

A.G. Heinen

J.A.P. Hagenaars

1. Inleiding

In de afgelopen jaren zijn aan de sociale faculteit van de Katholieke Hogeschool te Tilburg bij enkele survey-studies gegevens over de interviewers verzameld. Hieronder zullen enkele resultaten van een eerste analyse van dit materiaal weergegeven worden; tevens zal kort aangegeven worden welke analyses er nog gepland zijn. Een en ander werd mogelijk gemaakt door subsidie van Z.W.O.

2. Een drietal modellen ten aanzien van interviewereffecten

De theorieën rond "response effects" in het algemeen en interviewer-effecten in het bijzonder zijn van een zeer karig gehalte. Het formuleren van empirisch toetsbare hypothesen die theoretisch relevant zijn, is dan ook nog verre toekomstmuziek. Wel is het mogelijk uit de bestaande theoretische studies over interviewereffecten enige min of meer globale denkmodellen te destilleren. Een drietal van deze modellen zullen kort beschreven worden.

In de eerste plaats kan een behaviouristisch model onderscheiden worden. In de vroegere gedachtengang rond het survey onderzoek werden interviewers als uitwisselbare en identieke registratie-mechanismen gezien, die niet van invloed zijn op de antwoorden van de respondenten. Al snel echter wordt de opvatting gemeengoed dat bepaalde kenmerken van interviewers een vertekende werking kunnen hebben op antwoorden van respondenten, waarbij men echter nog geen aandacht schenkt aan mogelijke interacties tussen respondent- en interviewerkenmerken. Het grootste deel van het onderzoek naar interviewereffecten heeft dit tweede model als uitgangspunt.

Pas de laatste tijd wordt er meer de nadruk gelegd op de mogelijke interacties tussen respondent- en interviewerkenmerken. Men stelt in dit derde model dat niet alleen bepaalde interviewerkenmerken, maar ook vooral bepaalde combinaties van respondent- en interviewerkenmerken de antwoorden van respondenten kunnen vertekenen.

3. Het analysemateriaal

De drie bovenomschreven modellen zijn getoetst aan gegevens afkomstig uit een in 1970 en 1971 gehouden panelstudie. Het betreft een

verkiezingsonderzoek dat in het kader van de Tweede Kamer verkiezingen van 28 april 1971 werd verricht. Voorafgaande aan deze verkiezingen werd een vaste groep respondenten zesmaal benaderd. Deze interviews werden gehouden in opdracht van de KRO actualiteitenrubriek "Brandpunt". De uitvoering van het project lag in handen van Bureau Intomart te Hilversum.

Door Intomart is een getrapte adressensteekproef getrokken. Per adres werd door de interviewer met behulp van een lootschema vastgesteld wie van de personen, die op dat adres woonachtig waren, én ouder waren dan 17 jaar, geïnterviewd moest worden. De (ongewogen) omvang van de panelgroep was ongeveer 600 respondenten.

Voor iedere respondent die tot de panelgroep behoorde werd zodoende een grote hoeveelheid informatie verzameld. Aan dit geheel van informatie werden per respondent en per wave nog enige gegevens toegevoegd over de interviewer die de respondent in die betreffende wave had ondervraagd. Dit laatste geldt echter alleen voor die respondenten, die aan tenminste vier van de zes interviews hebben deelgenomen.

De volgende interviewergegevens werden geregistreerd:

- geboortjaar van de interviewer
- geslacht van de interviewer
- opleidingsniveau van de interviewer (algemeen vormend onderwijs en beroepsonderwijs)
- duur van het dienstverband van de interviewer bij Intomart.

Omdat ervan uitgegaan werd dat interviewereffecten o.a. duidelijk naar voren zullen komen bij "bedreigende" onderwerpen, zijn de analyses uitgevoerd op data die voor dit predikaat "bedreigend" in aanmerking lijken te komen: 38 vragen rond pornografie, homofilie en abortus¹⁾.

Deze 38 variabelen werden als afhankelijke variabelen beschouwd. Gelet op de beschikbare interviewervariabelen leken leeftijd en geslacht van de interviewer samen met leeftijd en geslacht van de respondent het meest bruikbaar als onafhankelijke variabelen. Op basis van de 38 eerder vermelde vragen zijn nog enkele andere afhankelijke variabelen geconstrueerd, te weten:

- a) een somscore voor de scores "geen antwoord", "geen mening", "nooit over nagedacht", en "wil niet zeggen".
- b) het aantal omstandigheden dat men opsomde, waaronder abortus is toegestaan, en
- c) een drietal factorscores gebaseerd op 22 van de 38 variabelen.

Deze factorscores weerspiegelden respectievelijk een tolerantie-dimensie, een deviantiedimensie en een pornografiedimensie.

4. De analysetechnieken

Om het materiaal te analyseren is gebruik gemaakt van een tweetal technieken, nl. log-lineaire analyse en regressieanalyse met dummy-variabelen²⁾.

De drie bovenomschreven modellen kunnen met behulp van deze technieken geformaliseerd worden. Model 1, het behaviouristische model, veronderstelt dat er alleen effect uitgeoefend wordt door de variabelen leeftijd en geslacht van de respondent, waarbij deze variabelen elkaars invloed op de afhankelijke variabele kunnen versterken of verzwakken. Model 2, het traditionele model, is een uitbreiding van model 1 in die zin dat er nu ook een effect wordt verondersteld van de interviewerkenmerken leeftijd en geslacht en eventueel van interacties tussen de interviewerkenmerken, maar niet van de interacties tussen de interviewerkenmerken en de respondentkenmerken. Deze laatste interacties worden wel opgenomen, naast alle eerder genoemde effecten, in model 3, het interactiemodel.

Bij de log-lineaire analyse voert men voor elk model een "goodness of fit"-toets uit met behulp van de likelihoodratio L en tevens is het mogelijk de modellen ten opzichte van elkaar te toetsen door de verschillen tussen de likelihoodratio's te toetsen.

Bij de regressieanalyse is iets soortgelijks mogelijk. Voor elk model berekent men de multiële correlatie R en de significantie van het verschil tussen deze multiële correlaties kan via een F -toets nagegaan worden.

De uiteindelijke keuze voor een bepaald model berustte op de volgende procedure. Bij de log-lineaire analyse werd allereerst nagegaan of model 1 of 2 goed bij de data paste. Indien dit niet het geval was werd gekozen voor model 3, een zgn. verzadigd model dat per definitie een perfecte "fit" geeft. Indien model 1 verworpen moest worden, en model 2 niet, werd natuurlijk voor model 2 gekozen. Indien model 1 bevredigende resultaten te zien gaf (evenals model 2) werd nagegaan of model 2 significant "beter" was dan 1. Indien dit het geval was werd model 2 gekozen, in het andere geval model 1. Bij de regressieanalyse werden allereerst de drie multiële correlaties R_1 , R_2 en R_3 (behorende bij respectievelijk model 1, 2 en 3) onderling getoetst. In het geval R_3 significant "hoger" werd bevonden

dan R_1 of R_2 werd voor model 3 gekozen. In de overige gevallen werd voor model 1 gekozen indien R_2 niet significant afweek van R_1 , en voor model 2 indien dit wel het geval was.

Op deze wijze werd geprobeerd na te gaan in hoeverre interviewer-effecten "nodig zijn" ter verklaring van de antwoorden der respondenten.

Tot slot van deze paragraaf nog een drietal opmerkingen. De hierboven beschreven procedure is voor elke afhankelijke variabele afzonderlijk uitgevoerd. Een dergelijke werkwijze loopt uiteraard het risico geringe, maar "systematisch" aanwezige interviewer-effecten te missen.

De procedure is verder twee maal toegepast; in de ene variant was leeftijd een dichotomie, in de andere een trichotomie, dit omdat het willekeurig dichotomiseren van dergelijke variabelen ongewenste consequenties voor de uitkomsten kan hebben³⁾.

En, tenslotte, een bijkomende doelstelling van de uitgevoerde analyses was het vergelijken van de uitkomsten van de beide gebruikte analyse-technieken⁴⁾.

5. De resultaten

Allereerst is hieronder een overzicht gemaakt, voor de oorspronkelijke 38 variabelen, waarin vermeld staat hoe vaak gekozen is voor model 1, 2 of 3 (De analyse is ook uitgevoerd voor de op basis van deze 38 variabelen geconstrueerde somscores en factorscores, maar de uitkomsten hiervan waren sterk afhankelijk van de gekozen categorisering van deze variabelen!). Het overzicht vermeldt de resultaten afzonderlijk voor het geval leeftijd gedichotomiseerd dan wel getrichotomiseerd is en eveneens afzonderlijk voor twee significantieniveaus.

Tabel 1: Leeftijd als dichotomie; significantieniveau = 5%

	log-lineaire model				
		behaviourist...	tradition..	interact..	totaal
regressie-analyse met dummy's	behaviouristisch	28			28
	traditioneel	2	3		5
	interacties	2		3	5
	totaal	32	3	3	38

Tabel 2: Leeftijd als dichotomie; significantieniveau = 10%

regressie-analyse met dummy's	log-lineaire model				
		behaviourist.	tradition.	interact.	totaal
	behaviouristisch	22	2		24
	traditioneel	1	6		7
	interacties	1		6	7
	totaal	24	8	6	38

Tabel 3: Leeftijd als trichotomie; significantieniveau = 5%

regressie-analyse met dummy's	log-lineaire model				
		behaviourist.	tradition.	interact.	totaal
	behaviouristisch	25			25
	traditioneel	5	4		9
	interacties	1	1	2	4
	totaal	31	5	2	38

Tabel 4: Leeftijd als trichotomie; significantieniveau = 10%

regressie-analyse met dummy's	log-lineaire model				
		behaviourist.	tradition.	interact.	totaal
	behaviouristisch	18			18
	traditioneel	2	12		14
	interacties	2	1	3	6
	totaal	22	13	3	38

In zeer globale termen kan men stellen dat in ongeveer 2/3 van de gevallen het model 1 zonder interviewereffecten een bevredigende verklaring van de data geeft. In ongeveer 1/3 van de gevallen is er dus sprake van interviewereffecten.

Het is ondoenlijk in een kort bestek aan te geven, hoe deze interviewereffecten er precies uit ziet. Daarvoor moet verwezen worden naar het in voetnoot 1 genoemde onderzoeksrapport. Enkele grote lijnen zijn wel aan te geven:

- Jonge interviewers krijgen bij vragen rond pornografie en homofilie relatief meer progressieve antwoorden dan oudere interviewers. Dit leeftijdseffect is doorgaans sterker bij mannelijke dan bij vrouwelijke interviewers.
- De variabele geslacht speelde in de analyses slechts een zeer bescheiden rol. Een typische uitzondering hierop werd gevormd door de vragen die over abortus handelden. Bij deze vragen kregen vrouwelijke interviewers meer pro-abortus antwoorden dan mannelijke interviewers.
- De oudste groep respondenten gaf bij vrouwelijke interviewers aanmerkelijk meer tolerante antwoorden dan bij mannelijke interviewers.
- Bij sterk controversiële items lijkt het leeftijdsverschil tussen respondent en interviewer een belangrijke rol te spelen; hoe groter dit verschil, des te groter is de kans dat er interviewereffecten optreden.

Vanuit methodisch oogpunt is het tenslotte nog interessant dat, zoals ook uit bovenstaande tabellen blijkt, de log-lineaire analyse en de regressieanalyse in grote lijnen dezelfde uitkomsten te zien geeft. Natuurlijk zijn er wel vele detailverschillen. Zo blijkt ook hierboven dat met behulp van regressieanalyse vaker voor een "gecompliceerder" model gekozen wordt dan via de log-lineaire analyse. Maar daar speelt natuurlijk ook in mee dat de procedure om een bepaald model te kiezen in beide analysetechnieken niet identiek is.

6. Programma voor verder onderzoek

Dankzij een subsidie van de Nederlandse Organisatie voor Zuiver-Wetenschappelijk Onderzoek, Z.W.O., kunnen in de toekomst op het beschikbare materiaal nog een aantal analyses worden uitgevoerd. Aan de volgende vier thema's zal in dit verband nadere aandacht worden geschonken:

- a) Bestudering van de invloed van de interviewerkenmerken op de stabiliteit van antwoorden van respondenten in de tijd.
In feite staat hier de vraag centraal of respondenten die meermalen door dezelfde interviewer zijn ondervraagd, meer betrouwbare antwoorden geven, dan respondenten die telkens door verschillende interviewers zijn onderzocht.
Op grond van de tot nu toe uitgevoerde analyses, zou deze

vraag ontkennend beantwoord moeten worden. Bij de onderzochte variabelen worden geen verschillen van betekenis tussen beide groepen respondenten gevonden.

Vermeldenswaard in dit verband is nog dat in het kader van de instrumentontwikkeling ten behoeve van deze analysefase gezocht is naar latente waarschijnlijkheidsmodellen, waarbij aan de assumptie van lokale onafhankelijkheid niet hoeft te zijn voldaan⁵⁾.

- b) Bestudering van de invloed van interviewerkenmerken bij diverse typen van vraagformulering (open vs. gesloten vragen etc.).
- c) Bestudering van de correspondentie (of het ontbreken daarvan) tussen interviewerscores en respondenten scores op "dezelfde" variabelen.
- d) Toetsing van enige veronderstellingen die in de theoretische literatuur vermeld zijn.

Noten

1. Voor de specifieke vraagstelling en een uitgebreide weergave van de analyseresultaten zij de lezer verwezen naar: A. Heinen en J. Hagenaars (1977), "Invloed van interviewers; onderzoek naar de invloed van geslacht en leeftijd van interviewers op antwoorden van respondenten op vragen rond pornografie, homofilie en abortus", Tilburg.
2. Zie voor een beschrijving van het log-lineaire model bijvoorbeeld Bishop, Y., Fienberg, S. and Holland, P. (1975): "Discrete Multivariate Analysis", Cambridge, Mass., MIT Press, en Fienberg, S. (1977): "The Analysis of Cross-Classified Categorical Data", Cambridge, Mass., MIT Press.
Voor regressieanalyse met dummy-variabelen, zie o.a. Johnston, J. (1972): "Econometric Methods", Tokyo: McGraw-Hill, inc.
3. Zie:
Reynolds, H. (1977): "Some comments on the causal analysis of surveys with log-linear models", American Journal of Sociology, 83; 127-143.
4. Voor een discussie over de relatie tussen beide analysetechnieken, zie o.a.
Goodman, L. (1975): "The relationship between the modified and more usual multiple regression approach to the analysis of dichotomous variables". In: D.R. Heise (ed.): "Sociological Methodology", 1975, San Francisco.
Knoke, D. (1975): "A comparison of log-linear and regression models for systems of dichotomous variables". Sociological Methods and Research, vol. 3, no. 4, mei 1975.
Gillespie, M. (1977): "Log-linear techniques and the regression analysis of dummy-dependent variables: further bases for comparison". Sociological Methods and Research, vol. 6, no. 1, augustus 1977.
5. Voor een eerste aanzet hiertoe, zie J.A. Hagenaars (1978), "Latent Probability Models with direct effects between indicators". In: Quality and Quantity, jrg. 12, 205-222. Verdere studies hiervan worden nog ondernomen.