

**COMPUTER GESTUURD ENQUETEREN:
TELEFONISCH OF PERSOONLIJK?
een literatuurstudie**

Ger J.M.E. Snijkers *)

Samenvatting

De persoons- en gezinsenquêtes die het Centraal Bureau voor de Statistiek (CBS) houdt, worden afgenomen met behulp van een drietal interviewmethoden: het face-to-face interview bij de respondent aan huis, het telefonisch interview, en de schriftelijke vragenlijst, die de respondent zelfstandig invult. Deze interviewmethoden kunnen op de klassieke manier worden gehanteerd met pen-en-papier of met behulp van de computer. In dit artikel wordt computer gestuurd telefonisch enquêteren (CATI) op grond van een literatuurstudie vergeleken met computer gestuurd persoonlijk enquêteren (CAPI), onder andere met betrekking tot kosten en kwaliteit van de gegevens. Om enig inzicht te krijgen in de consequenties die een overgang van pen-en-papier naar computer gestuurd enquêteren heeft voor de ontwerpfase, het veldwerk en de verwerkingsfase zullen eerst deze laatste twee interviewmethoden met elkaar worden vergeleken.

Trefwoorden: computer gestuurd enqueteren (CAI), CATI, CAPI.

*) Centraal Bureau voor de Statistiek, Hoofdafdeling Statistische Methoden.
Postbus 959, 2270 AZ VOORBURG. Tel.: 070-337 4924.

De in dit artikel weergegeven opvattingen zijn die van de auteur en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

1. Inleiding

De onderzoeken die het Centraal Bureau voor de Statistiek (CBS) houdt, worden afgenomen met behulp van een drietal interviewmethoden: het face-to-face (of persoonlijk) interview, waarbij een enquêtrice bij de respondent aan huis komt, het telefonisch interview, waarbij de respondent door een enquêtrice telefonisch wordt ondervraagd, en de schriftelijke vragenlijst, waarbij de respondent een vragenlijst zelfstandig invult. Deze interviewmethoden kunnen op de klassieke manier worden gehanteerd met pen-en-papier (de PAPI- Paper And Pencil Interviewing -methoden) of, zoals tegenwoordig steeds meer gebeurt, met behulp van de computer (de CAI- Computer Assisted Interviewing -methoden). De relatie tussen deze interviewmethoden is weergegeven in tabel 1 (Sikkel, 1986; Saris, 1989). Deze interviewmethoden kunnen ook gecombineerd worden, bijvoorbeeld in een panel-onderzoek waarbij na een werving en/of eerste ronde met de face-to-face methode voor de doorgangers wordt overgeschakeld op de telefonische methode.

Tabel 1. Interviewmethoden

PAPI (Paper And Pencil Interviewing) *)	CAI (Computer Assisted Interviewing)
face-to-face interviewing	CAPI (Computer Assisted Personal Interviewing)
telefoon interviewing	CATI (Computer Assisted Telephone Interviewing), ook wel COTEL genoemd
schriftelijke vragenlijst	CAMI (Computer Assisted Mail Interviewing) of het Tele-interview

*) De in deze kolom vermelde begrippen worden verder in dit artikel gedefinieerd als interviewmethoden met pen en papier.

Veel CBS-onderzoeken worden met de face-to-face of CAPI-methode gehouden. Deze methoden zijn nogal kostbaar vanwege de (reis)kosten van de enquêtrices. In de literatuur over enquêteren staat veelvuldig te lezen dat telefonisch interviewen kostenbesparend is en dat bovendien de gegevens sneller beschikbaar zijn (Nicholls & Groves, 1986; Kerssemakers, 1985; Kerssemakers, De Mast & Remmerswaal, 1987; Bronner, 1980). In dit artikel zullen we bezien of dat ook geldt voor CATI ten opzichte van CAPI. Dit doen we door middel van literatuurstudie. Daarbij zal de nadruk liggen op CBS-literatuur en CBS-ervaringen.

Kerssemakers (1985) geeft een overzicht van de stand van zaken van telefonisch enquêteren binnen het CBS. Datzelfde doen Kerssemakers, De Mast & Remmerswaal (1987), De Mast (1990a) en De Heer, Akkerboom en Israëls (1990) specifiek gericht op de non-respons ontwikkeling. Van Eck & Kazemier (1990) geven een uitgebreide beschrijving van een onderzoek naar zwart inkomen. Daarin gaan zij onder andere voor verschillende interviewmethoden na of deze geschikt zijn voor het onderzoeken van een gevoelig onderwerp als zwart inkomen.

We geven daarnaast een selectie van literatuur van buiten het CBS. De Leeuw (1989a, 1990) beschrijft beknopt de verschillen tussen de drie interviewmethoden aan de hand van een meta-analyse van literatuur en een uitgevoerd experiment. Nicholls & Groves gaan in twee uitgebreide artikelen in op de stand van zaken voor CATI met betrekking tot kosten (Nicholls & Groves, 1986) en de kwaliteit van de data (Groves & Nicholls, 1986). Körmendi & Noordhoek (1989) bespreken de resultaten van een aantal experimenten (onderzoek naar de werkomgeving en het absentisme bij de Deense PTT, 1983/84, en een omnibus-enquête, 1984) waarin telefonisch interviewing wordt vergeleken met face-to-face interviewing. Bij dit onderzoek moet worden opgemerkt dat de respondenten in het algemeen zeer gemotiveerd waren om aan de onderzoeken mee te werken. Steel & Boal (1988) gaan in op de telefonische bereikbaarheid in Australië. Wilson, Blackshaw & Norris (1988) geven een evaluatie van het telefonisch uitgevoerde Britse 'Labour Force Survey'. Catlin, Ingram & Hunter (1988) doen hetzelfde voor het Canadese 'Labour Force Survey'. Aquilino & Lo Sciuto (1990) beschrijven de effecten van de interviewmethode (telefoon en face-to-face) op drug- en alcoholrapportage in New Jersey (VS) in 1986-'87.

In paragraaf 2 worden de consequenties van computer gestuurd enquêteren ten opzichte van de traditionele pen-en-papier methode beschreven. In paragraaf 3 wordt ingegaan op de vergelijking CATI-CAPI. In paragraaf 4 wordt de conclusie geformuleerd. De algemene vragen achter dit artikel zijn: Wat zijn de voor- en nadelen van CATI ten opzichte van CAPI en in welke situaties kan CATI worden gebruikt?

2. *CAI versus PAPI*

Het is moeilijk om uit de bestaande literatuur een vergelijking tussen CATI en CAPI te destilleren. Veel literatuur gaat niet over de vergelijking van CATI en CAPI maar over CATI en PAPI. Bijgevolg wordt meer aandacht besteed aan Computer Assisted Interviewing (CAI) ten opzichte van de klassieke pen-en-papier methode dan aan telefonisch in vergelijking tot persoonlijk enquêteren met behulp van de computer.

Deze verschillen tussen CAI en PAPI worden in alle artikelen meer of minder gedetailleerd beschreven (zie o.a. Nicholls & Groves, 1986; Groves & Nicholls, 1986; House, 1985; Kerssemakers, 1985). Omdat bij CAI andere eisen worden gesteld aan het ontwerp van de vragenlijst, is het noodzakelijk de kenmerken van CAI (in vergelijking tot PAPI) hier beknopt te vermelden (Sikkel, 1988; De Leeuw, 1989b; Saris, 1989, 1990).

De voordelen van CAI zijn:

- 1) De mogelijkheid om een complexe routingstructuur te implementeren, die in het veld door de computer wordt bestuurd, waardoor ingewikkelde theoretische concepten (beter) gemeten kunnen worden. Dit betekent echter wel dat het veldwerk aan flexibiliteit verliest: er zijn bijna geen uitwijkmogelijkheden, en de vragenlijst is gesegmenteerd doordat vraag voor vraag op het scherm verschijnt. De enquêtrice verliest hierdoor het overzicht. Als een enquêtrice in een verkeerd deel van de vragenlijst terecht komt is het haast ondoenlijk om weer op het rechte pad terug te keren.

- 2) De mogelijkheid om interne controles (edits) op de antwoorden in te bouwen, die er voor zorgen dat elk formulier intern consistent is. Een probleem hierbij is echter de eventuele controle achteraf; als in het veld niet wordt bijgehouden welke antwoorden welke consistentie checks activeren en welke antwoorden daarna worden veranderd, kan achteraf niet meer worden nagegaan of er toch iets mis is gegaan; de enquêtrices zijn niet altijd inhoudelijk deskundig en het antwoord van de respondent moet (bij gesloten vragen) in een van de antwoordcategorieën passen.
- 3) De mogelijkheid om de formulering van vragen variabel te maken, zodat de vragen een persoonlijker karakter krijgen en toegespitst worden op de persoonlijke situatie van de respondent. Dit vergroot de betrouwbaarheid van de antwoorden. Samen met de automatische routing en de consistency checks zorgen deze elementen voor een betere kwaliteit van CAI-data.
- 4) De mogelijkheid om snel en eenvoudig veranderingen in een bestaande vragenlijst door te voeren. In de brontekst van het interviewprogramma worden de veranderingen aangebracht, waarna de vragenlijst opnieuw wordt gecompileerd. Verspreiding van het nieuwe interviewprogramma kan via diskettes of via een modem plaatsvinden. Bij grotere veranderingen en aanpassingen in de routing is het echter noodzakelijk dat het interviewprogramma opnieuw wordt getest.
- 5) Reductie van de benodigde tijd voor de verwerking van de gegevens en controle op de antwoorden achteraf, aangezien de gegevens direct in de computer worden ingevoerd. Samen met de mogelijkheid om rechtstreeks in de vragenlijsttekst veranderingen aan te kunnen brengen, levert dit een aanzienlijke besparing in papierwerk, tijd en geld.
- 6) De mogelijkheid om bij panelstudies gebruik te maken van de in eerdere ronden verzamelde gegevens.
- 7) De mogelijkheid om antwoorden op open vragen automatisch te coderen of te berekenen.

Een nadelig punt van CAI is de langere ontwikkelingstijd van het interviewprogramma. Het ontwerpen van een vragenlijst voor CAI houdt namelijk in dat de aspecten van elke fase uit het PAPI-ontwerpproces een stap eerder aan bod komen (Kerssemakers, 1985; De Leeuw, 1989b):

- Van veldwerk naar ontwerpfase:

Aangezien het verloop van een interview met de computer veel gestructureerder en rigider is dan een interview met pen en papier en de controles tijdens het veldwerk plaatsvinden, moet de ontwerper van een computer gestuurde vragenlijst inhoudelijk van het onderwerp van het onderzoek op de hoogte zijn en moet hij zich perfect kunnen inleven in elke voorkomende situatie, zowel voor wat de respondent als wat de enquêtrice betreft. Zowel de routing als de consistentie checks moeten zeer accuraat worden gedefinieerd, waarbij voorkomen moet worden dat de enquêtrice het spoor bijster raakt (ze heeft immers geen overzicht over de vragen, door de routing en de beperkte schermgrootte), en correcties moeten eenvoudig kunnen worden aangebracht. Verder moeten de routingvragen duidelijk en ondubbelzinnig zijn en de antwoordcategorieën van de routingvragen daadwerkelijk uitputtend en uitsluitend. Is de vragenlijst, inclusief routing en consistentie checks, helemaal klaar, dan ligt daarmee ook de hele structuur van het uiteindelijke data-bestand vast. Dit heeft ook consequenties voor de analyses.

- Van verwerkings- en correctie-fase naar veldwerk:

Deze verschuiving van activiteiten betreft in de eerste plaats de invoer van de gegevens in de computer, wat niet achteraf gebeurt maar tijdens het veldwerk plaatsvindt. Dit betekent intoetswerk voor de enquêtrices, waar ze misschien niet aan gewend zijn. Dit kan bovendien extra meetfouten introduceren (verkeerd intoetsen van antwoorden) door het verloren gaan van de coördinatie van hand en oog. In de tweede plaats vindt de correctie van inconsistente antwoorden tijdens het veldwerk plaats (op grond van in de ontwerpfase gedefinieerde consistentie checks) en niet meer in een iteratief correctie-proces na afloop van het veldwerk. In principe betekent dit dat we na beëindiging van het veldwerk beschikken over (per record) consistente gegevens, die direct beschikbaar zijn voor analyse.

Gezien de beperkingen/mogelijkheden die de computer de ontwerper van een computer gestuurde vragenlijst oplegt/geeft, is het niet aan te raden om domweg een papieren vragenlijst te vertalen naar een computer gestuurde vragenlijst. Immers, zoals gebleken is met betrekking tot de definiëring van de controle-structuur en het personifiëren van de vraagteksten, voor het maken van een interviewprogramma is meer informatie nodig dan in een papieren vragenlijst aanwezig is. Bovendien zal de opzet van de gehele vragenlijst aangepast moeten worden.

Ook de omgekeerde weg levert problemen op. Aangezien de routingstructuur van een computer gestuurde vragenlijst vele malen ingewikkelder kan zijn dan de routingstructuur van een papieren vragenlijst is het heel moeilijk om een vragenlijst uit de computer, inclusief routing, netjes op papier te zetten. Daarenboven is het voor de uitvoering van het veldwerk praktisch onmogelijk om de controle-structuur en variabele vraagteksten te converteren naar papier.

3. CATI versus CAPI

Naast de computer is ook de gekozen interviewmethode (telefonisch of face-to-face) van invloed op het ontwerp van de vragenlijst. De interviewmethode wordt onder andere bepaald door het onderwerp van het onderzoek en heeft consequenties voor de kosten van het onderzoek en de kwaliteit van de gegevens. In de literatuur wordt echter nog weinig ingegaan op specifieke aspecten die voortvloeien uit telefoon dan wel face-to-face computer gestuurd interview. Bovendien zijn er nauwelijks experimenten beschikbaar waarin op een systematische wijze CATI en CAPI met elkaar worden vergeleken.

In paragraaf 3.1 zullen we de kenmerken van CATI in relatie tot die van CAPI vermelden, voor zover ze betrekking hebben op de interviewmethode. Daarna worden in de paragrafen 3.2 en 3.3 CATI en CAPI systematisch met elkaar vergeleken voor wat kosten en kwaliteitsaspecten van de data betreft.

3.1. Kenmerken van CATI

- 1) Het belangrijkste kenmerk van CATI heeft betrekking op het gebruikte communicatiemedium: met de telefoon is het alleen mogelijk om via de spraak te communiceren. Dit heeft een aantal beperkingen tot gevolg.
 - In de eerste plaats is het moeilijk om visuele hulpmiddelen, zoals antwoordkaarten en ander toonmateriaal, te gebruiken. Dit in tegenstelling tot CAPI. Een mogelijke oplossing om bij een telefonisch interview toch toonmateriaal te gebruiken is het versturen van het toonmateriaal samen met de aanschrijfbrief, met het verzoek om het materiaal naast de telefoon klaar te leggen. Ook kan schriftelijk materiaal worden meegezonden, dat gedurende een bepaalde periode wordt ingevuld. De informatie wordt dan tijdens het telefonisch interview verzameld.
 - In de tweede plaats zijn, door het ontbreken van visueel contact, pauzes in het telefoongesprek hinderlijk, zowel voor respondent als enquêtrice. Om de voortgang van het gesprek niet te veel te belemmeren zal de respondent wellicht te weinig bedenktijd nemen voor het formuleren van antwoorden. Het opzoeken van gegevens is hierdoor vrijwel onmogelijk, evenals het stellen van vragen die een zwaar beroep doen op het geheugen. Voor de enquêtrice betekent het vermijden van pauzes dat ze wellicht te weinig tijd neemt om antwoorden op open vragen volledig in te typen, waardoor de antwoorden cryptisch kunnen worden (Catlin, Ingram & Hunter, 1988; Körmendi & Noordhoek, 1989). Dit betekent tevens dat de gesprekssnelheid van een telefoongesprek groter is dan van een face-to-face gesprek (Kerssemakers, 1985). Telefonische interviews zijn dan ook slechts geschikt voor vragen waarop direct, zonder bedenktijd geantwoord kan worden.
 - En tot slot kan het ontbreken van direct visueel contact ook gevolgen hebben voor de antwoorden die gegeven worden, wat betreft sociaal-wenselijke antwoorden of onderrapportage van gevoelige onderwerpen. Welke kant dit opwerkt is onduidelijk (Groves, 1989). Volgens De Leeuw (1989a) komen bij telefonisch interviewing sociaal-wenselijke antwoorden en onderrapportage (item non-respons) bij gevoelige onderwerpen iets meer voor dan bij face-to-face interviewing. Körmendi & Noordhoek (1989) vonden echter minder onderrapportage voor gevoelige onderwerpen bij telefonisch interviewing. Sykes & Collins (1988) vonden geen verschillen tussen beide methoden.

Als verklaring voor dit onduidelijke beeld zegt Groves (1989) dat de telefoon weliswaar een grotere sociale afstand creëert, maar dat het moeilijker is om de respondent te overtuigen van de legitimiteit van de enquêtrice. Deze twee aspecten werken in tegengestelde richting. Welk aspect de doorslag geeft is moeilijk te zeggen, waarschijnlijk spelen hierbij sociaal-culturele achtergronden een rol. Zo melden Aquilino & Lo Sciuto (1990) slechts kleine verschillen tussen beide methoden in onderzoek naar drug- en alcoholgebruik voor blanken (iets lager voor telefoon) in New Jersey (VS), terwijl het telefonisch interview voor zwarten aanzienlijk lagere schattingen opleverde.

- 2) CATI heeft een aantal voordelen met betrekking tot kosten en kwaliteit van de gegevens ten opzichte van CAPI. Een aantal van deze voordelen geldt slechts wanneer de enquêtegesprekken centraal vanuit een belzaal worden gevoerd, zoals bij het CBS gebeurt (Kerssemakers, 1985).
- Bij CATI wordt bij het CBS gewerkt op een netwerk met een beperkt aantal PC's. De hardware voor CATI hoeft slechts één keer te worden aangeschaft en is daarna beschikbaar voor alle onderzoeken en alle enquêtrices (die elkaar aflossen). Bovendien kunnen op hetzelfde netwerk meerdere interviewprogramma's tegelijkertijd draaien. Bij CAPI heeft elke enquêtrice haar eigen handheld-computer, waarop weliswaar meerdere interviewprogramma's kunnen draaien (één per diskette), maar niet tegelijkertijd. Bij CATI worden de hardware en de software dan ook efficiënter gebruikt.
 - Er zijn geen kosten voor het verzenden van het enquête-materiaal naar de enquêtrices en voor het terugsturen van de gegevens. Bovendien zijn de data sneller beschikbaar. Natuurlijk zijn er, net als bij CAPI, verzendkosten voor de aanschrijfbrieven, gericht aan de respondenten.
 - Er zijn geen reiskosten. Dit aspect wordt voor CATI als het meest kostenbesparend genoemd. Onder andere hierdoor wordt in het algemeen aangenomen dat de kosten van telefonisch enquêteren lager zijn dan van face-to-face enquêteren (Kerssemakers, 1985; De Leeuw, 1989a; Van Eck & Kazemier, 1990; Groves, 1990). In plaats van de reiskosten tellen bij CATI de telefoonkosten. De telefoonkosten zullen hoger zijn naarmate de vragenlijst langer is, en daarmee het interview gemiddeld gesproken langer zal duren.
 - Er zijn minder enquêtrices nodig. Voor eenzelfde aantal respondenten, bij gelijke vragenlijsten voor CAPI en CATI, is het aantal ondervraagde personen per enquêtrice bij CATI groter dan bij CAPI, aangezien er geen tijd verloren gaat met reistijd.
 - Er zijn geen problemen met de verdeling van enquêtrices over het land.
 - De verdeling van de interviews over de enquêtrices kan daardoor aselekt plaatsvinden, waardoor eventuele enquêtrice-effecten niet systematisch tot regionale verschillen leiden.
 - Bovendien bestaat via de telefoon een andere relatie tussen respondent en enquêtrice. De enquêtrice heeft wel enig effect op het interview, maar algemeen wordt aangenomen dat de enquêtrice-variantie in een telefonisch interview lager is dan in een face-to-face interview (uitgaande van een gelijk aantal respondenten per enquêtrice) (Pannekoek, 1988; De Leeuw, 1989a).
 - Door het werken op een netwerk in een centrale belzaal is een permanente supervisie mogelijk, zodat het interviewproces en het enquêtricedrag voortdurend kan worden bijgesteld.
- 3) Een derde belangrijk aspect van CATI zijn de eventuele niet-waarnemingsfouten die ontstaan door onderdekking.

Onderdekking heeft betrekking op het te gebruiken steekproefkader. Het steekproefkader moet de doelpopulatie dekken, dat wil zeggen dat mensen die tot de doelpopulatie behoren ook in het steekproefkader moeten voorkomen. Een probleem bij

het steekproefkader voor telefonisch enquêteren is dat niet iedereen in Nederland een telefoonaansluiting heeft. Bovendien zijn er mensen met een geheim nummer.

Uit CBS-onderzoek (Kerssemakers, De Mast en Remmerswaal, 1987) blijkt dat jongeren, ouderen en werklozen door een geringer telefoonbezit slechter bereikbaar zijn. Uit de Gezondheidsenquête 1987 (Van den Berg, 1987) volgt bovendien dat relatief meer gescheiden personen, lager opgeleide personen en mensen in de lagere inkomensgroepen geen telefoonaansluiting hebben. Om dezelfde reden concluderen Steel & Boal (1988) voor Australië dat het beter is niet te kiezen voor de telefoon als lage inkomensgroepen, jonge mensen, en mensen in huurwoningen tot de doelpopulatie behoren. In Engeland (Wilson, Blackshaw & Norris, 1988) blijkt het aantal telefoon-aansluitingen onder werklozen, uitgesplitst naar leeftijd en economische activiteit, systematisch lager te zijn dan onder werkenden. Groves (1989) concludeert op grond van het face-to-face National Health Interview Survey 1985-86, dat er minder telefoonaansluitingen zijn bij jongeren tot 25 jaar, de lagere inkomensgroepen (minder dan \$15 000), niet-blanken (vooral negers), alleenstaanden of gescheidenen van tafel en bed, en in landelijke gebieden (waar een telefoon-abonnement duurder is). Wat de geheime nummers betreft volgt uit de Gezondheidsenquête 1987 (Van den Berg, 1987) dat relatief meer gescheiden personen, mensen met een lager opleidingsniveau, mensen in de lagere en mensen in de hogere inkomensgroepen een geheim nummer hebben.

Als de groepen die telefonisch moeilijker bereikbaar zijn een (belangrijk) deel van de doelpopulatie vormen, kan er door deze onderdekking selectiviteit in de respons optreden, resulterend in vertekening van schattingen van de doelvariabelen, die in het algemeen slechts ten dele kan worden gecorrigeerd door te herwegen.

Voor telefonische onderzoeken worden bij het CBS eerst adressen getrokken uit een steekproefkader van adressen/postafgiftepunten van de PTT. Daarna worden bij deze adressen de telefoonnummers (zonder geheime nummers) gezocht. Deze methode heeft tot voordeel dat, ter completering van het onderzoek, adressen zonder telefoon of met een geheim nummer altijd op een andere wijze (persoonlijk of schriftelijk) kunnen worden benaderd (Kerssemakers, 1985). Om niet voor elk telefonisch onderzoek een parallel-onderzoek te moeten houden, kan overwogen worden om met een bepaalde frequentie bij de niet telefonisch bereikbare adressen een omnibus-enquête te houden.

- 4) De telefoon als communicatiemedium stelt verder nog een aantal eisen aan de ontwikkeling van de vragenlijst:
 - Het wordt afgeraden om lange en ingewikkelde vraagteksten te formuleren, aangezien mensen dat niet kunnen onthouden (De Leeuw, 1989a; Van Eck & Kazemier, 1990; Saris, 1990).
 - Ook wordt afgeraden om uitgebreide lijsten van antwoordcategorieën te gebruiken (Kerssemakers, 1985; De Leeuw, 1989a; Van Eck & Kazemier, 1990; Saris, 1990).
 - De duur van het interview mag niet te lang zijn, om de uitval door vermoeidheid te beperken. Daarmee is ook de lengte van de vragenlijst beperkt. In de literatuur wordt een gemiddelde interviewduur van 15-20 minuten aangegeven, met een uitloop naar 30 minuten (Kerssemakers, 1985; De Leeuw, 1989a). De Heer, Akkerboom en

Israëls (1990) noemen een gemiddelde interviewduur bij het CBS van slechts 8 minuten voor telefonische interviews, tegen ten minste 45 minuten voor face-to-face interviews. Körmeni & Noordhoek (1989) concluderen echter dat in Denemarken een gemiddelde interviewtijd van 37 minuten voor de telefoon nog geen problemen opleverde.

3.2. Kosten

Een systematische studie over de kosten van CATI in vergelijking met CAPI is er niet. Wel wordt in de literatuur CATI vergeleken met PAPI-methoden. Zo noemt Kerssemakers (1985) besparingen voor CATI ten opzichte van face-to-face van 40-50% als de kosten voor het hele traject van het steekproefontwerp en de opzet van de vragenlijst tot en met het schone bestand in rekening worden gebracht. Bij korte interviews zijn de besparingen iets groter (50-60%) en bij langere interviews kleiner (25%). Een conclusie die in dezelfde richting wijst is te vinden in Van Eck & Kazemier (1990), waar het een vergelijking van telefonisch en face-to-face interviewing betreft. Nicholls & Groves (1986) komen tot de conclusie dat CATI niet goedkoper is dan de telefonische PAPI-methode, maar ook niet duurder. Zij brengen daarbij de hoge startkosten in rekening, zoals aanschaf van apparatuur en software, training, en extra ondersteuningspersoneel. Kostenbesparingen zijn te vinden op het terrein van interviewing en verwerking van de gegevens. Groves (1989) vindt op grond van een empirische vergelijking uit 1979 van telefonisch (centraal) en face-to-face interviewing dat het telefonisch onderzoek ongeveer de helft heeft gekost van het face-to-face onderzoek (zie tabel 2). Hierbij dient aangetekend te worden dat de gemiddelde lengte van het telefonisch interview ongeveer de helft was van het face-to-face interview. Groves (1989) noemt nog een aantal andere onderzoeken, weliswaar met een andere kostenverdeling, maar met een conclusie die in dezelfde richting wijst.

Kerssemakers (1985) en Groves (1989) geven een schema dat hier als basis kan dienen voor een overzicht van kostenaspecten (zie tabel 2). In tabel 2 is verder vermeld of voor deze aspecten een besparing wordt verwacht, of dat ze extra kosten met zich meebrengen als gecentraliseerd CATI wordt vergeleken met CAPI: met een '-' wordt aangegeven dat er sprake is van lagere kosten, een '+' betekent hogere kosten, een '0' wil zeggen dat de kosten voor CATI ten opzichte van CAPI ongeveer gelijk zijn. Ter illustratie zijn de bedragen zoals door Groves (1989) in 1979 berekend eveneens vermeld.

Aspecten die van invloed zijn op de kosten van het enquêteren zijn te verdelen in drie clusters: de startkosten, kosten voor de ontwikkeling, opzet en uitvoering van het veldwerk, en kosten voor de verwerking van de gegevens.

De totale kosten voor de aanschaf van benodigde apparatuur en hardware worden voor CATI lager verondersteld dan voor CAPI. De totale kosten voor de opzet en de uitvoering van het veldwerk worden bij CATI eveneens lager ingeschat dan bij CAPI. De redenen hiervoor zijn genoemd in paragraaf 3.1, punt 2. De kosten voor de opleiding en begeleiding van enquêtrices wordt bij CATI lager ingeschat, omdat er minder enquêtrices zijn, de opleiding minder tijd kost, en de begeleiding geen extra communicatie- en reiskosten met zich meebrengt (De Mast, 1990b). De kosten voor aanschaf van software en

ondersteuning door automatiseringsdeskundigen voor hardware en software, en de kosten voor de ontwikkeling van een interviewprogramma zullen bij beide CAI-methoden ongeveer gelijk zijn. En verder hebben we reeds in paragraaf 2 gezien dat de controle- en correctie-fase bij alle CAI-methoden kan worden overgeslagen als de routing- en controle-structuur correct in het interviewprogramma is geïmplementeerd, zodat dit geen kosten meer met zich meebrengt.

Tabel 2. Kostenvergelijking van CATI (gecentraliseerd) ten opzichte van CAPI

kostenaspect	kosten van CATI t.o.v. CAPI	kostenvergelijking van Groves (1989)			
		telefoon		face-to-face	
	1)	£	\$	£	\$
aanschaf van benodigde apparatuur en hardware	-				
aanschaf software	0				
ondersteuning (administratief, hardware, software)	0	6,8	2597,29	5,9	5042,64
ontwikkeling en testen van interviewmachine	0	5,5	2098,41	5,6	4773,25
aantal enquêtrices	-	27,0	10241,05	32,2	27321,04
opleiding van de enquêtrices	-	5,4	2066,34	11,2	9523,61
supervisie over de enquêtrices	-	6,7	2550,36	12,8	10858,52
trekken en uitzetten van steekproef	-	2,5	955,27	10,1	8547,15
verzenden van enquête materiaal en gegevens	-	0,0	0,00	5,0	4223,86
reiskosten	-	0,0	0,00	13,2	11194,76
telefoonkosten	+	41,6	15793,60	2,1	1756,45
controle en correctie van de gegevens	0	4,3	1637,47	1,9	1622,64
totaal		100,0	37939,79	100,0	84863,92

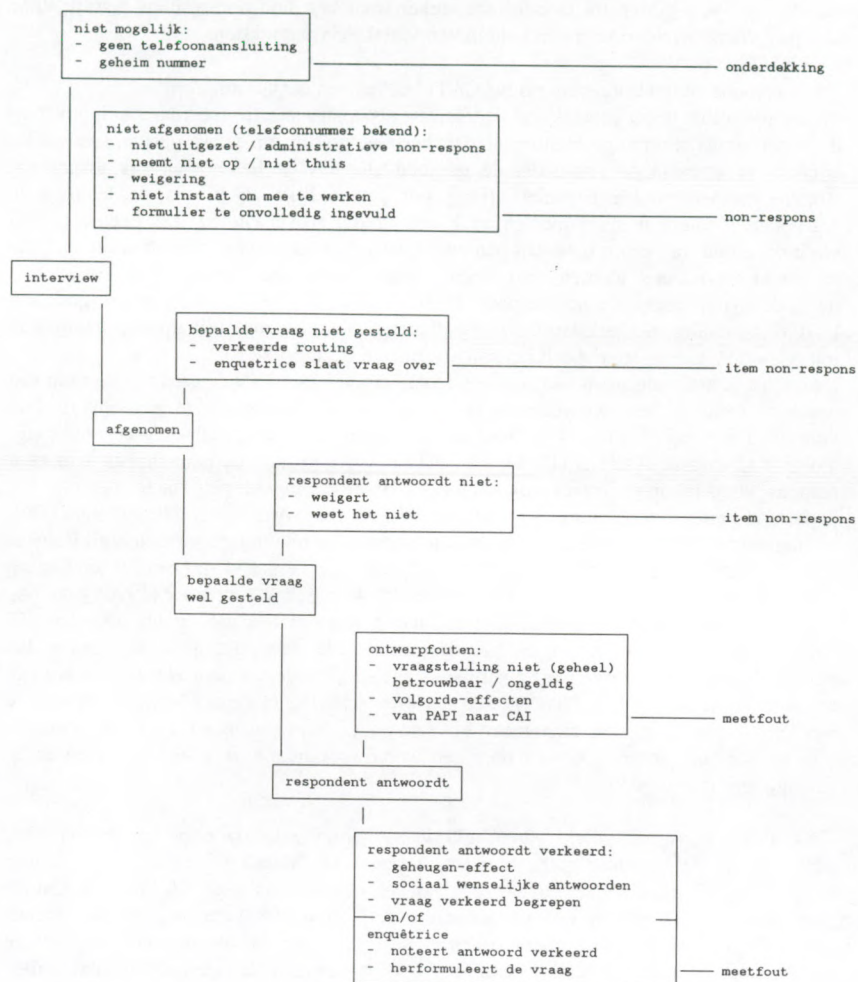
1) -: lagere kosten, 0: budgetneutraal, +: hogere kosten

Concluderend kunnen we stellen dat CATI goedkoper zal zijn dan CAPI. De kostenbesparingen betreffen vooral de lagere kosten voor de aanschaf van hardware, efficiënter gebruik van de hardware, het kleinere aantal enquêtrices en de afwezigheid van reiskosten. De besparingen door CATI ten opzichte van CAPI zullen groter zijn naarmate de duur van een interview korter en de steekproefomvang groter is.

3.3. Kwaliteitsaspecten

Tot de kwaliteitsaspecten van gegevens verkregen met behulp van enquêtes worden onderdekking, non-respons en meetfouten gerekend. De onderdekking is in paragraaf 3.1 (punt 3) al genoemd. Non-respons heeft betrekking op het niet verkrijgen van gegevens in die gevallen waarin een interview had moeten worden afgenomen. Als op slechts enkele vragen geen antwoord wordt gegeven, spreken we van item non-respons. Meetfouten ontstaan onder andere door het verkeerd weergeven van antwoorden door de enquêtrice of het geven van verkeerde antwoorden door de respondent. De relatie tussen onderdekking, non-respons en meetfout is weergegeven in figuur 1.

Figuur 1. De relatie tussen onderdekking, non-respons en meetfout (CATI).



We zullen nu non-respons en meetfout bespreken met het oog op de specifieke aspecten van CATI. Als vergelijkingsmateriaal nemen we gegevens van face-to-face interviewing, aangezien over CAPI geen gegevens bekend zijn. De non-respons kan bij CAPI echter wel verschillen in vergelijking tot face-to-face, zeker waar het item non-respons betreft, door duidelijker vraagformuleringen met behulp van variabele vraagteksten.

Non-respons en item non-respons bij CATI hebben een aantal oorzaken:

- In de eerst plaats is het gebruik van de telefoon als communicatie-medium van invloed op de non-respons. Sommige mensen vinden het niet prettig om lange gesprekken via de telefoon te voeren. Ze gebruiken de telefoon alleen voor het maken van afspraken. Andere mensen vinden het niet prettig om persoonlijke zaken via de telefoon te bespreken. Oudere mensen moeten vaak een drempel overwinnen. De Leeuw (1990) vindt op grond van een uitgevoerd pen en papier-experiment waarin telefonisch en face-to-face interviewing werden vergeleken, geen verschillen tussen beide interviewmethoden voor selectieve non-respons op belangrijke achtergrond variabelen (geslacht, leeftijd, opleiding, burgerlijke staat). Bij allochtonen zal het moeilijk zijn een telefonisch interview te houden. Hier speelt de taalbarrière een belangrijke rol.
- Daarnaast is het onderwerp van het onderzoek van belang. Zo weigeren mensen om een bepaalde vraag te beantwoorden als de vraag voor de respondent te gevoelig is. Een voorbeeld van zo'n vraag is de (per natie verschillend beoordeelde) inkomensvraag. Volgens Groves & Nicholls (1986), en De Leeuw (1990) zijn de percentages item non-respons voor inkomensvragen bij telefonisch interviewing vergelijkbaar met die van paper and pencil interviewing. Kerssemakers (1985) en Körmendi & Noordhoek (1989) vinden echter hogere percentages item non-respons bij telefonisch enquêteren. Volgens Kerssemakers (1985) bedraagt voor dit soort vragen de item non-respons bij telefonisch interviewing al snel het dubbele van die bij PAPI. Van Eck & Kazemier (1990) gaan nog een stapje verder wat gevoeligheid betreft, met hun onderzoek naar zwart inkomen. Zij concluderen dat in het algemeen het telefonisch interview niet geschikt is voor het verzamelen van informatie over moeilijke of gevoelige onderwerpen. Het onderwerp kan als dermate bedreigend worden ervaren dat men zelfs in het geheel weigert om mee te werken. Echter, als men zich meer bij een onderwerp betrokken voelt en positieve effecten van het onderzoek voor de eigen woon/werksituatie verwacht, zal men eerder geneigd zijn mee te werken.

Wat de responsgeneidheid van CATI in vergelijking tot face-to-face interviewing betreft, spreekt de literatuur zichzelf tegen. Volgens De Mast (1990a) wordt veelvuldig gemeld dat telefonisch enquêteren tot een hogere non-respons leidt. Zo melden Catlin, Ingram & Hunter (1988) en Wilson, Backshaw & Norris (1988) een hogere non-respons voor CATI, waarbij in beide onderzoeken is uitgegaan van de non-respons exclusief de onderdekking. Uit veel recente literatuur blijkt echter dat er geen noemenswaardige verschillen in non-respons percentages tussen telefonisch en face-to-face interviewing bestaan. Dit is ook de ervaring bij het CBS. Volgens Kerssemakers (1985) liggen de totale non-respons percentages voor de bij het CBS uitgevoerde CATI en face-to-face onderzoeken op eenzelfde niveau. Bovendien bleken er voor de verschillende categorieën van non-respons (weigering, neemt niet op/niet thuis, etc., zie figuur 1) weinig verschillen te bestaan in percentages tussen beide interviewmethoden. Van Eck & Kazemier (1990)

komen in hun onderzoek naar zwart inkomen tot een soortgelijke conclusie. Körmendi & Noordhoek (1989) zijn eenzelfde mening toegedaan.

Momenteel zijn de respons-percentages voor CATI bij het CBS zelfs gunstiger dan voor CAPI. Tot eenzelfde conclusie komt De Leeuw (1990) in haar experiment: 51% respons voor face-to-face tegen 66% voor het telefonisch onderzoek. Bovendien vertonen de percentages non-respons en weigering bij CATI bij het CBS geen stijgende tendens, terwijl dit bij CAPI/face-to-face wel het geval is (De Heer, Akkerboom & Israëls, 1990). Verder is de administratieve non-respons bij CATI aanzienlijk lager dan de 4-5% bij CAPI-onderzoeken (Kerssemakers, 1990). Daarbij zijn bij telefonisch enquêteren meer pogingen tot het maken van contact mogelijk dan bij face-to-face interviewing, zodat minder vaak 'niet thuis' verwacht kan worden.

Meetfouten kunnen ontstaan door de telefoon als communicatiemedium (zie paragraaf 3.1, punt 1). Ook het onderwerp van het onderzoek kan tot meetfouten leiden. En verder kunnen meetfouten ontstaan door geen rekening te houden met de eisen ten aanzien van de ontwikkeling van de vragenlijst (paragraaf 3.1, punt 4). Verdere effecten voor CATI zijn onder andere het ontvouwingeffect (als een schaal met uitgebreide antwoordcategorieën wordt gesplitst in een aantal vervolgvragen met deelschalen: men is dan eerder geneigd om de middencategorie te kiezen), of het extremitateffect (als toch voor een uitgebreide antwoordschaal wordt gekozen: men is dan geneigd om de extremen te kiezen) (Kerssemakers, 1985; De Leeuw, 1989a; Saris, 1990).

Concluderend kunnen we stellen dat de kwaliteit van de gegevens van telefonisch onderzoek voor een belangrijk deel afhankelijk is van dezelfde factoren die de kwaliteit bij CAPI bepalen, zoals de gevoeligheid van het onderwerp, de betrokkenheid van de doelpopulatie bij het onderwerp, de vaardigheid en de ervaring van de enquêtrices, de periode van het jaar en het tijdstip van de dag waarop het interview moet plaatsvinden (zie ook figuur 1). Op grond van het uitgevoerde experiment concludeert De Leeuw (1990) dat er geen verschillen zijn in de kwaliteit van de gegevens voor telefonisch en face-to-face interviewing. De keuze voor de telefoon heeft echter zeker effect op de selectiviteit van de gegevens als de doelpopulatie moeilijk telefonisch bereikbaar is (onderdekking in figuur 1).

Selectiviteit kan worden verminderd door telefonisch moeilijk bereikbare groepen met een andere interviewmethode, bijvoorbeeld CAPI, te benaderen. Dit brengt wel extra kosten met zich mee. Andere nadelige effecten, die ontstaan door het gebruik van de telefoon als communicatie-medium, kunnen volgens Groves & Nicholls (1986) worden tegengegaan door een goed ontworpen interviewprogramma waarin rekening wordt gehouden met de extra eisen die de telefoon stelt aan de vragenlijst (zie paragraaf 3.1, punt 4), een goede training van de enquêtrices die voorbereid moeten zijn op alle mogelijke situaties tijdens een telefoongesprek, en een goede organisatie voor het maken van afspraken (belschema's) en supervisie. Voor een CATI-programma betekent een goed ontwerp onder andere dat in de vragenlijst gebruik moet worden gemaakt van de faciliteit om vragen te personifiëren met behulp van variabele vraagteksten. Wat het gebruik van consistentie checks betreft kan overwogen worden om niet te veel controles op te nemen, aangezien deze de voortgang van het gesprek kunnen bemoeilijken. Anderzijds moet er naar gestreefd

worden intern consistente records te krijgen. Zo kan er voor gekozen worden om alleen de meest noodzakelijke consistentie checks te definiëren. En natuurlijk is een correcte routing vanzelfsprekend.

4. Conclusie

We hebben gezien dat de keuze van CATI invloed heeft op de kosten van het onderzoek en op de kwaliteit van de gegevens. In het algemeen kan gesteld worden dat CATI (met een centrale enquëtering) over de hele linie van aanschaf van hardware en software tot verwerking van de gegevens goedkoper is dan CAPI. Wat de kwaliteit van de gegevens betreft kan geen algemene conclusie worden geformuleerd. De kwaliteit van de gegevens is afhankelijk van het onderwerp van het onderzoek, het ontwerp van de vragenlijst en de bereikbaarheid van de doelpopulatie.

Wat het onderwerp betreft is CATI minder geschikt als:

- het onderwerp gevoelig of bedreigend is;
- een sterk beroep wordt gedaan op het geheugen;
- gegevens moeten worden opgezocht.

Met betrekking tot het ontwerp van de vragenlijst geldt voor CATI dat:

- geen gebruik kan worden gemaakt van uitgebreide visuele hulpmiddelen (antwoordkaarten, toonmateriaal);
- geen vragen gesteld mogen worden waarbij veel (meer dan 5) antwoordcategorieën tegelijk worden onderscheiden;
- geen vragen gesteld mogen worden met een ingewikkelde formulering;
- geen vragen gesteld mogen waarbij veel uitleg, zoals formele definiëring van begrippen of extra uitleg over het onderzoek, noodzakelijk is;
- de gemiddelde duur van het interview niet langer dan 20-30 minuten mag zijn;
- niet veel open vragen gesteld mogen worden, waarop uitgebreide antwoorden worden verwacht;
- een goed gebruik wordt gemaakt van variabele vraagteksten;
- een goede controle-structuur wordt opgesteld, waarbij vooral niet te veel consistentie checks worden gedefinieerd.

De doorgaans lagere kosten en de soms lagere kwaliteit van de gegevens (met betrekking tot selectiviteit) voor CATI kunnen we tot slot met elkaar in verband brengen door de kostenbesparingen van CATI in te ruilen voor extra maatregelen ter verbetering van de kwaliteit van de gegevens. Hierbij valt onder andere te denken aan een grotere steekproefomvang. Zo kunnen bij onderzoeken naar gebeurtenissen die bij bepaalde groepen met een vrij grote regelmaat terugkomen, meerdere personen ondervraagd worden met een kortere retrospectieperiode in plaats van minder personen met een langere retrospectieperiode. Dit heeft met name zin bij onderzoeken waarbij voor de belangrijke doelvariabelen de variantie tussen personen/huishoudens veel groter is dan de variantie in de tijd (zoals bij dagrecreatie), en men niet specifiek in het longitudinale aspect van de informatie is geïnteresseerd. Het waarnemen van een tweede periode bij dezelfde persoon levert dan veel minder informatie op dan wanneer men hiervoor een andere persoon enquëteert. Bij onderwerpen waarbij de variantie tussen personen/huishoudens relatief kleiner is, is zo'n voordeel kleiner. Dit is bijvoorbeeld het geval bij aankoop van duurzame artikelen of slachtofferschap.

Literatuur

- Aquilino, W.S. en L.A. Lo Sciuto, 1990, Effects of interview mode on self-reported drug use. *Public Opinion Quarterly*, Vol. 54, no. 3, blz. 362-395.
- Berg, J. van den, 1987, Telefonische bereikbaarheid respondenten face-to-face enquêtes. Interne CBS-nota (concept gebaseerd op de Gezondheidsenquête 1987, Centraal Bureau voor de Statistiek, Heerlen).
- Bronner, F., 1980, Telefonisch onderzoek. *Kwantitatieve Methoden*, Vol. 5, no. 2, blz. 145-156.
- Catlin, G., S. Ingram en L. Hunter, 1988, The effect of CATI on data quality: A comparison of CATI and paper methods. In: *Proceedings of the Fourth Annual Research Conference* (U.S. Bureau of the Census, Washington), Arlington (Virginia), blz. 291-299.
- Eck, R. van en B. Kazemier, 1990, Zwarte inkomsten uit arbeid. Een nadere analyse van de uitkomsten van enquête-onderzoek. *Statistische onderzoekingen M36* (Centraal Bureau voor de Statistiek, Voorburg/Heerlen).
- Groves, R.M. en W.L. Nicholls, 1986, The status of Computer-Assisted Telephone Interviewing: Part II - Data quality issues. *Journal of Official Statistics*, Vol. 2, No. 2, blz. 117-134.
- Groves, R.M., 1989, *Survey errors and survey costs* (Wiley, New York).
- Heer, W.F. de, J.C. Akkerboom en A.Z. Israëls, 1990, Ideas for nonresponse investigations; Contributions to the Nonresponse Workshop. Voorlopige versie (Centraal Bureau voor de Statistiek, Heerlen/Voorburg).
- House, C.C., 1985, Questionnaire design with Computer Assisted Telephone Interviewing. *Journal of Official Statistics*, Vol. 1, No. 2, blz. 209-219.
- Kerssemakers, F.A.M., 1985, Telefonisch enquêteren. In: *CBS Select 3* (Centraal Bureau voor de Statistiek, Voorburg/Heerlen), blz. 211-230.
- Kerssemakers, F.A.M., 1990, Interne memo (Centraal Bureau voor de Statistiek, Heerlen).
- Kerssemakers, F.A.M., F.A.C. de Mast en P.W.M. Remmerswaal, 1987, Computer Assisted Telephone Interviewing, some response findings. In: *CBS Select 4* (Centraal Bureau voor de Statistiek, Voorburg/Heerlen), blz. 119-131.

- Körmendi, E. en J. Noordhoek, 1989, Data quality and telephone interviews. A comparative study of face-to-face and telephone data collecting methods. (Danmarks Statistik, Kopenhagen).
- Leeuw, E.D. de, 1989a, Een schriftelijke vragenlijst, een telefonisch interview, of een face-to-face interview? Een moeilijke keuze.
In: Van der Zouwen, J. en W. Dijkstra, 1989, Sociaal-wetenschappelijk onderzoek met vragenlijsten. Methoden, knelpunten, oplossingen (VU Uitgeverij, Amsterdam), blz. 17-34.
- Leeuw, E.D. de, 1989b, Computergeleid enquêteren; een overzicht van technologische vernieuwingen. Tijdschrift voor Onderwijsresearch, Vol. 14, no. 4, blz. 201-213.
- Leeuw, E.D. de, 1990, Data quality in mail, telephone and face-to-face surveys; a mode comparison in the Netherlands. Response effects in surveys, Technical report nr. 4, Vrije Universiteit, Amsterdam.
- Mast, F.A.C. de, 1990a, Non-respons in persoons- en gezins-enquêtes. Interne CBS-nota (Hoofdafdeling Persoons- en Gezins-enquêtes, Centraal Bureau voor de Statistiek, Heerlen).
- Mast, F.A.C. de, 1990b, Interne memo (Centraal Bureau voor de Statistiek, Heerlen).
- Nicholls, W.L. en R.M. Groves, 1986, The status of Computer-Assisted Telephone Interviewing: Part I - Introduction and impact on cost and timeliness of survey data. Journal of Official Statistics, Vol. 2, No. 2, blz. 93-115.
- Pannekoek, J., 1988, Interviewer variance in a telephone survey. Journal of Official Statistics, Vol. 4, No. 4, blz. 375-384.
- Saris, W.E., 1989, A technological revolution in data collection. Quality & Quantity, Vol. 23, No. 3-4, blz. 333-349.
- Saris, W.E., 1990, Microcomputers in data collection (eerste versie). Universiteit van Amsterdam (Vakgroep Statistiek en Methodologie), Amsterdam.
- Sikkel, D., 1986, Computer Assisted Data Collection. CBS-rapport (Hoofdafdeling Statistische Methoden, Centraal Bureau voor de Statistiek, Voorburg).
- Sikkel, D., 1988, Quality aspects of statistical data collection. Proefschrift (Sociometric Research Foundation, Amsterdam).
- Steel, D. en P. Boal, 1988, Accessibility by telephone in Australia: Implications for telephone surveys. Journal of Official Statistics, Vol. 4, No. 4, blz. 285-297.

Sykes, W. en M. Collins, 1988, Effects of mode of interview: Experiments in the UK.

In: Groves, M., P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls en

J. Waksberg, 1988, Telephone survey methodology (Wiley, New York).

Wilson, P., N. Blackshaw en P. Norris, 1988, An evaluation of telephone interviewing on the British Labour Force Survey. Journal of Official Statistics, Vol. 4, No. 4, blz. 385-400.

ontvangen	20- 7-1991
geaccepteerd	20- 12-1991