgelopen half jaar precies af; anderen niet. – ² Voor R4, R8, R9 en R13 is het moeilijk om een gemiddelde over het half jaar te berekenen door hetzij een regelmatig patroon in een afwijkende tijdschaal (R4: 6 weken, R9: volcontinu rooster) hetzij door onregelmatigheden (R13 zuipvakantie, R8 sinds 2 mnd sterk geminderd) + ³ R7 interpreteert dit als vraag naar <i>veel</i> drinken wat in haar geval duidelijk minder is dan 6 glazen + ⁴ R8, R15 en R16 rekenen in te grote glazen of flesjes + ⁵ R8 vraagt of het nu alleen om bier gaat – ⁷ antwoordcategorien sturend: R7 wil 2 à 3 invullen en moet kiezen tussen 1 à 2 en 3 à 5; R16 kiest ten onrechte de eerste 'ja' (= ja, elke dag)

* Een gesignaleerd probleem is gemerkt met '+' als dit ook bij het andere onderzoek naar voren kwam, en met '-' als het in het andere onderzoek niet naar voren is gekomen.

Vraag 3¹

bureauonderzoek

veldonderzoek

Hoeveel dagen per maand drinkt²
u gemiddeld³ genomen?^{4,5}

- 1 ☐ 28 of meer
2 ☐ 24 t/m 27
3 ☐ 21 t/m 23⁶
4 ☐ 15 t/m 20⁶
5 ☐ 12 t/m 14
6 ☐ 9 t/m 11
7 ☐ 6 t/m 8
8 ☐ 3 t/m 5
9 ☐ 2 of minder

±¹ volgorde: logischer lijkt het deze vraag te laten voorafgaan aan de vorige vraag (6+ dagen)
–⁴ er wordt niet expliciet vermeld dat het om alcoholhoudende drank gaat, maar zal waarschijnlijk evident zijn
+⁵ de vraag is over welke periode het 'gemiddeld genomen' wordt berekend; er wordt hier geen periode genoemd i.t.t. vorige vraag
+⁶ de antwoordcategorieën suggereren een hoge mate van precisie; R-en zullen waarschijnlijk beginnen met een schatting per week en dit extrapoleren door vermenigvuldiging met vier. In dat geval zijn de antwoordcategorieën onhandig gekozen en zal het gemiddelde niet erg nauwkeurig bepaald worden

±¹ De neiging om bij 'drinkdagen' alleen te denken aan dagen dat er veel gedronken is, is wellicht mede effect van de vraagvolgorde (vraag 2 ging over 6+ glazen)
–² R8, R9 en R12 vatten *drinkdagen* op als dagen waarop veel gedronken hebben. Een pilsje bij het eten is voor hen geen *drinken*. Bij R14 maakt deze formulering het moeilijk om '28 of meer' in te vullen, hoewel hij elke dag een borrel neemt
–³ R3 vindt het moeilijk om een algemeen gemiddelde te berekenen vanwege verschillende periodes
+⁵ Alle R denken hier aan een periode van meerdere maanden zoals bedoeld. Maar: R8 wiens drinkpatroon recent gewijzigd is, vult in volgens het huidige patroon; R7 benadrukt seizoensverschil zomer-winter; R5 meldt verandering in patroon
+⁶ De meeste respondenten extrapoleren weekpatroon. R4 rekent 'eens per 4 a 5 dagen' om naar maand

* Een gesignaleerd probleem is gemerkt met '+' als dit ook bij het andere onderzoek naar voren kwam, en met '-' als het in het andere onderzoek niet naar voren is gekomen.

Vraag 4	bureauonderzoek	veldonderzoek
<i>Als u op een dag ¹ alcohol drinkt, hoeveel glazen drinkt u dan gemiddeld ² (halve glazen ³ naar boven afronden)?</i> 1 ☐ 11 of meer glazen 2 ☐ 7-10 glazen 3 ☐ 6 glazen 4 ☐ 4-5 glazen 5 ☐ 3 glazen 6 ☐ 2 glazen 7 ☐ 1 glas	+ ² vraag is of R echt gaat rekenen om aan een rekenkundig gemiddelde te komen of dat zij 'gemiddeld' opvat als 'normaal gesproken' waarbij zij eerder een mediane of modale waarde zal invullen. Vooral bij R met sterke wisseling in consumptie zal dit tot problemen leiden + ³ het is zeer de vraag of het onderscheid tussen hele en halve glazen, dat hier wordt geïntroduceerd, voor R-en hanteerbaar is + ⁴ verschil in range van antwoord-categorieën	- ¹ R1 rekent niet per <i>drinkdag</i> maar per dag R5 kiest modale drinkdag = dag dat zij wijn drinkt; op een 'bierdag' drinkt zij veel meer R8 rekent alleen dagen dat hij <i>veel</i> drinkt + ² R4 berekent een echt rekenkundig gemiddelde; alle anderen geven eerder een modale of mediane waarde - ² R1 en R7 geven een 'normatief' antwoord – nl. wat zij voor zichzelf nastreven. - ² R10 vult haar maximum in (3 glazen); dit is hoger dan het gemiddelde + ³ R1 en R5 merken op dat dit onderscheid niet werkt; R8, R15 en R16 drinken (vaak) uit grotere dan standaardglazen. + ⁴ R7 maakt opmerking over de 'opbouw van de antwoordcategorieën'

* Een gesignaleerd probleem is gemerkt met '+' als dit ook bij het andere onderzoek naar voren kwam, en met '-' als het in het andere onderzoek niet naar voren is gekomen.

Vraag 5	bureauonderzoek	veldonderzoek
<i>Zijn er in de afgelopen¹ tien jaar periodes geweest dat u beduidend² meer of minder dronk dan nu³?</i>	+ ¹ de referentieperiode 'afgelopen tien jaar' overlapt met het 'afgelopen halfjaar' dat eerder als referentie voor het actuele drinkpatroon is gehanteerd + ² <i>beduidend</i> en <i>ongeveer gelijk</i> zijn vage termen die de R geen houvast geven	+ ¹ R8 legt grens tussen 'afgelopen 10 jaar' en 'nu' bij 2 maanden geleden toen hij begon met minder drinken. + ² Het begrip 'beduidend' wordt verschillend geïnterpreteerd. R5 selecteert extreme periode (jaar Amerika); R3 in 'think aloud': "zo'n woordje als 'beduidend'". Er zijn verschillende periodes, o.a. afhankelijk van deelname aan cursussen - ³ R1, R7 en R9 vergelijken afgelopen 10 jaar als geheel met voorafgaande periode i.p.v. nu - ⁴ R10 stuit op een foute routing na het antwoord 'nee' (code 1) gaat zij ten onrechte naar vraag 6 die niet meer van toepassing is
1 ☐ nee, mijn drinkpatroon is ongeveer gelijk gebleven ⁴		
2 ☐ ja, ik heb beduidend meer gedronken dan ik nu doe		
3 ☐ ja, ik heb beduidend minder gedronken dan ik nu doe		

Vraag 6	bureauonderzoek
<i>Kunt u globaal¹ aangeven hoe lang u beduidend² meer of minder heeft gedronken?</i>	- ¹ 'globaal' is een onduidelijke term - ² 'beduidend' idem
..... jaren maanden	veldonderzoek

* Een signaleerd probleem is gemerkt met '+' als dit ook bij het andere onderzoek naar voren kwam, en met '-' als het in het andere onderzoek niet naar voren is gekomen.

Bijlage 2: Veldonderzoek: Het Drie Stappen Test Interview (DSTI)

Bij het ontbreken van een standaardmethode hebben wij voor dit onderzoek een drie-stapenprocedure uitgewerkt (zie Hak en Van der Veer 1999 voor een uitgeschreven protocol voor deze procedure):

1. *Hardop denkend invullen.* Aan de respondent werd gevraagd om de zes vragen schriftelijk te beantwoorden en om bij de beantwoording zoveel mogelijk 'hardop' te denken. Sommige respondenten hadden moeite met deze opdracht, hetgeen op de band wordt gedocumenteerd door (soms lange) pauzes. Het resultaat van dit onderdeel van het interview bestaat uit twee delen: een ingevulde vragenlijst en een geluidsopname die het denkproces documenteert.

2. *Cognitief interview achteraf* over de wijze waarop de vragen zijn beantwoord. Dit onderdeel was bedoeld als een exploratie van het denkproces zoals dat in de voorgaande fase van het 'hardop denken' plaatsvond. Wij gebruikten daarbij de volgende itemlijst, mede gebaseerd op het bureauonderzoek. Deze lijst werd echter niet als een dwingende handleiding gebruikt:

Vraag 1

- Hoe heb je geredeneerd bij het invullen van de vraag?
- Had je moeite deze vraag in te vullen? Zo ja, waarom?
- Wat vind je van de antwoordcategorieën?
 - Ontbreekt er een soort drank die voor jou belangrijk is?
 - Drink je wel eens alcoholarm bier? Zo ja, heb je die meegeteld bij de vervolgvragen?

Vraag 2

- Hoe heb je geredeneerd bij het invullen van de vraag?
- Had je moeite deze vraag in te vullen? Zo ja, waarom?
 - Heb je je afgevraagd wat een *glas* is? Hoe ben je daarmee omgegaan bij het beantwoorden?
 - Hoe ben je omgegaan met de opdracht deze vraag te beantwoorden voor het *afgelopen halfjaar*?
 - Heb je geteld? Wat? Hoe?
- Wat vind je van de antwoordcategorieën?
 - Hoe heb je berekend wat jouw juiste antwoord is?

Vraag 3

- Hoe heb je geredeneerd bij het invullen van de vraag?
- Had je moeite deze vraag in te vullen? Zo ja, waarom?
 - Hoe heb je *gemiddeld genomen* geïnterpreteerd?
- Wat vind je van de antwoordcategorieën?
 - Hoe heb je berekend wat jouw juiste antwoord is?

Vraag 4

- Hoe heb je geredeneerd bij het invullen van de vraag?
- Had je moeite deze vraag in te vullen? Zo ja, waarom?
 - Hoe heb je *gemiddeld* geïnterpreteerd?
 - Hoe heb je het aantal glazen vastgesteld? Hoe heb je afgerond?
- Wat vind je van de antwoordcategorieën?
 - Hoe heb je berekend wat jouw juiste antwoord is?

Vragen 5-6

- Hoe heb je geredeneerd bij het invullen van deze vragen?
- Had je moeite deze vragen in te vullen?

3. *Interview over drinkpatronen (inhoudelijk interview)*. In dit derde onderdeel werd getracht om een zo precies mogelijk beeld te krijgen van het drinkpatroon van de betreffende respondent, om zo te kunnen vaststellen of de in de eerste fase verkregen antwoorden die (uiteraard ook subjectieve) werkelijkheid wel voldoende benaderen en om eventuele problemen met de eerdere vragen nader te specificeren. Ook hierbij werd van een itemlijst uitgegaan, die niet dwingend werd gevolgd:

1. Drinkpatronen

De meeste mensen hebben gewoonten. Voorbeelden zijn:

- een borrel voor het slapen gaan
- twee glazen wijn bij het eten
- een borrel voor het eten
- naar het stamcafé (met vaste consumptie)

Heb jij één of meer van dergelijke patronen?

- Zo ja, wil je ze beschrijven?
 - Consumptie per keer: aantal 'glazen'?
 - Frequentie: dagelijks? Waarvan hangt dit af?

2. Soorten dagen

De meeste mensen hebben een typisch drinkpatroon voor 'gewone' dagen ('doorde weekse' dagen of 'werkdagen') en een ander drinkpatroon voor 'weekend' dagen. Geldt dat ook voor jou?

- Zo nee, is dit omdat eigenlijk *elke* dag weer anders is wat dit betreft?
- Zo nee, is dit omdat je drinkpatroon *constant* is, ongeacht het soort dag?
- Zo ja:
 - Welk onderscheid maak je hier? Een tweedeling (of driedeling, enzovoort)?
 - Beschrijf die verschillende patronen.

3. *Speciale dagen*

Tot nu toe hebben we het gehad over het normale drinkpatronen. Maar er kunnen speciale gebeurtenissen zijn waarbij dingen anders gaan, zoals verjaardagen, carnaval, oudejaar, vakantie, recepties. Geldt dat ook voor jou? Kun je een lijst maken van soorten gebeurtenissen (of periodes in het jaar) waarin je drankgebruik hoger is dan op een normale drinkdag?

- Zo ja, geef per geval aan waarin het verschilt van een normale drinkdag: soort drank; hoeveelheid.
- Idem, voor periodes met minder drankgebruik.

N.B. in de latere interviews (11 t/m 16) werd een checklist gebruikt van bijzondere gelegenheden.