



EVALUATIE-ONDERZOEK NAAR HET EFFECT VAN EEN INTERVIEWTRAINING (*)

GEERT LOOSVELDT

1. Probleemstelling

In het onderzoek naar de kwaliteit van gegevens bekomen via het survey-interview kunnen verschillende onderzoeksstrategieën onderkend worden. Elke onderzoeksstrategie vertrekt vanuit een verschillende benadering van de kwaliteit van gegevens. Zo maakt het experimenteel onderzoek naar het effect van interviewtechnieken (Cannell e.a., 1979; Cannell e.a., 1981) vooral gebruik van evaluatiecriteria die zijn afgeleid uit het accuraatheidsonderzoek. Binnen het accuraatheidsonderzoek worden er uitspraken gedaan over de kwaliteit van gegevens door de antwoorden van de respondent te vergelijken met gegevens uit secundaire bronnen

* Dit artikel is gebaseerd op een deel van het doctoraatsproefschrift dat door de auteur werd verdedigd op 25 maart 1985. Promotor van het proefschrift was Prof. dr. J. Billiet.

waarvan verondersteld wordt dat ze juist zijn of in ieder geval minder onderhevig aan fouten (Parry en Crossley, 1950; Don Cahalan, 1968; Clausen, 1968; Traugott en Katosh, 1979; Andersen e.a., 1979). Dit impliceert dat de effectiviteit van de interviewtechnieken die tijdens een training werden aangeleerd, wordt beoordeeld door een analyse van bepaalde kenmerken van de antwoorden (hoeveelheid, belangrijkheid en sociale wenselijkheid).

Wanneer dit vertaald wordt naar het schema van een evaluatieonderzoek betekent dit dat bij het onderzoek naar het effect van bepaalde interviewtechnieken enkel een effectevaluatie wordt uitgevoerd en dat er geen procesevaluatie gebeurt. Bij een procesevaluatie ligt de nadruk op de wijze waarop een bepaald resultaat tot stand komt. Procesevaluatie heeft als bedoeling de interne validiteit van de effectevaluatie te verhogen (Freeman, 1977; Patton, 1980; Rutman, 1977). Het uitvoeren van een procesevaluatie bij het onderzoek naar het effect van interviewtechnieken komt neer op een analyse van de interactie tussen interviewer en respondent. In die zin wordt bij een procesevaluatie gebruik gemaakt van een kwaliteitsbenadering die zich richt op de gedragscontext waarbinnen een antwoord geregistreerd wordt. Daarbij wordt verondersteld dat adequaat gedrag of een adequaat verloop van een interactie een noodzakelijke voorwaarde is om te komen tot niet vertekende informatie.

Procesevaluatie en effectevaluatie vertrekken dus vanuit een verschillende kwaliteitsbenadering van gegevens. De verbinding van de resultaten van beide evaluatieve invalshoeken ontbreekt nagenoeg volledig in het tot nu toe verricht onderzoek naar de kwaliteit van gegevens bekomen via het survey-onderzoek.

In dit artikel zal geprobeerd worden om bij het bepalen van het effect van een interviewtraining op de kwaliteit van de gegevens, resultaten van een procesevaluatie in verband te brengen met kenmerken van de antwoorden om zodoende te komen tot wat Freeman noemt een "comprehensive evaluation" waarbij elementen van proces- en effectevaluatie met elkaar in verband gebracht worden.

2. Onderzoeksopzet

Om het effect van de interviewtraining op de geregistreerde antwoorden te meten werd een veldexperiment opgezet. De experimentele groep bestond uit respondenten die geïnterviewd werden door getrainde interviewers en een controlegroep die werd geïnterviewd door een groep niet-

getrainde interviewers.

Algemeen kan de interviewtraining omschreven worden als het aanleren van een gedragsrepertoire dat de interviewers moet toelaten om binnen verschillende actiestructuren die zich kunnen voordoen tijdens een interview, adequaat te handelen. Dit gedragsrepertoire structureert zich rond drie aandachtspunten : de elementaire interviewregels, het induceren van de rol van betrouwbare respondent door het gebruik van nauwkeurige en systematische instructies en het toedienen van negatieve en positieve feedback en ten derde reactiegedrag van de interviewer ten aanzien van inadequaat respondentengedrag. Er werd geopteerd voor een taakgerichte stijl van interviewen. Bij deze stijl gaat de interviewer een zakelijke relatie aan met de respondent en richt hij zijn aandacht volledig op het verwerven van informatie.

De feitelijke doelstelling en het daarmee verbonden experimenteel opzet van het onderzoek werden tijdens de training niet meegedeeld. Om te vermijden dat de interviewers daarvan ook naar enig vermoeden zouden krijgen werden tijdens de training weinig of geen voorbeelden gegeven van experimenteel onderzoek over het interview. Het onderzoek werd inhoudelijk voorgesteld als een onderzoek naar de leefsituatie van de gehuwde vrouw. De training bestond uit vijf bijeenkomsten van ongeveer drie uur.

Gezien het groot aantal factoren die een mogelijke invloed kunnen hebben op de geregistreerde antwoorden en dus als interveniërende variabelen kunnen optreden, werd geopteerd voor een homogene respondentengroep en interviewersgroep. Hiermee verhoogt de kans dat de verschillen tussen experimentele groep en controlegroep enkel veroorzaakt worden door het inwerken van de training. De respondentengroep bestond enkel uit gehuwde vrouwen tussen 22 en 55 jaar. Bij de selectie van de interviewers werden enkel vrouwen met weinig of geen interviewervaring weerhouden.

Om regiovariantie te vermijden werd de steekproef getrokken binnen een beperkt niet-stedelijk geografisch gebied (Hageland).

In totaal werden 285 respondenten geïnterviewd waarbij het interview op band werd opgenomen : 123 door zes getrainde interviewers en 162 door zeven niet getrainde interviewers.

3. Procesevaluatie : Analyse van het interactieverloop

Interactieanalyse van interviews gebeurt meestal op basis van de uitgeschreven bandopnamen van de interviews. Elke actie (taalhandeling) van

de interviewer en de respondent wordt opgevat als een variabele. De categorieën van elke variabele zijn specificaties van die acties. Wanneer de opeenvolgende acties van een vraag-antwoord sequentie worden weergegeven dan bekomt men een interactieketen voor een bepaalde vraag in een bepaald interview. Een vraag-antwoord sequentie begint met het stellen van een vraag en eindigt met de actie voorafgaand aan het stellen van de volgende vraag. De gezamenlijke weergave van de interactieketens met betrekking tot een bepaalde vraag of een bepaald soort vraag voor alle interviews gebeurt door middel van boomdiagrammen (cfr. figuur 1 en 2).

Het codeerschema dat voor de procesevaluatie werd opgesteld, sluit aan bij het schema van Cannell, Lawson en Hausser (1975) voor de evaluatie van interviewgedrag, het schema van Brenner (1982) met als uitgangspunt de conceptualisatie van het interview als regelgeleid handelen en het daarop geïnspireerde schema van Dijkstra (1983). De interactieanalyse werd uitgevoerd bij de eerste zes en de laatste vijf uitspraken van een reeks van 44 uitspraken in verband met ziekte en gezondheid (2). In de volgende paragraaf worden enkele resultaten van deze analyse weergegeven.

3.1. Interactieanalyse met betrekking tot elf uitspraken

Gezien de interactieanalyse gebeurt in het kader van een procesevaluatie waarbij het effect van een interviewtraining wordt nagegaan, is het duidelijk dat het interactieverloop in de getrainde en de niet getrainde conditie met elkaar vergeleken worden. Er wordt verwacht dat het interactieverloop in de getrainde conditie als "beter" kan beschouwd worden.

Algemeen gesteld gebeurt deze evaluatie vanuit de doelstelling van het interview nl. het op betrouwbare wijze bekomen van geldige informatie. Concreet betekent dit o.a. dat een interviewer de instructies en de uitspraken leest zoals ze in de vragenlijst staan, adequaat doorvragen stelt, geen antwoordmogelijkheden suggereert, een uitspraak indien nodig adequaat herhaalt of verduidelijkt en op een juiste wijze positieve feedback geeft. Adequaat respondentengedrag betekent bij het beantwoorden van een uitspraak voornamelijk een keuze maken uit de voorziene antwoordmogelijkheden en indien nodig vragen om de uitspraak te herhalen en te verduidelijken.

Bij de interactieanalyse werden enkel de 252 respondenten betrokken

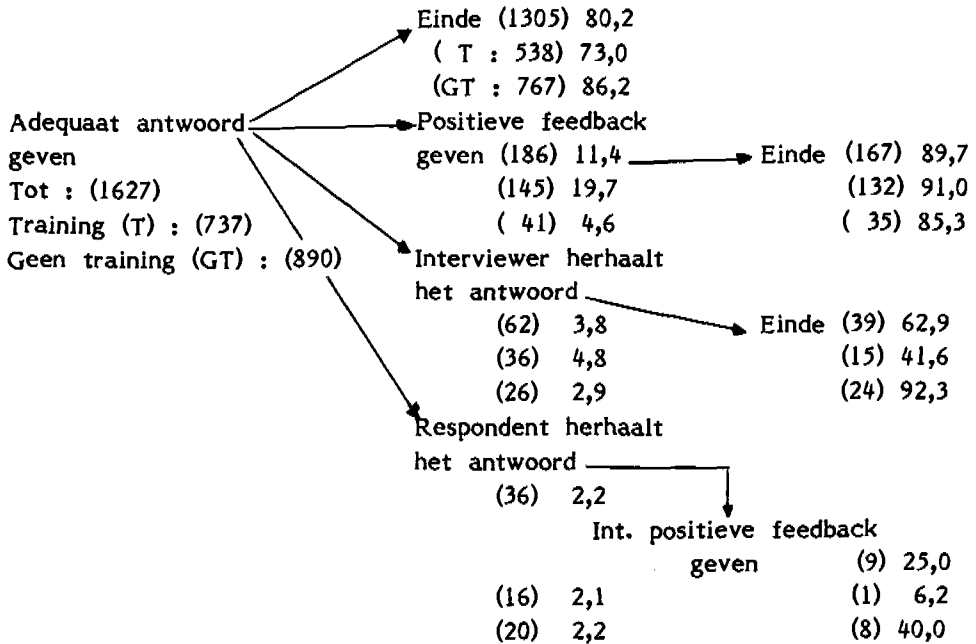
waarvan er informatie beschikbaar was voor alle elf uitspraken. Er werden dus 2772 vraag-antwoord sequenties geanalyseerd. Dit resulteerde in 669 verschillende interactieketens die allemaal starten met het adequaat lezen van een uitspraak. De maximale lengte werd vastgelegd op 20 acties. Dit vrij groot aantal interactieketens is het resultaat van de optie om het gedrag zo gedetailleerd mogelijk te coderen. Een overzichtelijke presentatie in een boomdiagram wordt door dit aantal echter onmogelijk. Soms wordt omwille van die overzichtelijkheid, de interactieketen afgebroken na 4 of 5 acties of na knooppunten met een frequentie die lager is dan 5 (Brenner, 1982) of 15 (Dijkstra, 1980). Een gevolg daarvan is dat interactieketens met vermoedelijk vrij unieke karakteristieken niet worden gerepresenteerd en uit de analyse verdwijnen. Interactieanalyse heeft in die zin veel weg van het kijken naar een landschap met een telescoop. Men doet veel moeite om alles gedetailleerd te bekijken maar men slaagt er niet in om een totaalbeeld op te hangen van hoe het landschap eruit ziet.

Uit de interactieanalyse blijkt dat na het voorlezen van een uitspraak in bijna 60% (3) onmiddellijk een adequaat antwoord volgt. Dit gebeurt in meerdere mate in de getrainde conditie (training : 64,4 %; geen training : 54,7 %). In de niet-getrainde conditie wordt een uitspraak daarentegen meer gevolgd door een gedeeltelijk antwoord (training : 16,8 %; geen training : 23,9 %). Hiermee zijn ook de twee meest voorkomende acties onmiddellijk volgend op het lezen van een uitspraak aangeduid.

Hoe ziet het verdere interactieverloop er nu uit nadat de respondent een adequaat antwoord heeft gegeven.

Uit figuur 1 blijkt dat in 80 % van de interacties het adequaat antwoord het einde betekent van de vraag-antwoord sequentie. Hiermee is dan ook de meest voorkomende interactieketen (47,08 % van alle interacties) omschreven : het lezen van een uitspraak gevolgd door een adequaat antwoord. In de getrainde conditie komt deze korte interactie iets minder voor omdat er meer positieve feedback gegeven wordt en het antwoord in iets grotere mate herhaald wordt. Deze versterkende reacties sluiten een groot aantal interacties af.

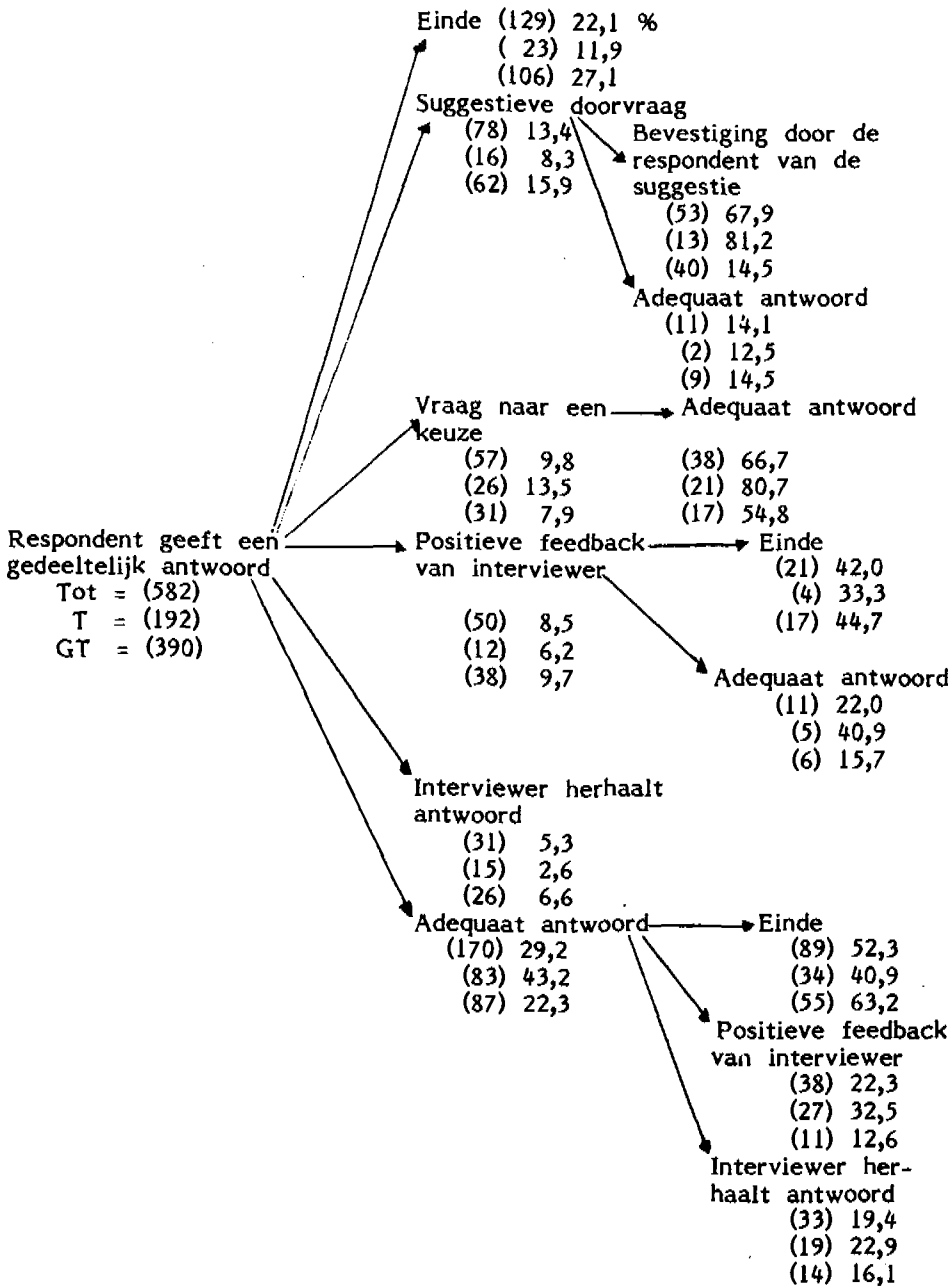
Figuur 1. Acties volgend op het geven van een adequaat antwoord (n = 1627) naar trainingsconditie (training : n = 737; geen training : n = 890) (absolute aantallen tussen de haakjes gevolgd door percenten) (4)



In figuur 2 is het interactieverloop weergegeven als de respondent na het lezen van de uitspraak een gedeeltelijk antwoord geeft. Bij een vergelijking met figuur 1 valt onmiddellijk op dat het interactieverloop veel minder eenduidig is. Dit blijkt onder andere uit de vaststelling dat de 4 reacties in figuur 1 (niet reageren inbegrepen) 97,7 % van de reacties omvatten. De 6 reacties in figuur 2 zijn goed voor 88,5 % van de reacties (niet reageren inbegrepen).

In de niet-getrainde conditie wordt in 27,18 % van de gevallen de vraag antwoord sequentie afgesloten met een onvolledig antwoord. In de getrainde conditie gebeurt dit in mindere mate (11,9 %). Hiermee wordt duidelijk hoe vooral in de niet getrainde conditie het vrij grote percentage uitspraken gegenereerd wordt waarbij de respondent niet komt tot een adequaat antwoord. De interviewster leest de uitspraak waarop de

Figuur 2. Acties volgend op het geven van een gedeeltelijk antwoord (n= 582) naar trainingsconditie (training : n = 192; geen training : n = 390)



respondent reageert met een gedeeltelijk antwoord. Het is dan wellicht de interviewer die zelf een antwoordcategorie kiest. Er moet hierbij opgemerkt worden dat in de vragenlijst op een paar uitzonderingen na, telkens een van de voorziene antwoordcategorieën werd aangeduid.

Het onvolledig antwoord is in meer dan één vierde van de interacties (meer bij getrainde dan bij niet-getrainde) een tussenstap naar een adequaat antwoord. De verschillen tussen de getrainde en de niet getrainde conditie in reactie t.a.v. dit adequaat antwoord zijn vergelijkbaar met de bevindingen in figuur 1.

Opvallend is verder het verschillend doorvraaggedrag tussen de condities. Men reageert in ongeveer gelijke mate met een doorvraag maar de verhouding tussen een suggestie doen en de vraag om een antwoord te kiezen is duidelijk verschillend en kan in de getrainde conditie als beter beschouwd worden. Dit is ook het geval voor het geven van positieve feedback en het herhalen van een antwoord.

Deze resultaten van de procesevaluatie illustreren duidelijk dat een interview geen aaneenschakeling is van vragen en antwoorden, maar dat zelfs een op het eerste zicht vrij eenvoudige taak binnen een interview resulteert in een grote diversiteit van interacties waarin heel wat meer gebeurt dan vragen stellen en antwoorden geven.

Verder blijkt dat de experimentele condities verschillen zowel qua interviewer- als qua respondentengedrag en dat de interviewtraining inwerkt bij de getrainde interviewsters. De interviewpraktijk van de getrainde interviewsters is op een aantal punten beter dan die van hun niet-getrainde collega's. De wijze waarop de antwoorden bekomen worden, verschilt tussen de condities. De verschillen kunnen echter niet gesteld worden in termen van aan- of afwezigheid van bepaalde acties maar wel in termen van *meer of minder*. De verschillen zijn *m.a.w.* *gradueel*.

Alvorens de relatie tussen informatie van de procesevaluatie en kenmerken van de antwoorden te analyseren lijkt het daarom aangewezen om de verschillen tussen de condities wat betreft het interactieverloop verder te preciseren.

3.2. Onderlinge vergelijkbaarheid van het interactieverloop per trainingsconditie

De vraag of alle interviewsters binnen een conditie onderling vergelijkbaar en van de andere groep onderscheidbaar gedrag vertonen, is hier aan de orde. Om deze vraag te beantwoorden werd gebruik gemaakt van correspondentieanalyse (5).

De kruistabel die hier geanalyseerd wordt, is gevormd door de variabele

interviewster (13 categorieën) en de variabele interactieverloop. De categorieën van deze laatste variabele zijn de verschillende interactieketens. Gegeven het groot aantal unieke ketens en ketens met een lage frequentie worden enkel interactieketens opgenomen met een frequentie groter dan 25. Het gaat om volledige interactieketens. De acht interactieketens die in tabel 1 worden beschreven bevatten 65 % van de interacties. Om de niet opgenomen interacties een zeker gewicht te geven in de analyse, werd een categorie "andere" mee in de analyse opgenomen (categorie I).

Tabel 1. Omschrijving van de gebruikte interactieketens

Interactieketens	AC
A : adequaat antwoord geven	1304
B : adequaat antwoord geven door positieve feedback van de interviewer	167
C : adequaat antwoord geven gevolgd door herhaling van het antwoord door de interviewer	39
D : adequaat antwoord geven gevolgd door een herhaling door de respondent	26
E : gedeeltelijk antwoord	129
F : suggestie van een antwoord en bevestiging daarvan door de respondent	32
G : gedeeltelijk antwoord gevolgd door een adequaat antwoord	89
H : gedeeltelijk antwoord gevolgd door een adequaat antwoord gevolgd door positieve feedback van de interviewer	32

Indien alle interviewsters op een gelijke wijze interviewen betekent dit dat de verdeling van de variabele interactieverloop voor alle interviewers gelijk is aan de marginale proportionele verdeling van die variabele. De afwijkingen (afstand) van die conditionele verdeling t.o.v. de marginale verdeling (zwaartepunt) is dus een eerste indicatie van de mate waarin een interviewer afwijkt van het doorsnee interviewgedrag. In tabel 2 (onder interactieverloop) zijn de chi-kwadraat afstanden tussen de interviewsters en het zwaartepunt weergegeven.

Er kan worden aangetoond dat deze afstand een bijdrage is tot de chi-kwadraat van de tabel. Gerelateerd aan de totale chi-kwadraat bekommt men dus de proportionele bijdrage van een interviewer tot de totale associatie in de tabel.

Tabel 2. Chi-kwadraat afstanden tussen interviewsters tot hun zwaartepunt en de proportionele bijdrage tot de totale associatie (interactieverloop en antwoordresultaat)

Interviewster (*)	chi-kwadraat afstand		proportionele bijdrage	
	interactie- verloop	antwoord- resultaat	interactie- verloop	antwoord- resultaat
T1	0,12	0,02	0,05	0,01
T2	0,07	0,12	0,03	0,06
T3	0,09	0,05	0,03	0,02
T4	0,17	0,12	0,07	0,06
T5	0,62	0,06	0,24	0,03
T6	0,10	0,02	0,04	0,02
GT1	0,09	0,01	0,03	0,01
GT2	0,14	0,06	0,05	0,03
GT3	0,26	0,05	0,10	0,03
GT4	0,42	0,98	0,16	0,50
GT5	0,36	0,32	0,13	0,16
GT6	0,08	0,13	0,03	0,06
GT7	0,11	0,01	0,04	0,01

(*) T : training; GT : geen training

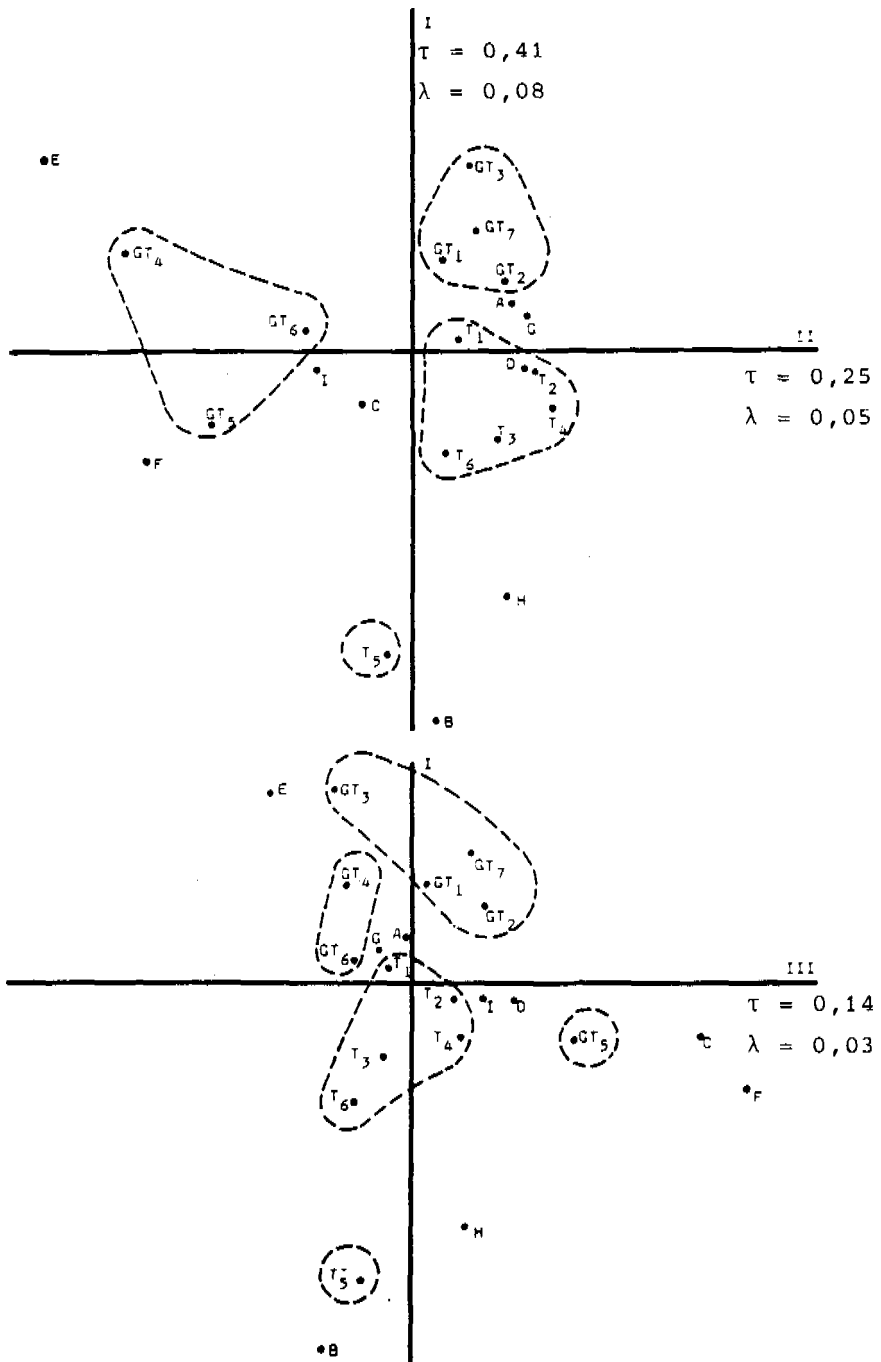
De getrainde interviewer T5 en de niet-getrainde interviewers GT3, GT4 en GT5 realiseren gezamenlijk 63 % van de associatie. Vooral de getrainde interviewer T5 draagt in vergelijking tot de andere interviewers aanzienlijk bij tot de samenhang in de tabel.

De resultaten wijzen ook in de richting dat een aantal getrainde en niet-getrainde interviewers een vergelijkbaar interactieverloop kennen. Dit wordt ondersteund door de resultaten van een hiërarchische clusteranalyse (Ward's algoritme) op de afstanden tussen de verschillende interviewsters.

Uit de clusteranalyse blijkt dat de getrainde interviewsters, behalve T5, vrij vlug samengevoegd worden in één cluster. Bij de niet-getrainde interviewsters gebeurt de clustering in meerdere stappen. Zo wordt de cluster met vier niet-getrainde interviewsters (GT1, GT2, GT3, GT7) eerst samengevoegd met de cluster van de getrainde en pas daarna worden de drie andere niet-getrainde (GT4, GT5, GT6) daaraan toegevoegd.

De representatie van de tabel gevormd door het interactieverloop en de interviewster, in een drie-dimensionele ruimte (81,8 % van de associatie)

Figuur 3. Representatie van de interviewsters naar interactieketens in een drie-dimensionele ruimte (6)



biedt een mogelijkheid om deze clusters inhoudelijk te duiden (figuur 3).

De eerste as in figuur 3 kan gerelateerd worden aan interviewergedrag en in het bijzonder aan het geven van feedback. Hieruit blijkt o.a. dat T5 zich duidelijk van de rest onderscheidt door het geven van feedback. De cluster getrainde interviewers situeren zich duidelijk aan de kant die geïnterpreteerd kan worden als het regelmatig geven van feedback.

De tweede as wordt duidelijk bepaald door interacties waarbij de respondent niet komt tot een adequaat antwoord of waarbij de interviewer hierop blijkbaar anticipeert door onmiddellijk een antwoord te suggeren dat dan door de respondent bevestigd wordt. In die zin wijst de tweede as op een zekere respondentenbekwaamheid en op de moeilijkheid om bepaalde respondenten te interviewen. Drie niet-getrainde interviewers onderscheiden zich op deze as.

De derde as kan, wanneer het geven van feedback buiten beschouwing wordt gelaten, geïnterpreteerd worden in termen van intervieweractiviteit. De as loopt van afwezigheid van interviewergedrag over het herhalen van het antwoord tot het stellen van suggestieve doorvragen, wat als overactiviteit kan beschouwd worden. De niet getrainde interviewsters GT5 scoort vrij hoog op deze as.

Opmerkelijk is wel dat de cluster van de getrainde interviewsters steeds rond de oorsprong ligt terwijl de niet-getrainde grotere spreiding hebben op de assen.

Een probleem bij de gevolgde werkwijze is wel dat de interpretatie moet gebeuren aan de hand van 65 % van de interactieketens. Wanneer het lezen van de uitspraken buiten beschouwing wordt gelaten betekent dit dat slechts ongeveer één vierde van het totaal aantal gecodeerde acties gebuikt wordt. De categorie "andere" bevat m.a.w. heel wat informatie om te komen tot een evaluatie van het interactieverloop tussen interviewer en respondent. De moeilijkheid om alle informatie die vervat zit in alle ketens te gebruiken, komt hiermee terug aan de orde. Om informatie over alle interactieketens in de analyse te betrekken en aldus het gestelde probleem enigszins te ondervangen werden alle ketens beoordeeld op hun antwoordresultaat. Er werden drie mogelijkheden onderscheiden :

- de interactieketen bevat een antwoord dat als adequaat kan beschouwd worden (82,8 % van de interacties)
- de interactieketen resulteert in een vermoedelijk vertekend antwoord (10,3 %)
- de interactieketen bevat slechts een deel van de antwoordmogelijkheden (6,9 %).

Bijna één vijfde van de interactieketens bevat dus geen adequaat antwoord. Bij de getrainde interviewsters gebeurt dit in 8,39 % van de interactieketens, bij de niet-getrainde in 23,4 % van de interacties.

Uit de resultaten van de correspondentieanalyse op de tabel gevormd door de variabele interviewster en de variabele antwoordresultaat blijkt dat de interviewsters GT4 en GT5 de grootste afwijkingen vertonen en twee derden van de totale associatie vertegenwoordigen (cfr. tabel 1, onder antwoordresultaat).

Het clusterpatroon van de interviewsters komt op een iet of wat andere wijze tot stand dan bij de clusteranalyse met de variabele interactieverloop. De niet-getrainde interviewsters GT1, GT2 en GT7 worden in vergelijking met die clusteranalyse vlugger bij de getrainde interviewsters samengevoegd. Ook de getrainde interviewster T5 die zich onderscheidde door het geven van positieve feedback behoort tot deze cluster.

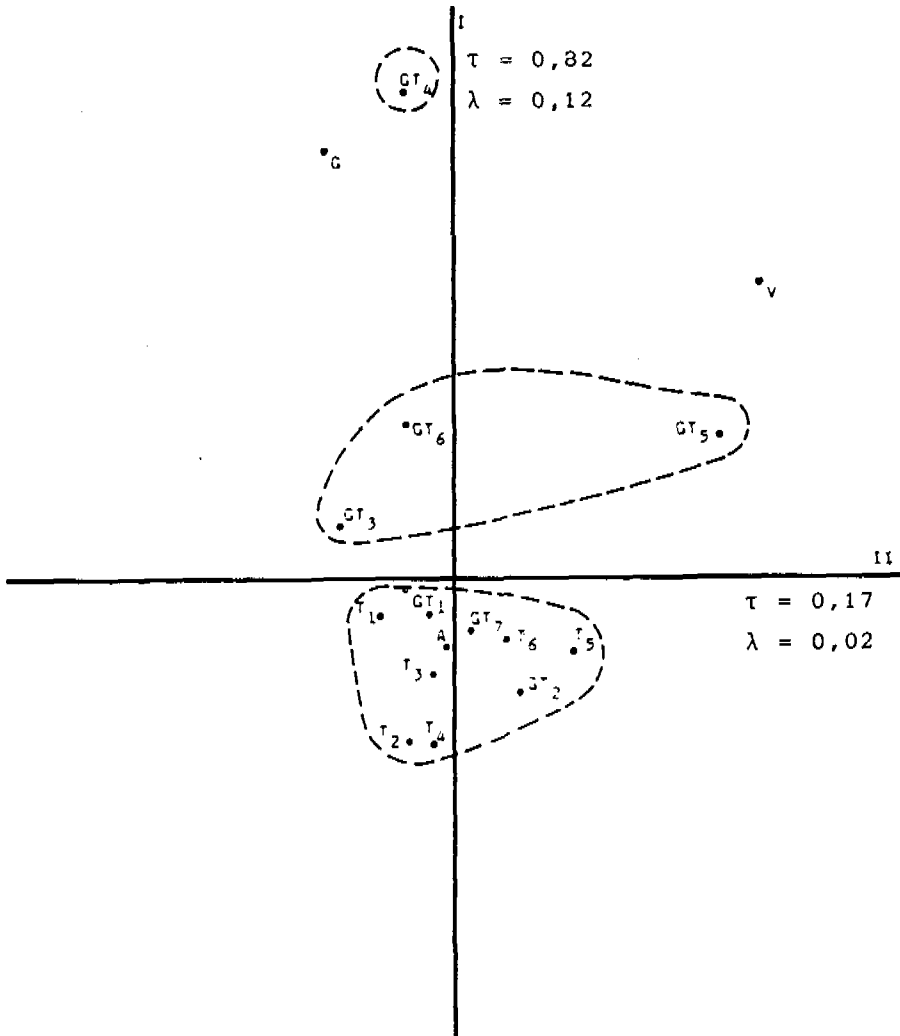
De grafische voorstelling van de tabel (cfr. figuur 4) maakt het terug mogelijk om het interviewgedrag binnen de clusters te duiden.

De eerste as van de grafische voorstelling (82 % van de associaties) kan beschouwd worden in termen van het al of niet komen tot een adequaat antwoord met als extreme punten een adequaat antwoord en geen antwoord. De cluster met de 9 interviewsters situeert zich aan de adequate zijde van de as. De tweede as wordt bepaald door het vertekend karakter van het antwoord. De niet-getrainde interviewster GT5 die ook in de eerste analyse kon opgemerkt worden door een zekere overactiviteit scoort op deze as vrij hoog. Ook hier blijkt terug dat sommige niet-getrainde interviewsters (GT1, GT2, GT7) zich weinig onderscheiden van de getrainde interviewsters.

Op basis van de resultaten van de analyse van de onderlinge vergelijkbaarheid van het interactieverloop per trainingsconditie kan gesteld worden dat de interviewtraining zoals verwacht het interviewgedrag standaardiseert. De getrainde interviewsters interviewen op ongeveer dezelfde wijze. De niet-getrainde interviewsters verschillen onderling sterker van elkaar dan de getrainde interviewsters. Sommige van de niet-getrainde interviewsters (GT1, GT2, GT7) onderscheiden zich echter weinig van de getrainde interviewsters.

De vraag die zich nu opdringt is of de vastgestelde verschillen en gelijkenissen in de wijze waarop de antwoorden worden bekomen, gerelateerd kan worden aan inhoudelijke kenmerken van de antwoorden.

Figuur 4. Presentatie van de interviewsters naar antwoordresultaat in een twee-dimensionele ruimte



3.3. Relatie procesevaluatie en bekomen antwoorden

Gegeven de verschillen die tijdens de procesevaluatie werden vastgesteld tussen de trainingscondities wat betreft het interactieverloop bij de 11 uitspraken kan, wanneer wordt aangenomen dat de wijze waarop geïnterviewd wordt de geregistreerde antwoorden kan beïnvloeden, bij de antwoorden op deze uitspraken een trainingseffect verwacht worden. Uit variantieanalyses met de trainingscondities als onafhankelijke variabele en de antwoorden op de uitspraken als afhankelijke variabele (niet opgenomen) blijkt echter dat dit niet het geval is. Er zijn, behalve bij één uitspraak, geen significante verschillen tussen de antwoorden.

Wanneer op basis van de informatie van de procesevaluatie de niet-getrainde interviewsters opgedeeld worden in twee groepen en er dus met drie groepen wordt gewerkt, dan zou er kunnen verwacht worden dat het verschil tussen de getrainde conditie en de groep niet-getrainde interviewsters die een vergelijkbaar interactieverloop kennen (GT1, GT2, GT7) kleiner is dan het verschil tussen deze laatste groep en de overige niet-getrainde interviewsters. Bij 7 van de 11 uitspraken ligt het gemiddelde van groep 2 dicht bij het gemiddelde van de getrainde groep dan bij het gemiddelde van de andere niet-getrainde condities (cfr. tabel 3).

Tabel 3. Gemiddelde score voor de elf uitspraken naar trainingsconditie en naar drie groepen interviewsters (afzonderlijk)

uitspraak	training wel	geen tr. = tr.	geen train.	SS	df	F	Prob
uitspraak 1	1,74	1,57	1,51	2,54	2	< 1	n.s
uitspraak 2	1,29	1,32	1,03	3,76	2	1,06	n.s
uitspraak 3	1,57	2,06	1,74	1,74	2	2,72	0,06
uitspraak 4	1,89	1,87	2,17	4,80	2	1,23	n.s
uitspraak 5	1,60	1,65	1,47	1,38	2	< 1	n.s
uitspraak 6	1,37	1,15	1,54	5,29	2	1,37	n.s
uitspraak 7	0,61	0,68	0,49	1,36	2	< 1	n.s
uitspraak 8	2,27	1,88	2,29	7,45	2	< 1	n.s
uitspraak 9	1,82	1,65	1,71	1,28	2	< 1	n.s
uitspraak 10	2,44	2,25	2,52	2,80	2	< 1	n.s
uitspraak 11	0,72	0,76	0,81	0,38	2	< 1	n.s
n	104	63	84				

Er dient echter opgemerkt te worden dat behalve bij uitspraak 3 er geen significante verschillen tussen de gemiddelden werden vastgesteld. De resultaten zijn echter niet van die aard om te stellen dat groep 2 een tussenpositie inneemt.

Dit alles leidt tot de merkwaardige vaststelling dat juist bij die vragen (uitspraken) waarbij de interactieanalyse werd uitgevoerd en op basis waarvan gesteld werd dat beide condities verschillen qua interactieverloop er geen effect van de training op de antwoorden kan worden vastgesteld.

Zelfs wanneer informatie van de procesevaluatie expliciet in de analyse werd opgenomen (werken met drie groepen) werden eveneens geen effecten vastgesteld.

De gelijkensis in interactieverloop die tijdens de procesevaluatie werd vastgesteld van de drie niet-getrainde interviewsters met de getrainde groep, kan dus niet doorgetrokken worden naar een gelijkensis in bekomen antwoorden bij de 11 uitspraken.

Een directer werkwijze om het verband tussen interactieverloop en antwoorden te analyseren bestaat erin om kenmerken van het interactieverloop te relateren aan de antwoorden. Op basis van het aantal adequate antwoorden op de 11 uitspraken (NAA) werd het interactieverloop als vlot (meer dan 8 adequate antwoorden) of als moeizaam (8 en minder adequate antwoorden) beschouwd.

Kan er nu een verband gelegd worden tussen deze beoordeling van het interactieverloop en de antwoorden op de uitspraken ?

Tabel 4. Gemiddelde scores op 11 uitspraken naar interactieverloop

uitspraak	interactieverloop		SS	df	F	Prob
	vlot	moeizaam				
uitspraak 1	1,73	1,27	10,00	1	6,11	0,01
uitspraak 2	1,30	1,04	3,08	1	1,74	n.s
uitspraak 3	1,70	1,88	1,48	1	< 1	n.s
uitspraak 4	1,98	1,98	0,00	1	< 1	n.s
uitspraak 5	1,66	1,29	6,49	1	4,13	0,04
uitspraak 6	1,36	1,41	0,13	1	< 1	n.s
uitspraak 7	0,67	0,32	5,52	1	6,64	0,01
uitspraak 8	2,18	2,19	0,00	1	< 1	n.s
uitspraak 9	1,73	1,79	0,16	1	< 1	n.s
uitspraak 10	2,40	2,50	0,46	1	< 1	n.s
uitspraak 11	0,77	0,70	0,22	1	< 1	n.s
n	190	62				

Bij drie uitspraken (1, 5 en 7) is er een significant verschil tussen de vlotte en de moeizame conditie. Inhoudelijk kunnen er bij deze drie uitspraken niet direct verschillen onderkend worden.

In de vlotte conditie ligt de gemiddelde score bij die drie uitspraken telkens hoger. Dit wijst op een zekere systematiek in de antwoordpatronen. Het onderzoek van de relatie tussen het interactieverloop en voorkomen van bepaalde antwoordcategorieën lijkt dan ook zinvol (tabel 5).

Tabel 5. Gemiddeld aantal van voorkomen van een antwoordcategorie naar interactieverloop

Antwoord categorie	Interactieverloop		SS	df	F	Prob
	vlot	moeizaam				
- volledig eens	2,71	3,72	48,18	1	10,76	0,00
- gedeeltelijk eens	3,71	2,67	50,40	1	14,68	0,00
- geen mening	1,43	1,62	1,82	1	< 1	n.s
- gedeeltelijk oneens	1,57	1,25	4,81	1	2,09	n.s
- volledig oneens	1,54	1,66	0,60	1	< 1	n.s
N	190	62				

Wanneer de interactie tekenen van een moeizaam verloop vertoont resulteert dat in significant meer uitspraken waarmee men volledig eens is. Bij een vlot verloop daarentegen antwoordt men significant meer dat men het gedeeltelijk eens is. De nuance van gedeeltelijk is dus sterker aanwezig wanneer de interactie vlot verloopt. Hetzelfde patroon, maar in erg afgezwakte vorm is te vinden bij de variabelen die betrekking hebben met oneens zijn. Een moeizaam verloop van de interactie gaat dus gepaard met een tendens naar de extreme antwoordmogelijkheden (vooral volledig eens = de "ja-zeggars").

Deze meer directe koppeling van informatie van de procesevaluatie met de antwoorden leert dus wel iets over de wijze waarop het interactieverloop het antwoordresultaat bepaalt. Wel dient hier nog opgemerkt te worden dat wanneer de trainingscondities mee in de analyse worden opgenomen er geen interactie-effecten worden vastgesteld.

Tot nu toe werd enkel geprobeerd om kenmerken van het verloop van een onderdeel van het interview in verband te brengen met de antwoorden die tijdens dit deel van het interview werden geregistreerd. Het

lijkt echter zinvol om na te gaan of de informatie van de procesevaluatie van een deel van het interview gerelateerd kan worden aan antwoorden op vragen die niet in de procesevaluatie waren betrokken. Indien dit zou mogelijk zijn betekent dit dat een beperkte interactieanalyse zou volstaan om het interactieverloop in zijn geheel te beoordelen en te relateren aan de bekomen antwoorden.

Om dit te onderzoeken werd gebruik gemaakt van antwoorden die betrekking hebben op een open en een gesloten vraag naar klachten op gebied van ziekte en gezondheid gedurende de laatste drie weken, twee open vragen naar voordelen en nadelen van de woonomgeving en ja-nee vragen, eventueel gevolgd door doorvragen, naar tijdschriften en aankopen op krediet.

Op basis van de bevindingen van het accuraatheidsonderzoek waarbij voornamelijk onderrapportering werd vastgesteld, kan gesteld worden dat meer informatie op die vragen kan beschouwd worden als betere informatie. De interviewtechnieken, die tijdens de training werden aangeleerd, zijn trouwens gericht om de onderrapportering tegen te gaan.

Er werd dan ook verwacht dat in de getrainde conditie een hoger percentage respondenten zou zijn dat minstens een feit of ervaring m.b.t. de bevroegde onderwerpen zou rapporteren en dat het gemiddeld aantal feiten of ervaringen hoger zou liggen (7). De verschillen in bekomen antwoorden bij deze vragen tussen de trainingscondities werden op basis hiervan geëvalueerd. Hierdoor werd het mogelijk om een uitspraak te doen over de relatie tussen de kwaliteit van de interviewpraktijk en de kwaliteit van de antwoorden. Dit kon in veel mindere mate gebeuren met de antwoorden op de uitspraken omdat de evaluatie van die antwoorden in termen van kwaliteit niet zo duidelijk is.

Verder werd ook de non-respons ("weet niet", "geen antwoord" en blanco) bij vragen in verband met partijvoorkeur, aspecten van de financiële situatie en seksuele relatie, in de analyse betrokken (volledigheid).

Uit de effectevaluatie blijkt (Loosveldt, 1985 : 121-147) dat in de getrainde conditie een "zwaarder" klachtenprofiel wordt geregistreerd dan in de niet-getrainde conditie : er zijn meer klagers en de klagers hebben meer klachten. Ook bij de andere vragen wordt er meer informatie bekomen in de getrainde conditie. Het verschil situeert zich hier niet zozeer in het al of niet vermelden van een feit en gebeurtenis maar wel in de frequentie ervan. In het licht van de bevindingen van het accuraatheidsonderzoek zouden deze gegevens van een betere kwaliteit zijn. Verder was ook de non-respons aanzienlijk kleiner in de getrainde conditie (8).

Gegeven de resultaten van de procesevaluatie en de resultaten van de effectevaluatie lijkt het aannemelijk om te stellen dat een betere interviewpraktijk resulteert in betere antwoorden. Deze conclusie gebeurt echter door het achtereenvolgens en afzonderlijk uitvoeren van een proces- en effectevaluatie. De vraag die terug kan gesteld worden is of informatie van de procesevaluatie direct kan gerelateerd worden aan de antwoorden op de vragen die bij de effectevaluatie werden gebruikt. Daarvoor werd terug onderzocht of de niet-getrainde interviewers die weinig verschillen qua interactieverloop (bij de 11 uitspraken) met de getrainde interviewsters, antwoorden bekomen die vergelijkbaar zijn met die groep.

Op basis van de percentages respondenten in de drie groepen interviewsters (cfr. tabel 6) kan enkel bij de variabele volledigheid gesteld worden dat de middengroep (geen training = training) aanleunt bij de getrainde conditie. Bij de variabele klachten (open) is er nog een klein verschil tussen de middengroep en de groep met de andere niet getrainde interviewsters. Bij de overige variabelen kan zeker niet gesteld worden dat de bekomen antwoorden van de middengroep meer gelijken op de getrainde groep dan op de niet getrainde groep. De gemiddelde aantallen van de kenmerken vertonen een ander patroon. Daar neemt de middengroep (in zekere mate) een tussenpositie in. Bij het paarsgewijze testen van de gemiddelden is die tussenpositie echter enkel bij het aantal klachten (open) duidelijk te onderkennen omdat het gemiddelde van de middengroep niet significant verschilt van de getrainde groep maar ook niet significant verschilt van de niet getrainde groep. Bij de variabele onvolledigheid is er geen significant verschil tussen de gemiddelden van de getrainde groep en de middengroep. Bij het gemiddeld aantal voordelen, nadelen en tijdschriften is er geen verschil tussen de middengroep en de niet getrainde conditie. Bij deze variabelen is er dus enkel een significant verschil van de getrainde groep met de twee andere groepen. Bij de overige twee variabelen (krediet aankopen en klachten gesloten) zijn er geen significante verschillen tussen de drie groepen.

Uit deze resultaten blijkt dus dat slechts bij twee variabelen de informatie van de procesevaluatie van een bepaald onderdeel van het interview in de verwachte zin kan gerelateerd worden aan de antwoorden van vragen die niet bij de procesevaluatie betrokken waren.

Tabel 6. Percentage respondentent met een bepaald kenmerk naar drie groepen interviewsters

	training	= geen train.	tr. geen train.	chi- kwadraat	df	Prob
volledigheid	65,3	60,3	44,0	9,01	2	0,01
klachten open	53,8	38,1	34,1	8,32	2	0,01
klachten gesloten	85,5	80,9	79,7	1,22	2	n.s
voordelen	98,0	93,6	95,1	2,21	2	n.s
nadelen	67,3	73,0	70,2	0,62	2	n.s
tijdschriften	87,3	84,1	84,3	0,48	2	n.s
kredietaanbopen	29,7	18,0	22,6	3,02	2	n.s
n	104	63	84			

Gemiddelde aantallen van een bepaald kenmerk naar drie groepen interviewsters

Het verband tussen interactieverloop en antwoorden kan tenslotte terug directer onderzocht worden aan de hand van de beoordeling van het interactieverloop als vlot of moeizaam.

	training	= geen training	train geen training	F	df	Prob
onvolledigheid	1,52(36)	1,72(25)	2,68(47)	7,86	2	<0,01
klachten open	1,71(56)	1,41(24)	1,25(28)	3,70	2	<0,03
klachten gesloten	3,89(89)	4,13(51)	3,74(67)	<1	2	n.s
voordelen	2,80(102)	2,40(59)	2,10(79)	8,78	2	<0,01
nadelen	1,95(70)	1,60(46)	1,37(59)	6,77	2	<0,01
tijdschriften	3,04(90)	2,22(53)	2,31(70)	5,77	2	<0,01
kredietaanbopen	1,51(31)	1,18(11)	2,21(19)	1,09	2	n.s

Uit tabel 7 (percentages) blijkt dat het verband tussen dit kenmerk en het al of niet registreren van een bepaald kenmerk van de respondent onbestaande is. Het grootste verschil tussen een vlot en een moeizaam interactieverloop wordt zoals verwacht, gevonden bij de variabele volledigheid. Wanneer de variabele trainingsconditie aan de analyse wordt toegevoegd, wordt enkel bij de variabele klachten (open) een interactie-effect vastgesteld. In de getrainde conditie ligt het percentage klagers het hoogst bij een vlot interactieverloop (55,3 % tegenover 40 %). Bij de niet getrainden gebeurt net het omgekeerde. Als de interactie vlot verloopt worden er 30 % klagers geregistreerd, bij een moeizaam verloop stijgt dit percentage tot 46 %. Dit betekent dus dat het trai-

ningseffect bij deze variabele het grootst is bij een vlot verloop. Dit wekt de indruk dat een vlot verloop bij de niet getrainde interviewsters vooral een snelle wijze van interviewen in de hand werkt.

Het beeld van de gemiddelde aantallen van deze kenmerken is terug ietwat duidelijker. Wanneer de interactie op basis van een beperkt onderdeel van het interview als moeizaam wordt bestempeld dan blijkt de non-respons bij vragen die niet in de interactieanalyse werden betrokken hoger te zijn dan bij een vlot verloop van de interactie. Er werden verder geen interactie-effecten vastgesteld wanneer de trainingsconditie in de analyse wordt opgenomen.

Tabel 7. Percentage respondenten met een bepaald kenmerk naar interactieverloop

	Interactieverloop		chi- kwadraat	df	Prob
	vlot	moeizaam			
volledigheid	58,9	50,8	1,24	1	n.s
klachten open	42,6	45,1	0,12	1	n.s
klachten gesloten	82,1	83,6	0,07	1	n.s
voordelen	95,7	96,6	0,09	1	n.s
nadelen	69,3	70,9	0,06	1	n.s
tijdschriften	86,7	81,6	0,96	1	n.s
krediet aankopen	23,1	28,3	0,66	1	n.s
n	190	62			

Gemiddelde aantallen van een bepaald kenmerk naar interactieverloop

	Interactieverloop		F	df	Prob
	vlot	moeizaam			
onvolledigheid	1,89(78)	2,53(30)	4,03	1	0,04
klachten open	1,53(81)	1,51(27)	< 1	1	n.s
klachten gesloten	3,73(156)	4,45(51)	2,85	1	0,09
voordelen	2,55(182)	2,22(58)	3,59	1	0,06
nadelen	1,72(131)	1,50(44)	1,91	1	n.s
tijdschriften	2,76(164)	2,06(49)	6,81	1	<0,01
krediet aankopen	1,41(43)	1,22(18)	< 1	1	n.s

4. Besluit

Op basis van de procesevaluatie bij 11 uitspraken werd gesteld dat het interactieverloop tussen de trainingscondities verschilt. Er werd vastgesteld dat de training het interviewgedrag op een positieve wijze beïnvloedt. De interviewpraktijk in de getrainde conditie werd als beter beschouwd. Bij verdere analyse was het echter mogelijk om een groep niet getrainde interviewsters te detecteren waarbij het interactieverloop vergelijkbaar was met dat van de getrainde interviewsters. De antwoorden die deze groep bekomt vertonen echter slechts een geringe gelijkenis met die van de getrainde interviewsters. De gelijkenis in interactie resulteert dus niet in een grotere gelijkenis van de antwoorden. Anders uitgedrukt betekent dit dat een min of meer gelijke beoordeling van de kwaliteit van de interviewpraktijk niet overeenstemt met een zelfde beoordeling van de kwaliteit van de antwoorden. Er dient hierbij opgemerkt te worden dat bij de effectevaluatie de antwoorden op deze vragen in de getrainde conditie als beter beschouwd worden.

Hoewel de relatie tussen een bepaald kenmerk van het interactieverloop (vlot - moeizaam) en de bekomen antwoorden niet kan ontkend worden, zijn de resultaten ook hier niet erg overtuigend. Wel zijn er aanwijzingen dat kenmerken van het interactieverloop van een bepaald onderdeel van het interview veralgemeenbaar zijn naar andere vragen (cfr. non-respons). Dit dient echter nog verder onderzocht.

Uit dit alles blijkt dat de verschillende kwaliteitsbenaderingen die binnen een proces- en effectevaluatie gehanteerd worden, niet zonder meer met elkaar in verband kunnen worden gebracht en tot eenduidige resultaten leiden. Dit betekent ook dat de op het eerste zicht evidente stelling dat de vastgestelde verschillen tussen de condities bij de effectevaluatie tot stand komen door verschillende wijzen van interviewen, minder aannemelijk is.

SAMENVATTING

In het onderzoek naar de kwaliteit van gegevens bekomen via het survey-interview kunnen verschillende onderzoeksstrategieën onderkend worden. In dit artikel wordt geprobeerd om bij het bepalen van het effect van een interviewtraining, verschillende kwaliteitsbenaderingen op elkaar te betrekken. Hiervoor werd gebruik gemaakt van het onderzoeks-schema van een evaluatieonderzoek.

De gegevens werden verzameld binnen een veldexperiment waarbij de experimentele groep bestond uit 123 respondenten die geïnterviewd werden door zes getrainde interviewers en een controlegroep (n = 162) die werd geïnterviewd door 7 niet getrainde interviewers.

Uit de procesevaluatie blijkt dat de interviewpraktijk van de getrainde interviewers beter is. Uit de resultaten van de effectevaluatie blijkt ook dat de antwoorden op een aantal vragen als beter kunnen worden beschouwd. De analyse van de relatie tussen informatie uit de procesevaluatie en kenmerken van de antwoorden laat echter niet toe te besluiten dat het hanteren van verschillende kwaliteitsbenaderingen zonder meer tot eenduidige resultaten leiden.

ABSTRACT

The research concerning the quality of survey data uses several research strategies.

In this article an attempt is made to integrate - by determination of the effects of an interviewtraining - different approaches of quality of survey data. Therefore the research scheme of an evaluation research was used.

The data were collected during a field experiment in which the experimental group consisted of 123 respondents who were interviewed by 6 trained interviewers and a control group (n = 162) who was interviewed by 7 untrained interviewers.

According to the results of the process evaluation the interviewer performance of the trained interviewers is better.

There are also indications that - to some questions - trained interviewers obtain answers of a better quality.

The analysis of the relation between information resulting from process evaluation and the characteristics of the obtained answers may not lead to the conclusion that the application of different approaches of quality leads - without any exception - to unambiguous results.

NOTEN

- 1) Bij het onderzoeksontwerp werd verondersteld dat het op band opnemen van een interview effect kan hebben op de antwoorden (Belson, 1967 : 258). Om een verwarring tussen het effect van de training en de bandopname te vermijden werd het experimenteel opzet uitgebreid met twee bijkomende condities (wel en geen bandopname). Op deze wijze ontstonden er vier experimentele condities : training met band, training zonder band, geen training met band en geen training zonder band. Deze structuur maakte het mogelijk om zowel het effect van de training en de bandopname afzonderlijk als het interactie-effect tussen beide variabelen na te gaan. Gegeven de probleemstelling van dit artikel en het feit dat procesevaluatie slechts mogelijk is als het interview op band is opgenomen, wordt de analyse hier enkel uitgevoerd bij de respondenten waarvan het interview werd opgenomen.
- 2) De antwoordmogelijkheden waren : volledig eens - gedeeltelijk eens - geen mening - gedeeltelijk oneens en volledig oneens. Een keuze van een van deze antwoordmogelijkheden is een adequaat antwoord. Een antwoord waarbij het niet mogelijk is om zowel de richting van de opinie als de sterkte van de opinie te bepalen wordt als een gedeeltelijk antwoord beschouwd.
- 3) Gezien de frequentie van voorkomen van acties gerelateerd is aan het aantal uitspraken, worden de percentages berekend op het aantal uitspraken : training : $n = 1144$; geen training : $n = 1628$.
- 4) De percentages worden steeds berekend op de overeenkomstige categorie uit de voorgaande actie, bv. $(538/737) \times 100 = 73 \%$ van de interacties eindigt na een adequaat antwoord in de getrainde conditie.
- 5) Deze techniek analyseert een contingentietabel door de categorieën van de gekruiste variabelen voor te stellen in een laag dimensionale ruimte, zodat bepaalde kenmerken van de structuur van de tabel visueel zichtbaar worden. Twee aspecten zijn hier belangrijk : de mate waarin de verdeling van een variabele binnen een categorie van de andere variabele afwijkt van de marginale verdeling en de mate waarin deze conditionele verdelingen verschillen tussen twee categorieën (Matthijs en Daemen, 1985). Om alle interviewsters een gelijk gewicht te geven werd de analyse uitgevoerd op de tabel

waarbij de frequenties gedeeld werden door het aantal respondenten per interview.

- 6) De betekenis van de letters stemt overeen met de beschrijving in tabel 1 (I = de categorie andere).
- 7) In de analyse werd een onderscheid gemaakt tussen het al of niet aanwezig zijn van een kenmerk en de frequentie van voorkomen ervan indien het kenmerk aanwezig te is. Hierdoor werd het onderscheid gemaakt tussen afwezigheid van informatie en aanwezigheid van weinig (detail) informatie.
- 8) Bij de effectevaluatie werden alle respondenten, ook diegene waarbij het interview niet op band werd opgenomen, gebruikt.

BIBLIOGRAFIE

ANDERSEN, R., J. KASPER en FRANKEL, M.R.

1979 Total survey error, London, Jossey-Bass.

BELSON, W.A.,

1967 Tape recording : its effect on accuracy of response in survey interviews, in Journal of Marketing Research, 4 : 253-260.

BRENNER, M.,

1982 Response-effects of 'role-restricted' characteristics of the interviewer, in W. Dijkstra en J. Van der Zouwen, Response behavior in survey-interview. London, Academic Press : 131-165.

CAHALAN, D.,

1968 Correlates of respondent accuracy in the Denver validity survey, in Public Opinion Quarterly, 32 : 607-621.

CANNELL, C.F., OKSENBERG, L, en CONVERSE, J.M.,

1975 A technique for evaluating interviewer performance, Survey Research Center, The University of Michigan.

CANNELL, C.F., LAWSON, S.A., en HAUSSER, D.L.

1979 Experiments in interviewing techniques. Field experiments in health reporting 1971-1977, SRC, Ann Arbon.

- CANNELL, C.F., MILLER, P.V., en OKSENBURG, L.,**
 1981 Research on interviewing techniques, in S. Leinhardt (ed.),
Sociological Methodology. San Francisco, Jossey-Bass : 389-437.
- CLAUSEN, A.R.,**
 1968 Response validity : vote report, in Public Opinion Quarterly,
 32 : 588-605.
- DIJKSTRA, W.,**
 1983 Beïnvloeding van antwoorden in survey-interviews. Utrecht,
 Elinkwijk.
- DIJKSTRA, W., VAN DER VEEN, L., en VAN DER ZOUWEN, J.,**
 1980 Effects of interviewer style on response behaviour in survey-
 interviews. Paper presented at the Symposium "Uses of the
Research Interview" at the University of Surrey.
- FREEMAN, H.E.,**
 1977 The Present Status of Evaluation Research, in Evaluation Studies
Review annual, 2 : 17-51.
- LOOSVELDT, G.,**
 1985 De effecten van een interviewtraining op de kwaliteit van
gegevens bekomen via het survey-interview. Leuven, Faculteit
 der Sociale Wetenschappen, Departement Sociologie, Doctoraal
 proefschrift.
- MATTHIJS, K., en DAEMEN, L.,**
 1985 Correspondentie-analyse als methode voor het analyseren van de
 structuur van contingentietabellen, in Kwantitatieve Methoden,
 17 : 5-37.
- PARRY, H.J., en CROSSLEY, H.M.,**
 1950 Validity of responses to survey questions, in Public Opinion
Quarterly, 14 : 61-80.
- PATTON, M.,**
 1980 Qualitative evaluation methods. London, Sage Publications.
- RUTMAN, L., (ed.)**
 1977 Evaluation research methods : a basic guide. London, Sage
 Publications.

TRAUGOTT, M.W., en KATOSH, J.P.,
1979 Response validity in surveys of voting behaviour, in Public
Opinion Quarterly, 43 : 359-377.