

# is televisiekijken meetbaar?

jan van den bulck

Kijkcijfers spreken tot de verbeelding. Het accurate meten van televisiekijkgedrag is een van die zeldzame wetenschappelijke bezigheden die een belangrijke plaats in het maatschappelijke leven lijken in te nemen, naast, onder meer, meteorologische vaststellingen of de meting van het ozongehalte. Opmerkelijk is daarbij dat televisiekijken een zeer belangrijk commercieel aspect heeft. In de televisiewereld hangt van de meting van kijkcijfers de leefbaarheid van zenders en programma's af. Kijkcijfers zijn zelfs de basis van politieke en beleidsbeslissingen, zo blijkt uit de formuleringen van het recente McKinsey-rapport over de Vlaamse openbare omroep en de erop gebaseerde uitlatingen van de mediaminister.

Het meten van kijkcijfers is echter niet van vanzelfsprekend. In de wetenschappelijke wereld bestaat een ruim scala aan technieken om kijkgewoontes vast te stellen (cf. Van den Bulck, 1996), en ook de voor commerciële doeleinden verzamelde *kijkcijfers* staan regelmatig ter discussie (cf. De Morgen, 17 februari 1995; 25 maart 1995). Zowel de wetenschappelijke als de commerciële meting van het kijkgedrag hebben hun zwakke en sterke kanten. De communicatiewetenschap maakt vaak gebruik van enquêtes, waarin televisiekijken wordt gemeten aan de hand van zogeheten 'self-reports', waarbij mensen *zelf* en *post factum* hun kijkgedrag moeten inschatten. Televisiestations gebruiken daarentegen door privé-bedrijven verzamelde zogeheten 'set-meter data', waarbij van een panel van gezinnen dagelijks zeer nauwkeurig (per seconde!) wordt vastgesteld naar welke zender wordt gekeken.

Twee vragen laten zich stellen. Ten eerste, zouden wetenschappers niet beter gebruik maken van veel nauwkeurigere 'set-meter' dataverzamelingmethoden? Ten tweede zou men de vraag echter ook kunnen omdraaien: waarom duur set-meter onderzoek doen als eenvoudige enquêtes even goed zouden zijn? Op het eerste gezicht kan de tweede optie absurd overkomen, terwijl de eerste een na te streven doel kan lijken. Audimetrie (het mechanische meten van kijkcijfers met behulp van apparatuur) is immers veel nauwkeuriger dan het zelf inschatten van soms vaag gedefinieerde gewoontes. Het is echter belangrijk op te merken dat de nauwkeurigheid (en de daarbij horende betrouwbaarheid) van de meting slechts één aspect is. Minstens even belangrijk is immers de *validiteit* van de gehanteerde variabelen (cf. Mulac en Kunkel,

1989). Het verzamelen van data is immers altijd mogelijk, maar heeft slechts zin als die data *betekenisvol* zijn. Ze moeten, met andere woorden, verwijzen naar een concept dat ook in de realiteit bestaat. Methodologen weten dat wanneer in een vragenlijst die 'diepgewortelde' opinies meet, ter controle wordt gepeild naar opinies over niet bestaande issues, een groot deel van de respondenten toch een mening zal uiten (cf.: Shuman en Presser, 1981). Men heeft in zo'n geval wel iets gemeten, maar er beantwoordt in de realiteit niets aan deze data.

'Set-meter' en 'self reported' data zijn niet gelijkwaardig. Ze hebben niet dezelfde nauwkeurigheid en validiteit. Mechanisch gemeten kijkcijfers hebben een grotere nauwkeurigheid, maar zelf-gerapporteerde gegevens zouden voor heel wat vormen van wetenschappelijk onderzoek meer valide kunnen zijn. Belangrijker nog is de vaststelling dat ze niet hetzelfde meten (cf. Van den Bulck, 1995, 1996). Kijkcijfers meten vrij nauwkeurig het actuele kijkgedrag van het onderzochte panel, terwijl enquêtes peilen naar langdurige kijkgevoontes van de steekproef. Beide methodes zijn niet zonder meer verwisselbaar.

## PROBLEMEN VAN ENQUETE-GEGEVENS

Webster en Wakshlag (1985:41) menen dat respondenten er geen probleem mee hebben toe te geven dat ze zware kijkers zijn. Sociale wenselijkheids-effecten zouden in dat opzicht dus geen effect hebben op de betrouwbaarheid van enquêtegegevens.

Chaffee en Mutz (1988:38) menen dat mensen over het algemeen hun kijkgedrag juist kunnen inschatten. Televisieprogramma's hebben immers meestal een standaardduur (30 minuten). Dit zou inhouden dat de mentale processen waarop de respondent een beroep moet doen om in antwoord op enquêtevragen zijn kijkgedrag in te schatten, vrij eenvoudig zijn. Andere auteurs hebben echter beweerd dat respondenten hun kijkgedrag tot 25 procent over- of onderschatten (Mulac en Kunkel, 1989:63). Van den Bulck (1996) vroeg respondenten te schatten hoeveel uren ze televisiekeken, maar vroeg ook wanneer ze elke dag begonnen en stopten met televisiekijken. Als een respondent beweert drie uur per dag televisie te kijken, maar vervolgens indiceert om acht uur met kijken te beginnen en om tien uur te stoppen, is het duidelijk dat minstens een van beide schattingen fout moet zijn. In de aangehaalde studie, waarin respondenten nog een ruime marge kregen om zich te vergissen, bleek 15 à 20 procent het eigen kijkgedrag te overschatten.

Als respondenten hun kijkgedrag verkeerd inschatten, kan dat het gevolg zijn van het feit dat ze een foute heuristisch toepassen. Het is echter ook mogelijk dat kijkgedrag fout wordt gemeten omdat de vraagstelling verkeerd is. In veel onderzoek wordt er impliciet van uitgegaan dat 'veel' televisiekijken hetzelfde is als 'vaak' televisiekijken. Als gevolg daarvan vragen onderzoekers meestal alleen 'hoeveel uren kijkt u doorgaans televisie', en gaan ze er bij het schatten van het totale kijkgedrag vervolgens van uit dat de respondent *elke dag* zoveel kijkt. Het kijkvolume wordt dan bijvoorbeeld met zeven vermenigvuldigd om een schatting van het wekelijkse kijkgedrag te bekomen. In Van den Bulck (1996) wordt aangetoond dat slechts ongeveer vijftig procent van de 'veelkijkers' ook 'vaakkijkers' zijn. Met deze methode wordt het kijkvolume van bepaalde respondenten die veel kijken, maar slechts enkele dagen per week, dus systematisch overschat.

Zulke meetproblemen roepen vragen op over de geldigheid (validiteit) van de in veel onderzoek gebruikte concepten. Verschillende termen (kijkcijfers, kijkgedrag, kijkgewoonte, kijkvolume, kijkattitude, enz.) worden vaak door elkaar gebruikt, en in andere gevallen duidt hetzelfde woord op verschillende begrippen. Elk woord verwijst echter naar een bepaalde theoretische '*construct*'. Als het woord 'kijkvolume' wordt gebruikt als men het over set-meter data heeft, is de *construct* dat het over feitelijk, dagelijks kijkgedrag gaat. Als het woord echter op basis van zelf-gerapporteerde data wordt gebruikt, verwijst het naar veel algemenere kijkgewoontes. Bovendien ligt aan de basis van zo'n *construct* vaak een niet geëxpliciteerde (soms zelfs nog niet ontdekte) theorie (cf. o.m. Salomon en Cohen, 1978; Allen, 1981; Mulac en Kunkel, 1989). Als 'volume' alleen wordt gemeten aan de hand van het aantal kijkuren op bepaalde dagen, is de impliciete theorie bijvoorbeeld dat 'veelkijken' ook 'vaakkijken' impliceert.

Het is niet altijd eenvoudig om vast te stellen wat men precies heeft gemeten. De vraag 'hoeveel uren kijkt u gewoonlijk op een doorsnee dag?', zal niet voor elke respondent hetzelfde betekenen. Hoe kan de onderzoeker weten wat hij of zij heeft gemeten? Wat betekent 'gewoonlijk'? Hoe interpreteert de respondent de term 'doorsnee dag'? Meet zo'n vraag de kijkgewoonte en zo ja: wat is het daaraan verbonden tijdsperspectief? Rosengren (1994:21) merkt op dat verschillende mediatheorieën een verschillend tijdsperspectief hantieren. Nieuwsdiffusietheorieën onderzoeken wat er zich onmiddellijk na het bekend worden van een bepaald nieuwsfeit afspeelt. Agenda setting theorieën onderzoeken meestal ook effecten op de relatief korte termijn. Cultivatietoetsen gaan dan weer over het gedurende jaren of zelfs decennia gecumuleerde (meestal kleine) effect van een soort 'meta-boodschap' die door de expliciete inhoud van de media doorschemert. Over welk niveau kan men

nog met enige geldigheid uitspraken doen aan de hand van op transversaal onderzoek gebaseerde en met vrij algemeen geformuleerde vragen gemeten gegevens?

## PROBLEMEN VAN KIJKCIJFERMETING

Modern kijkersonderzoek dat wordt uitgevoerd in opdracht van televisiestations, gebeurt aan de hand van panelonderzoek. In Vlaanderen wordt bij zevenhonderdvijftig (tot voor kort: 600) gezinnen een apparaatje verbonden met alle aanwezige televisietoestellen. Dit apparaatje registreert per seconde op welk kanaal het televisietoestel afgestemd staat. Vermits van de respondenten wordt gevraagd op elk moment aan te geven welke gezinsleden en bezoekers aan het kijken zijn, wordt zo het dagelijkse kijkgedrag van ongeveer 2.250 mensen geregistreerd.

Vooraf de op basis van deze gegevens volgens verschillende methoden uitgevoerde puntschattingen spreken tot de verbeelding. De cijfers van het panel worden geëxtrapoleerd om een schatting te geven van de verdeling van de kijkende bevolking over programma's en zenders. De pers publiceert daarbij graag classificaties zoals een 'top 20'. Het weekblad Humo brengt zelfs zo'n top 10 voor elke Vlaamse zender. Dit is uiteraard niet de enige toepassingsvorm voor deze cijfers. Televisiekanalen en adverteerders zijn vooral geïnteresseerd in de samenstelling van het publiek volgens bepaalde demografische kenmerken. Ook voor wetenschappelijk onderzoek lijken kijkcijfers een schat aan informatie in zich te bergen. Zeer interessant lijkt bijvoorbeeld de *schakelanalyse*, waarbij per seconde wordt nagegaan wat een bepaalde groep kijkers heeft gedaan. Op die manier kan bijvoorbeeld worden onderzocht wanneer kijkers 'zappen'. Als dit onderzoek bovendien wordt gerelateerd aan de inhoud van het onderzochte programma, kan daaruit zelfs worden afgeleid waarom kijkers op een bepaald moment beslissen van kanaal te veranderen (bijvoorbeeld tijdens het muzikale intermezzo van een praatprogramma, cf. Plees, 1994).

Kijkcijfers zijn echter pas bruikbaar voor het trekken van conclusies op het niveau van de kijkende bevolking als ze aan bepaalde methodologische criteria voldoen. Meer bepaald dient te worden nagegaan hoe betrouwbaar, nauwkeurig en veralgemeenbaar ze zijn.

### Betrouwbaarheid

Een eerste belangrijke aspect van de betrouwbaarheid van de gegevens is de representativiteit van de steekproef. Van een toevalssteekproef die groot

genoeg is, kan worden gezegd dat ze binnen bepaalde foutenmarges een goede schatting kan opleveren van de verdeling van bepaalde gegevens in de bevolking. Op basis van de steekproef kan men met andere woorden uitspraken doen over heel de populatie waaruit de steekproef is getrokken. Dit kan echter pas als het om een echte toevalssteekproef gaat, omdat dan maximaal wordt vermeden dat bepaalde groepen systematisch worden genegeerd. Een telefoonenquête kan bijvoorbeeld niet leiden tot een representatieve steekproef, omdat alle leden van de populatie die niet over een telefoonaansluiting beschikken, per definitie niet in het onderzoek kunnen worden opgenomen.

Voor kijkersonderzoek wordt echter gebruik gemaakt van een gestratificeerde steekproef. Dit betekent dat er van bepaalde kenmerken wordt vooropgesteld hoe ze in de steekproef verdeeld dienen te zijn. Concreet houdt dit bijvoorbeeld in dat op voorhand wordt vastgesteld hoeveel mannelijke arbeiders tussen de 40 en 50 jaar de steekproef zal moeten bevatten. Welke mensen uiteindelijk zullen worden geselecteerd, wordt wel nog door het toeval bepaald. Volgens den Boon (1993:269) leidt dit ertoe dat het panel een goede afspiegeling van de totale bevolking vormt. Hoewel methodologen er van uitgaan dat het stratificeren in de praktijk nauwelijks een verbetering teweegbrengt (Pedhazur en Pedhazur Schmelkin, 1991:332), is het in theorie een goede manier om de representativiteit te verhogen (Hays, 1994).

Het uitvoeren van puntschattingen op basis van de verkregen gegevens is echter niet vanzelfsprekend. In de pers wordt graag gewag gemaakt van exacte schattingen van aantallen kijkers. In de praktijk moet echter een betrouwbaarheidsinterval worden berekend waartussen de verkregen schattingen kunnen schommelen. Vooral als subgroepen worden bestudeerd (zoals: alle mannelijke kijkers tussen de 25 en de 45 jaar met een inkomen tussen bepaalde grenzen), worden de schattingen steeds instabieler (cf. den Boon, 1993, voor een overzicht van de berekeningswijzen). Met de komst van kleine kanalen (cf. Kanaal 2 en VT4) is de neiging om met marginale cijfers te werken nog groter geworden. In De Morgen van 18 januari 1996 vermeldt Terry Verbiest, programmadirecteur van VT4, dat het marktaandeel van zijn kanaal recent van 8,7 naar 10,1 procent is gegaan. Hij voegt er aan toe dat ze in hun doelgroep (de 15 tot 44-jarigen) echter zelfs 15 procent halen. Zulke uitspraken zijn gebaseerd op de kijkcijfers van een zeer kleine groep respondenten uit de steekproef, waardoor de betrouwbaarheidsintervallen aanzienlijk worden (zie den Boon, 1993:283, voor voorbeelden). De ware cijfers voor de hele bevolking kunnen zowel veel hoger als veel lager liggen.

Problematischer is de vaststelling dat het kijkerspanel in feite een *clustersteekproef* vormt. Hoewel de gewoonte bestaat te extrapoleren op basis van

individuele kijkers, is de steekproef er eigenlijk één van huishoudens. Men dient dus met andere woorden de meeste betrouwbaarheidsintervallen hoogstens op het niveau van de 750 huishoudens te berekenen, en niet op dat van de meer dan 2000 potentiële kijkers. Het kijkgedrag van deze kijkers is immers niet onafhankelijk van het gedrag van andere kijkers uit hetzelfde cluster. Concreet betekent dit dat of een bepaalde jongere al dan niet naar een bepaald jongerenprogramma heeft gekeken, niet alleen afhangt van het kijkgedrag van die jongere, maar ook van dat van de andere leden van het gezin en van de gezags- en beslissingsstructuur in dat gezin (cf. Webster en Wakshlag, 1982, over de groepsprocessen bij programmakeuze).

## Nauwkeurigheid

Kijkcijfers zijn erg nauwkeurig. Per seconde wordt geregistreerd op welk kanaal het televisietoestel ingeschakeld staat. Tegelijkertijd moeten de leden van het gezin via een apparaat opgeven wanneer ze beginnen en stoppen met kijken. In persartikels over audiometrie worden soms leden van het panel opgevoerd die beweren dit na een tijdje na te laten (cf. De Morgen, 25 maart 1995). Het is precies om dit soort 'panelvermoeidheid' te vermijden dat het panel regelmatig wordt vernieuwd. Buitenlandse studies die 'toevallige controles' uitvoerden, suggereren nochtans dat de invloed van dit fenomeen minimaal is. Gemiddeld genomen zouden de leden van het panel dus relatief nauwkeurig zijn in het opgeven van hun gedrag (cf. Dahms, 1983; Lutz, 1987; Sang et al., 1994; Wiedemann, 1988).

Een belangrijke reden voor het tot op de seconde nauwkeurig registreren van kijkcijfers is het feit dat adverteerders willen weten wat het bereik is geweest van hun televisieadvertenties. Bij de oudere dagboekmethodes, waarbij het panel zelf moest invullen waarnaar het had gekeken, kon onmogelijk worden nagegaan of men nog voor zijn toestel zat op het ogenblik dat de commercials werden uitgezonden. Men zou bijvoorbeeld van kanaal kunnen veranderen op het ogenblik dat de reclame begint. Mechanisch geregistreerde kijkcijfers laten toe precies uit te maken op welk ogenblik een andere zender werd gekozen. Toch is dit geen bevredigende oplossing. Het is mogelijk dat kijkers de reclame gebruiken als een 'plaspauze', maar het de moeite niet vinden die korte onderbreking te 'melden' aan het apparaat. Misschien wordt het reclameblok gebruikt om met gezinsleden te overleggen of men niet beter naar een ander kanaal zou overschakelen. Om te weten of men een programma ook echt heeft *gezien*, zijn daarom methodes ontwikkeld om oogbewegingen te registreren (cf. Thorson, 1994; Grimes en Meadowcroft, 1995). De data worden daardoor steeds betrouwbaarder en nauwkeuriger (cf.: Mulac en Kunkel, 1989:63).

## Validiteit

De verfijning van de nauwkeurige meting van het televisiekijken hoeft niet te eindigen bij het registreren van oogbewegingen. Het is niet noodzakelijk zo dat het *zien* van een commercial ook betekent dat men hem heeft *opgemerkt*. Salomon en Cohen (1978:267) wijzen er immers op dat mediaboodschappen eerst moeten worden opgemerkt, voor ze *überhaupt* effect kunnen hebben. Het is vervolgens nog niet zo dat een boodschap die werd opgemerkt, ook werd begrepen. Dat hoeft op zijn beurt niet te betekenen dat er ook een effect is, en het effect hoeft, in het geval van advertenties, niet gelijk te zijn aan het door de opdrachtgever gewenste effect. Zoals blijkt uit een verzameling artikels uit een recente reader van Lang (1994), is er een trend naar steeds gesofistikeerder onderzoek dat pas een voorlopig eindpunt zal bereiken als de neurologische activiteit in het brein van de kijker wordt geregistreerd en geïnterpreteerd. Als dat soort onderzoek gemeengoed wordt, zal pas met enige zekerheid kunnen worden gezegd dat een bepaalde boodschap echt bij de ontvanger is 'aangekomen'.

Hoewel zulke vormen van onderzoek een zeer hoge nauwkeurigheid inhouden, stellen ze grote validiteitsproblemen. Het begrip validiteit of geldigheid verwijst immers naar de mate waarin wat geldt voor de respondenten uit het onderzoek, ook mag worden veralgemeend naar de populatie.

Over het algemeen wordt er altijd van uitgegaan dat het kijkgedrag van de leden van het kijkcijferpanel niet verschilt van dat van vergelijkbare personen uit de populatie. In het begin zouden nieuwe leden van het panel 'abnormaal' kijkgedrag vertonen (omdat ze zich bestudeerd weten, omdat ze vrienden en familie uitnodigen om naar dat 'kastje' te komen kijken, enz.), maar dit extreme gedrag verdwijnt na enige tijd, waaruit de uitvoerders van het kijkersonderzoek afleiden dat het kijkgedrag op dat ogenblik weer 'normaal' is (cf. De Morgen, 25 maart 1995). Duits onderzoek heeft nagegaan wat het effect is van een 'interventie' in een bestaand panel. Hierbij werden mensen van wie al enige tijd het kijkgedrag werd gemeten, bevraagd over hun kijkgedrag. Hoewel deze interventie een klein effect bleek te hebben, bleek uit de kijkcijfers dat dit effect snel weer verdween, en dat de cijfers vóór en na de interventie vergelijkbaar waren (cf. Sang et al., 1994). Het is per definitie echter onmogelijk om het gedrag van niet-bestudeerde mensen te vergelijken met dat van bestudeerde mensen. Vermits de deelnemers aan het kijkersonderzoek aan het apparaat moeten laten weten dat ze met televisiekijken beginnen, worden ze telkens bewust gemaakt van het feit dat ze aan het televisiekijken zijn, een handeling die weinig bewust gebeurt en slechts 'low involvement' inhoudt (cf. Shrum, 1995). Het introduceren van een 'bewustmaker' kan dan ook een niet te verwaarlozen effect hebben.

Als nieuwe kanalen aan het aanbod worden toegevoegd, wordt de televisiekijkapparatuur door technici bijgesteld. Ook dat is niet zonder effect. Het is best mogelijk dat bepaalde delen van de bevolking sommige nieuwe zenders niet eens programmeren (uit nalatigheid, bewuste keuze of gebrek aan technische kennis). Bovendien heeft onderzoek aangetoond dat televisiekijkers vaak een typisch repertorium van kanalen raadplegen bij het 'zappen' (cf. Heeter, 1985; Ples, 1994; Van den Bulck, 1995). De volgorde van de kanalen zoals de kijker die zelf instelt op zijn toestel en afstandsbediening, kan dus een niet te verwaarlozen invloed hebben op de keuze van programma's. Of een nieuw kanaal op nummer 4 dan wel op nummer 56 van de afstandsbediening werd geprogrammeerd, kan de vanzelfsprekendheid van op toevallige keuzes gebaseerd kijkgedrag sterk bepalen (cf. Van den Bulck, 1995:157).

Ook de representativiteit van de steekproef roept validiteitsvragen op. Zoals hoger werd gesteld, is een gestratificeerde steekproef geen hinderpaal om representativiteit te bekomen. Op voorwaarde dat de 'strata' overeenkomen met de werkelijke verdeling in de populatie, kan dit zelfs scheeftrekking helpen voorkomen. Binnen elk stratum moeten de respondenten echter willekeurig ('at random') worden gekozen. In geen geval mag de selectie van respondenten worden beïnvloed door een factor die de resultaten van het kijkcijferonderzoek kan vertekenen. Weimann et al. (1992) vermelden bijvoorbeeld dat bij dagboekstudies via een postenquête slechts 66% van de gecontacteerden data terugstuurden. De auteurs wijzen er op dat het al dan niet terugsturen ongetwijfeld mee beïnvloed zal zijn door factoren die samenhangen met het televisiekijken. Lichte kijkers vinden televisiekijken misschien onbelangrijk. Voor zware kijkers neemt televisie echter wellicht zo'n grote plaats in hun leven in, dat ze sneller geneigd kunnen zijn om op onderzoek te reageren. Als zulke redeneringen meespelen, zal onder meer het televisiekijken overschat worden. Volgens een recente Dimarso-studie, uitgevoerd in opdracht van Humo (2885), zouden zware kijkers significant meer bereid zijn dan lichte kijkers om aan kijkcijferonderzoek mee te doen. Als dit waar is, worden de kijkcijfers *overschat*. In een schriftelijke reactie liet Aspemar weten dat met deze mogelijke vertekening wordt rekening gehouden (Humo, 2886). Uit de brief blijkt dat 'de panelgezinnen werden (...) geselecteerd op basis van hun geproclameerde (gemiddelde) dagelijkse kijktijd'. Vermits de panelleden worden geselecteerd op basis van een enquête en in deze enquête wordt gepeild naar hun kijkgedrag, geldt veel van de kritiek die kan worden gegeven op enquêtes, ook voor automatische kijkcijfermeting. Met andere woorden: *de externe validiteit van de kijkcijfers kan nooit groter zijn dan die van een enquête.*



## DE BRUIKBAARHEID VAN KIJKCIJFERS EN ENQUÊTEGEGEVENS

Kijkcijfers en enquêtegegevens zijn geen concurrenten. Afhankelijk van de onderzoeksvraag zal één van beide vormen de beste meetmethode zijn.

Kijkcijfers vormen een enorme schat aan zeer nauwkeurig geregistreerde gegevens over het kijkgedrag van een grote, relatief representatieve groep mensen. Gebruik maken van deze cijfers om een top 20 op te stellen is niet altijd even zinvol. Vermits zulke rangschikkingen meestal worden voorgesteld alsof ze op *alle* kijkers slaan, gaat het om schattingen, die een betrouwbaarheidsinterval vereisen. Mochten zulke betrouwbaarheidsintervallen standaard worden weergegeven, dan zou wel eens kunnen blijken dat bijvoorbeeld de 'top 5' onderling verwisselbaar is.

De komst van kleine zenders heeft geleid tot een onevenredige aandacht voor veralgemeningen op basis van het kijkgedrag van minuscule subgroepen uit de steekproef. Deze cijfers zijn hoogst onbetrouwbaar. Niet alleen zijn de betrouwbaarheidsintervallen veel groter, de validiteit van veralgemeningen naar het niveau van de bevolking zijn op zijn minst bediscussieerbaar.

Kijkcijfers zijn echter wel uitermate geschikt voor het schatten van het kijkgedrag van grote groepen. Gelet op de inspanningen om de steekproef representatief te maken, kan men er vrij zeker van zijn dat de 'top 20' inderdaad twintig van de meest bekeken programma's omvat. Conclusies gebaseerd op grote aggregaten (verschillen tussen mannen en vrouwen, hoger en lager opleiden, enzovoort) zijn ongetwijfeld vrij accuraat en zeer waardevol.

Kijkcijfers worden in de eerste plaats gebruikt voor het berekenen van het 'marktaandeel' van zenders en programma's. De onderzoeksafdelingen van deze zenders hebben doorgaans niet de tijd en de middelen om zulke cijfers aan een meer diepgaande analyse te onderwerpen. Onder meer Sang et al. (1994) wijzen erop dat op die manier een enorm databestand met zeer gedetailleerde longitudinale gegevens onderbenut blijft. De auteurs wijzen erop hoe zelfs eenvoudige tijdreeksenanalyse (o.a. ARIMA) tot interessante en vernieuwende inzichten kan leiden. Ook voor de nauwkeurige analyse van het gedrag van de kijkers van één programma zijn deze gegevens, althans door academici, nog nauwelijks gebruikt. Zelfs de rudimentaire schakelanalyse, die Plees (1994) uitvoerde voor één programma (dat op twee kanalen tegelijkertijd werd uitgezonden), levert al boeiende resultaten op.

Het is echter niet zo dat kijkcijfermeting automatisch de beste methode is. Het gebruik van enquêtegegevens door academici is in de eerste plaats ingegeven door een gebrek aan financiële middelen (het verzamelen van setmeter data kost tientallen miljoenen per jaar). Dit betekent echter niet dat kijkcijfergegevens in alle gevallen een beter alternatief zouden zijn. Veel academisch onderzoek is immers geïnteresseerd in het bestuderen van televisiekijken op de langere termijn. Het is zeer de vraag of kijkcijfergegevens of gelijkaardige meetmethoden hierbij erg nuttig zijn. Het is bijvoorbeeld erg moeilijk om uit kijkcijfergegevens individuele 'programmatrouw' af te leiden (Webster en Wakshlag, 1982:447). Afhankelijk van de bestudeerde periode komt men steeds tot andere conclusies met betrekking tot kijkgedrag, selectiviteit, kijkerstypes, enz. (cf. Heeter en Greenberg, 1988; Webster en Wakshlag, 1982:447; Gunter, 1985:94; Van den Bulck, 1995:151). Er zijn namelijk heel wat veranderlijke factoren die bepalen of een bepaald individu op een bepaalde dag een bepaald programma zal zien, onafhankelijk van zijn of haar voorkeuren (voor een overzicht, cf. Van den Bulck, 1995). Zo blijken zelfs weersomstandigheden (Vandebosch en Roe, 1996) of de menstruele cyclus van vrouwelijke kijkers (Zillmann en Bryant, 1985:181) goede predictoren van het kijkgedrag te zijn. Als een gevolg van deze factoren kan het kijkgedrag op één dag of zelfs in één bepaalde periode niet zomaar worden veralgemeend om tot conclusies over de *kijkgewoontes* van de respondent te komen. De dure en omslachtige methode om mensen gedurende een week per kwartier in een dagboek te laten noteren waar ze wanneer en hoelang naar keken, is dan ook niet zo nuttig en betrouwbaar als zou kunnen lijken. Rosengren en Windahl (1989:30) draaien de vraag zelfs om, en stellen dat dagboek- en kijkcijfergegevens alleen bruikbaar zijn, als kan worden aangetoond dat de informatie die door de respondenten wordt gegeven, accuraat is, de referentieperiode waarin de studie werd uitgevoerd, een 'gemiddelde' periode is (als zoiets al bestaat en te definiëren is), en als bovendien het kijkgedrag van de respondent in die periode zijn of haar 'gemiddelde kijkgedrag' weerspiegelt. Als aan deze voorwaarden niet is voldaan, zijn de schijnbaar nauwkeurige gegevens niet meer dan accuraat gemeten, maar zeer vage indicatoren van de *kijkgewoontes*.

Enquêtegegevens zijn dus niet per definitie minderwaardig, zelfs integendeel. Bovendien leidt het inzicht dat ook kijkcijfers slechts indicaties (binnen de marges van een betrouwbaarheidsinterval) geven van het kijkgedrag van de bevolking en onderhevig zijn aan grote schommelingen, tot de conclusie dat ook een enquête tot de slotsom zou komen dat *Wittekerke*, *FC De Kampioenen* en *Jambers* tot de meest populaire programma's behoren. Dit betekent echter niet dat enquêtegegevens niet met grote argwaan moeten worden

bekeken. Afhankelijk van de gebruikte methode lopen de schattingen van kijktijd en kijkgewoontes sterk uiteen. Bovendien werd boven reeds aange- toond dat er grote aanwijzingen zijn dat minstens een gedeelte van de res- pondenten moeite heeft met het nauwkeurig inschatten van hun kijkgedrag. Onderzoekers zijn er in het verleden te vaak en te snel van uitgegaan dat ver- schillende onderzoekers hetzelfde bestudeerden, omdat ze allemaal het begrip 'kijkvolume' hanteerden.

## AANBEVELINGEN

Omwille van de grote variëteit aan methodologieën en de invloed die er van uitgaat, is het duidelijk dat elke onderzoeker zijn televisievariabelen nauw- keurig en expliciet moet definiëren. In onderzoek dat gebruik maakt van concepten als 'kijktijd', 'kijkvolume', 'kijkgewoonte', enzovoort, moet worden aangegeven op welke manier de resultaten werden bekomen. Allen (1981) en Salomon en Cohen (1978) wijzen erop dat alle televisieonderzoekers min- stens impliciet een 'televisietheorie' hanteren bij het uitvoeren van onderzoek, maar dat ze zich zelf vaak niet bewust zijn van de wetenschappelijke en theoretische *constructs* die door hun variabelen worden geïmpliceerd. Con- creet houdt dit bijvoorbeeld in dat een onderzoeker die alleen vraagt 'hoeveel uren kijkt u 's avonds televisie?', er van uitgaat dat mensen elke avond even- veel televisiekijken.

De oplossing van het Zweedse *Media Panel Programme* (cf.: Rosengren en Windahl, 1989) houdt in dat wordt gevraagd naar kijktijd *en* kijkfrequentie. Bovendien wordt rekening gehouden met de grote verschillen die bestaan tussen onder andere het weekend en een gewone weekdag. Deze benadering werd met andere oplossingen vergeleken in Van den Bulck (1996) en bleek beter te voldoen aan de validiteitsvereisten. Een mogelijke kritiek op deze werkwijze is dat ze op zijn minst een achttal vragen vooronderstelt (*hoeveel uren kijkt u op een weekdag, vrijdag, zaterdag, zondag, en hoe vaak kijkt u op een weekdag, vrijdag, zaterdag, zondag*). Vermits vragenlijsten altijd zo kort mogelijk moeten worden gehouden, betekent dit dat andere interessante onderzoeksvragen wellicht zouden moeten wijken. Uit dezelfde studie (Van den Bulck, 1996) blijkt dat ruwere benaderingen van het kijkvolume verge- lijkbare resultaten opleveren. Voor zulke (traditionele) benaderingen zijn slechts één of twee vragen nodig. Men zou zich kunnen afvragen of het dan niet economischer is om slechts de ruwe schatting te gebruiken. Zowel de constructvaliditeit als de externe validiteit van een onderzoek zijn echter essen- tiële elementen van de wetenschappelijke waarde ervan (cf. Mulac en Kunkel, 1989, voor een goede bespreking vanuit communicatiewetenschappelijk

oogpunt). Zo worden attitudes doorgaans gemeten aan de hand van een batterij vragen, waaruit de attitude dan als achterliggende *latente* variabele wordt geëxtraheerd. Besparen op de lengte van de vragenlijst kan een reden zijn om een attitude niet te bevragen, maar nooit om een goede methode te vervangen door een minder nauwkeurige. Voor televisie-onderzoekers zijn de kijkvariabelen de essentie van hun studies. Het is belangrijk dat er zoveel mogelijk aandacht naartoe gaat.

Tot slot zou het interessant zijn om de leden van het kijkerspanel via een schriftelijke of face-to-face enquête uitgebreid te bevragen over hun televisiekijkgedrag. Deze gegevens zouden dan kunnen worden gerelateerd aan een longitudinale studie van de kijkcijfers van dezelfde respondenten. Er zou kunnen worden nagegaan in welke mate kijkgedrag en kijkgewoonte overeenstemmen, en welke vraagstelling het dichtst bij de resultaten van het kijkcijferonderzoek staan. Bovendien zou nauwkeurig kunnen worden gemeten hoe goed kijkers hun eigen kijkgedrag kunnen inschatten, en wat goede indicatoren daarvan zijn. De mate waarin beide methodes onderling verwisselbaar zijn, zou kunnen worden nagegaan. Ook het automatische kijkcijferonderzoek zelf zou ermee gebaat zijn. Via zulke studies zouden immers accurate manieren kunnen worden ontwikkeld om via enquêtes naar actueel kijkgedrag te peilen. Op die manier zou de externe validiteit van de kijkcijfermeting worden verstevigd. Uit studies aangehaald door Sang et al. (1994) zou blijken dat zo'n 'interventie' in het panel geen groot en geen langdurig effect op de kijkcijfers zou hebben. Beide methoden van onderzoek hebben er dus alleen maar bij te winnen.

#### LITERATUURLIJST

- Allen, R.L. (1981) 'The Reliability and Stability of Television Exposure', *Communication Research*, 8 (2): 233-256.
- Chaffee, S.H. & Mutz, D.C. (1988) 'Comparing Mediated and Interpersonal Communication', in R.P. Hawkins, J.M. Wiemann & S. Pingree (eds.) *Advancing Communication Science: Merging Mass and Interpersonal Processes*. Newbury Park: Sage.
- Dahms, H. (1983) 'Wie Zuschauer fernsehen - zur Qualität des "Fernschkontaktes"', *Media Perspektiven*, 4: 279-286.
- den Boon, A. (1993) 'De Betrouwbaarheid van Kijkcijfers: een Onderzoek naar de Factoren die de Betrouwbaarheid van Kijkcijfers Beïnvloeden', *Massacommunicatie*, (4): 262-288.
- Grimes, T. & Meadowcroft, J. (1995) 'Attention to Television and Some Methods for Its Measurement', in B.R. Burleson (ed.) *Communication Yearbook nr. 18*. Thousand Oakes: Sage.
- Gunter, B. (1985) 'Determinants of Television Viewing Preferences', in D. Zillmann & J. Bryant (eds.) *Selective Exposure to Television*. Hillsdale: Lawrence Erlbaum.
- Hays, W.L. (1988) *Statistics*. Fort Worth: Holt, Rinehart and Winston.

- Heeter, C. (1985) 'Program Selection with Abundance of Choice: A Process Model', *Human Communication Research*, 12 (1): 126-152.
- Lang, A. (ed.) (1994) *Measuring Psychological Responses to Media Messages*. Hillsdale: Lawrence Erlbaum
- Lutz, B. (1987) 'Die Zuverlässigkeit elektronisch gemessener Fernsehnutzungsdaten. Untersuchung zur Reliabilität des GfK-Meters', *Jahrbuch der Absatz- und Verbrauchsforschung*, 33 (3): 226-241.
- Mulac, A. & Kunkel, D. (1989) 'Methodological Issues in the Study of Message Effects', in J.J. Bradac (ed.) *Message Effects in Communication Science*. Newbury Park: Sage.
- Pedhazur, E.J. & Pedhazur Schmelkin, L. (1991) *Measurement, Design and Analysis, An Integrated Approach*. Hillsdale: Lawrence Erlbaum.
- Plees, Y. (1994) *Zappen, Sociologische en Technologische Aspecten, een Profileren van de Vlaamse Zapper en de Reacties van de Media*. Leuven: departement communicatiewetenschap, eindverhandeling.
- Rosengren, K.E. (1994) 'Culture, Media and Society: Agency and Structure, Continuity and Change', in K.E. Rosengren (ed.) *Media effects and Beyond: Culture, Socialization and Lifestyles*. London: Routledge.
- Rosengren, K.E. & Windahl, S. (1989) *Media Matter, T.V. Use in Childhood and Adolescence*. Norwood: Ablex.
- Salomon, G. & Cohen, A.A. (1978) 'On the Meaning and Validity of Television Viewing', *Human Communication Research*, 4 (3): 265-270.
- Sang, F., Schmitz, B., Bretz, H.J., Stanat, P. & Tasche, K. (1994) 'Time-Series Analysis of Longitudinal People Meter Data', *European Journal of Communication*, 9: 91-106.
- Shrum, L.J. (1995) 'Assessing the Social Influence of Television, A Social Cognition Perspective on Cultivation Effects', *Communication Research*, 22 (4): 402-429.
- Shuman, H. & Presser, S. (1981) *Questions and Answers in Attitude Surveys, Experiments on Question Form, Wording and Context*. New York: Academic Press.
- Thorson, E. (1994) 'Using Eyes on Screen as a Measure of Attention to Television', in A. Lang (ed.) *Measuring Psychological Responses to Media Messages*. Hillsdale: Lawrence Erlbaum.
- Vandebosch, H. & Roe, K. (1996) 'Weer of geen weer, kijk of kijk (met) meer', *Communicatie*, 25 (3):
- Van den Bulck, J. (1995) 'The Selective Viewer: Defining (Flemish) Viewer Types', *European Journal of Communication*, 10 (2): 147-177.
- Van den Bulck, J. (1996) 'Veelkijkers versus Vaakkijkers: Betrouwbaarheid en Validiteit bij het Meten van het Kijkvolume', *Tijdschrift voor Communicatiewetenschap*, 24 (1): 40-56.
- Webster, J.G. & Wakshlag, J.J. (1982) 'The Impact of Group Viewing on Patterns of Television Program Choice', *Journal of Broadcasting*, 26 (1): 445-455.
- Webster, J.G. & Wakshlag, J.J. (1985) 'Measuring Exposure to Television', in D. Zillmann & J. Bryant (eds.) *Selective Exposure to Television*. Hillsdale: Lawrence Erlbaum.
- Weimann, G., Brosius, H.-B. & Wober, M. (1992) 'T.V.-Diets: Towards a Typology of T.V. Viewership', *European Journal of Communication*, 7: 791-515.
- Wiedemann, J. (1988) 'Das GfK-Meßsystem, geprüft und für gut befunden, Ergebnisse der Kontrolluntersuchungen zur Reliabilität der GfK-Daten', *Media Perspektiven*, 6: 366-374.
- Zillmann, D. & Bryant, J. (1985) 'Affect, Mood, and Emotion as Determinants of Selective Exposure', in D. Zillmann & J. Bryant (eds.) *Selective Exposure to Communication*. Hillsdale: Lawrence Erlbaum.