

De volledigheid van informatie over feiten in het verleden zoals gegeven door ondervraagden bij een enquête

J. M. M. van Tulder

1 Inleiding

De wijze waarop bij een sociologisch onderzoek gegevens worden verkregen is behandeld in vele boeken en artikelen. Het onderwerp is echter van zo fundamenteel belang, dat daarover ongetwijfeld nog veel geschreven zal worden en ook zal moeten worden. Uiteraard zal de aandacht zich dan steeds meer richten op detailspecten, die bij het verkrijgen van de gegevens van belang kunnen zijn. In dit artikel wordt een dergelijk detailspect onder de loupe genomen, namelijk het probleem van de volledigheid van de informatie over feiten in het verleden bij een vraagggesprek.

De bij een onderzoek te verzamelen gegevens kunnen in drie groepen worden ingedeeld: feiten, meningen en kennis.

Enige voorbeelden:

feiten: — hoe oud bent U?

— bezit U een bromfiets?

meningen: — bent U het eens of niet eens met de volgende uitspraken: (—)

— bent U van plan dit jaar een bromfiets te gaan kopen?

kennis: — heeft U gisterenavond tussen tien voor acht en acht uur naar de televisie gekeken (feit), en zo ja: weet U nog welke reclamespots er toen zijn uitgezonden?

Indien nu op een bepaald moment een vraagggesprek plaats vindt, dan kan men — theoretisch gesproken — voor gegevens behorende tot elk van de drie groepen, vragen naar de situatie op het moment van ondervraging en naar de situatie zoals die was ergens in het verleden. In het laatste geval geldt veelal dat de mate van nauwkeurigheid van het weergeven van die situatie afhankelijk is van het tijdsinterval tussen het 'nu' en het tijdstip in het verleden waarvoor men de situatie wil meten.

De vraag is nu of het mogelijk is een inzicht te krijgen in de invloed van dit tijdsinterval op de nauwkeurigheid van de meting.

Vooropgesteld zij dat de nauwkeurigheid van de meting niet alleen afhankelijk is van het tijdsinterval maar ook van een nog onbekend aantal variabelen die samenhangen met het object van de situatie die we willen meten.

Hierbij is het onderscheid in een feit-object, een meningsobject en een kennisobject primair.

In dit artikel blijven de laatstgenoemde twee objecten buiten beschouwing. Het vragen naar meningen en kennis is niet op één lijn te stellen met het vragen naar feiten en beperking is een eerste voorwaarde voor het verkrijgen van een inzicht in de onderhavige materie.

2 De nauwkeurigheid van de meting van een niet controleerbaar feit bij ondervraging van meerdere personen

Indien we onze gedachten over 'feiten' laten gaan, dan blijkt spoedig dat, gezien het doel van de overwegingen, een splitsing in twee groepen moet worden gemaakt en wel in:

- a controleerbare feiten en
- b niet-controleerbare feiten.

Deze tweedeling komt voort uit de gedachtengang dat sommige feiten ook na verloop van een zeker tijdsinterval nog gecontroleerd kunnen worden, terwijl andere feiten dan niet meer controleerbaar zijn. Een tweetal voorbeelden, waarop we hierna verder ingaan, kunnen dit verduidelijken.

Vragen we of iemand een televisietoestel of een bromfiets heeft, dan zal dit feit vrijwel altijd controleerbaar zijn. Vragen we evenwel of iemand in de afgelopen maand een bromfiets heeft gekocht dan kan dit feit gecontroleerd worden (tonen van de rekening) maar de situatie is al wat moeilijker.

Vragen we of iemand bepaalde bladzijden van een tijdschrift heeft gezien of gelezen dan is een echte controle van dit feit niet mogelijk omdat het lezen niet iets concreets heeft opgeleverd. Het feit is mogelijk alleen nog meetbaar aan de hand van de kennis over de inhoud van de desbetreffende bladzijden.

Een ander voorbeeld heeft betrekking op de verplaatsingen van iemand, als hij of zij de deur uitgaat. Gaan we deze verplaatsingen na door middel van een vraaggesprek dan is het feit van de verplaatsing reeds geschied, zodat een controle op de individuele antwoorden niet meer mogelijk is. Toch is er verschil tussen dit voorbeeld en dat van de bladzijden die gezien of gelezen zijn. Verplaatsingen worden vaak herhaald en kunnen dan beter worden weergegeven; het zien of lezen van bepaalde bladzijden van een tijdschrift geschiedt éénmaal of slechts enkele malen.

Het eerste voorbeeld heeft vaste referenties (de punten van de gemeente, die men bij de verplaatsing passeert), het tweede voorbeeld heeft variabele referenties (de, in de loop van de tijd qua inhoud steeds wisselende, bladzijden van een tijdschrift).

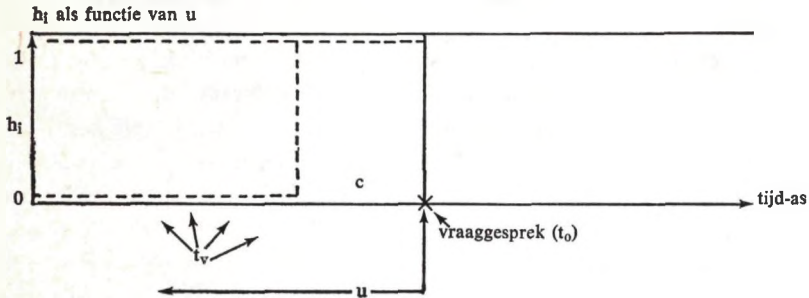
Het is duidelijk dat voor feiten in het heden het onderscheid in twee groe-

pen niet bestaat. Immers valt het moment van het vraaggesprek samen met het moment van het gebeuren (in feite 'observatie') dan is daarmee het feit gecontroleerd. Is daarentegen de observatie van het feit niet meer mogelijk dan zal de controle aan de hand van andere, op het tijdstip van het vraaggesprek wel waarneembare gegevens moeten gebeuren.

Zolang dit kan zal de weergave van het feit juist zijn. Is echter een directe controle niet meer mogelijk (zoals bij verplaatsingen) dan zijn we geheel aangewezen op het geheugen van de ondervraagde personen en zal de weergave van het feit meer of minder onnauwkeurig zijn. Opgemerkt zij hierbij dat dit niet behoeft te betekenen dat bij ondervraging van meerdere personen het gemiddelde van de gegeven antwoorden fout moet zijn.

Bij een onnauwkeurige weergave van de feiten kan het voorkomen dat het feit wordt overschat of wordt onderschat. Indien deze twee soorten mis-schattingen tegen elkaar opwegen, zal het totaalresultaat de werkelijkheid dicht kunnen benaderen. Evenwel de nauwkeurigheid van het eindresultaat is minder dan het geval zou zijn indien alle individuele antwoorden juist zouden zijn.

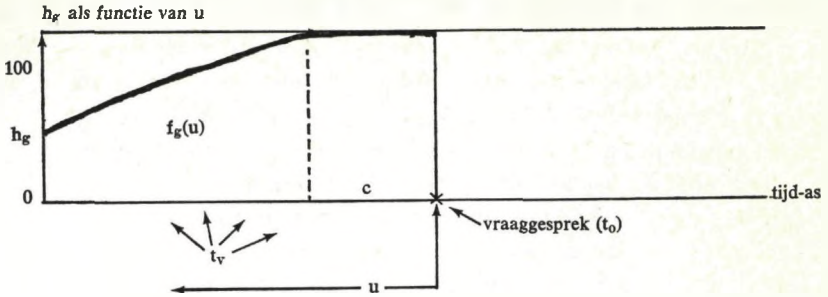
Vatten we het bovenstaande grafisch samen, dan ontstaat het volgende beeld:



Hierin is L_i de mate van nauwkeurigheid van de weergave van een feit door één individu en is u het tijdsinterval t_0-t_v , waarbij t_0 staat voor het moment van het vraaggesprek en t_v voor het moment van het gebeuren (in het verleden). Het tijdsinterval waarin controle op het feit nog mogelijk is wordt weergegeven met het symbool c .

Voor $u < c$ is dus controle mogelijk en is h_i gelijk aan 1. Voor $u > c$ kan h_i gelijk aan 1 zijn, doch ook $h_i = 0$ wordt dan mogelijk. Deze c zal niet voor elk individu even lang zijn. Bij een roker kan het voorkomen dat hij op de vraag naar het merk dat hij rookt geen pakje kan tonen. h_i kan dan 0 of 1 zijn.

Bij ondervraging van meerdere personen is het beeld als volgt:



Hierin is L_g de mate van nauwkeurigheid van de weergave van een feit door honderd individuen (m.a.w. het percentage juiste antwoorden).

Het verloop van de lijn $f_g(u)$ wordt bepaald door de verhouding van het aantal personen met $h_i = 1$ t.o.v. het totaal aantal personen.

Nogmaals, dit wil niet zeggen dat het totaalresultaat van de antwoorden op de vraag naar het feit verkeerd moet zijn.

Dit is mede afhankelijk van de fout in de antwoorden van personen met $h_i = 0$. Is hier een systematische fout aanwezig dan is het totaal-antwoord vertekend, waarbij kan worden opgemerkt dat deze systematische fout minder invloed heeft naarmate $h_g = f_g(u)$ dichter bij 100 ligt. Wat dit betreft is het dus van belang u zo klein mogelijk te nemen, liefst korter dan c . Voor sommige feiten is c echter gelijk aan 0 (alleen observatie helpt dan), met als extra complicatie dat de h_g curve soms zeer snel kan dalen (van rechts naar links gezien). In een dergelijke situatie moeten we proberen door middel van hulpvragen, die voor de ondervraagde persoon een referentiekader kunnen vormen voor het beantwoorden van de vragen waarom het gaat, de waarde van h_g te verbeteren. Dit zal niet altijd eenvoudig zijn, aangezien we de mogelijkheid niet mogen uitsluiten dat dergelijke hulpvragen bij de personen die toch $h_i = 0$ houden, leiden tot een verkeerde beantwoording van de vraag die systematisch is, waardoor, als de waarde van h_g niet aanzienlijk verbetert, het middel erger zou zijn dan de kwaal.

Bij dit alles mag de betekenis van de variabele u niet uit het oog worden verloren. Afhankelijk van de vorm van de curve $h_g = f_g(u)$ kan soms een kleine verkorting van de periode $u = t_0 - t_v$ reeds een aanzienlijke verbetering van het aantal juiste antwoorden opleveren.

Van veel belang is hierbij, dat het al of niet optreden van een systematische fout bij personen met $h_i = 0$ vaak samenhangt met u waarbij de richting van de systematische fout voor kleine u anders dan die voor grote u . In het navolgende wordt hier nader op ingegaan.

3 De mogelijkheid van het optreden van een systematische fout in onjuiste antwoorden

Zoals bekend, is het begrip voor tijdsintervallen bij vele mensen zeer zwak. Een periode van een dag kan men nog overzien; voor een week wordt het voor een aantal mensen al moeilijk om deze van het begin tot het einde juist te plaatsen. Eén maand is in de gedachten van de één, een periode van vier weken, voor een ander is het een vaag begrip — wat betreft de tijd

— waarbij het soms lastig is, na te gaan of een bepaald feit zich al of niet in precies die periode heeft voorgedaan. Bij gebeurtenissen die met niet te grote frequentie voorkomen (bijv. tweemaal in de drie maanden) kan men vaak de grens van de periode van een maand niet goed trekken en meent dan, dat de gebeurtenis de laatste keer binnen de afgelopen maand zal hebben plaatsgevonden, ook al was dit niet het geval. Hierdoor zal een over- of onderschatting van het gevraagde feit optreden.

Op grond van de ervaring gedurende vele jaren door schrijver dezes opgedaan kunnen de volgende indicaties worden gegeven. Bij een klein tijdsinterval is het soms een overschatting, bij een wat groter tijdsinterval een onderschatting omdat men feiten vergeet. Bij een zeer groot tijdsinterval geeft de ondervraagde vaak een antwoord dat hijzelf heeft verkregen door de frequentie van de feiten over een korte periode, die hij meent te kunnen overzien, te vermenigvuldigen met het aantal van die perioden in het zeer grote tijdsinterval.

Een voorbeeld hiervan is het bioscoopbezoek. Vraagt men geregelde bioscoopbezoekers naar het aantal malen per jaar dat zij naar de bioscoop gaan dan redeneren ze als volgt: 'Ik ga elke week éénmaal, dus dat is ongeveer 50 maal per jaar'. In werkelijkheid gaat men dan tussen de 40 en 45 maal per jaar. Dit komt omdat men de inderdaad bestaande gewoonte om éénmaal per week te gaan, gemakshalve met 50 vermenigvuldigt en dan meent met het weglaten van de resterende 2 weken in een jaar voldoende gecorrigeerd te hebben voor het overslaan van weken. Men is niet in staat de periode van een jaar te overzien wat betreft het overslaan en onderschat dit dan.

Helaas is over het optreden van deze systematische fout weinig bekend. Ten dele komt dit doordat het een gecompliceerde zaak betreft waarbij herinneringsverlies, periodeschattingfouten, frequentie van het voorkomen van het feit, uiteraard het soort feit en andere variabelen een belangrijke rol spelen. Van de andere kant koestert de onderzoeker vaak de 'valse' hoop dat het systematische effect zo erg niet is, zeker niet als hij vermoedt dat het aantal mensen dat een verkeerde opgave doet relatief klein zal zijn.

4 Het probleem nader bepaald

Bestaat het vermoeden van het optreden van relatief veel onjuiste antwoorden dan zal de onderzoeker, zoals reeds gezegd, het tijdsinterval tussen het optreden van het feit en het vragen ernaar inkorten, waarbij hij echter niet weet tot hoever hij terug moet gaan, wil hij:

- a een voldoende groot aantal juiste antwoorden krijgen en
- b bij de resterende personen met onjuiste antwoorden geen systematische fout in een hem onbekende richting introduceren.

Het is daarom gewenst dat systematisch onderzocht wordt: (1) of er bepaalde wetmatigheden zijn op te stellen voor het verband tussen het aantal juiste antwoorden en het tijdsverloop tussen het zich voordoen van het feit en het vragen naar het feit, (2) op welke wijze de periode waarin het feit direct of indirect nog controleerbaar is verlengd kan worden en (3) bij welke tijdsinterval voor personen met onjuiste antwoorden de systematische fout in hun antwoorden zo klein mogelijk is.

Het probleem is hierbij vooral, dat de drie genoemde componenten vaak moeilijk uit elkaar te halen zijn. Echter; ook zonder kennis van de werking van de afzonderlijke componenten is het mogelijk een – weliswaar onvolledig – inzicht in de onderhavige materie te verkrijgen.

In het navolgende worden, in het kort, twee projecten besproken waar onderzoek werd gedaan naar feiten die niet controleerbaar waren en waarbij de invloed van de lengte van het tijdsinterval nagegaan kon worden.

Het eerste project dat door de N.V. Publex te Amsterdam en de N.V. v/h Nederlandse Stichting voor Statistiek te 's-Gravenhage samen was opgezet betrof passanten bij een aantal punten in de stad Utrecht.¹

Het tweede project is ontwikkeld in nauwe samenwerking tussen

N.V. v/h Nederlandse Stichting voor Statistiek,

Interact B.V.,

Novum B.V. en

Admedia B.V.

De ontwikkelde methode kreeg de naam NINA-methode en heeft betrekking op het meten van het zien of lezen van bepaalde pagina's van tijdschriften (Pagina - Konfrontatie Kans - Onderzoek).²

5 Het passanten onderzoek

In 1964 is in de stad Utrecht een onderzoek ingesteld met het doel gegevens te verzamelen over het aantal passages bij bepaalde punten in de stad. Hierbij was het niet voldoende alleen de resultaten van tellingen te hebben. Er moesten ook gegevens komen over de passanten zelf (geslacht leeftijd

¹ De volledige resultaten van het passanten onderzoek zijn neergelegd in een rapport: 'Onderzoek Publex aanplakgelegenheden, Utrecht 1964, Aantal passages en attentiewaarde'. Nadere inlichtingen hierover zijn te verkrijgen bij de schrijver van dit artikel.

² De volledige resultaten van het Pagina Konfrontatie Kans (P.K.K.) Onderzoek volgens de NINA-methode zijn neergelegd in de twee delen van het boek: Pagina Konfrontatie Kans (P.K.K.), Uitgegeven door Novum B.V. in de NOVUM-reeks (1974, Haarlem). Het eerste deel van het boek geeft de methode van het onderzoek en de verwerking van de verzamelde gegevens; het tweede deel geeft de belangrijkste uitkomsten van het onderzoek.

enz.). In verband met dit laatste moest gebruik gemaakt worden van de techniek van de mondelinge ondervraging.

Het onderzoek werd ingesteld in de periode van 18 juni t/m 4 juli 1964 bij een representatieve steekproef van 328 personen uit de populatie van personen van 15 jaar en ouder in de stad Utrecht. Het onderzoek duurde vier weken. In elke week werden ongeveer evenveel personen ondervraagd. Bij het vraaggesprek is gebruik gemaakt van een, in een proefenquête geteste vragenlijst. Om de passages te bepalen is een plattegrond van Utrecht gebruikt waarop 300 genummerde punten waren gezet. Gevraagd werd naar alle routes die een ondervraagde de dag voor de ondervraging en de dag daarvoor had gemaakt.

Deze routes werden dan op de kaart aangewezen, waarbij de enquêteurs de nummers van de punten die gepasseerd waren op routeformulieren noteerden. De verdeling van de gesprekken over de dagen van de week was zodanig dat elke dag — met uitzondering van vrijdag en zaterdag — zowel 'gisteren' als 'eergisteren' kon zijn. Aangezien op zondag niet werd geënuquêteerd, kon zaterdag alleen als 'eergisteren' en vrijdag alleen als 'gisteren' voorkomen.

We hebben hier te maken met feiten (het passeren van punten), die bij de ondervraging niet gecontroleerd konden worden en met een tijdsinterval van één dag, resp. twee dagen.

Twee vragen zijn nu van belang:

- 1 wat is de invloed van het tijdsinterval op de resultaten?
- 2 worden er bij een bepaalde tijdsinterval uitkomsten verkregen die goed zijn?

Voor het beantwoorden van de eerste vraag is een analyse gemaakt waarbij elke dag in de positie van 'gisteren' werd vergeleken met dezelfde dag, maar dan in de positie van 'eergisteren'.

Voor het beantwoorden van de tweede vraag kon gebruik gemaakt worden van de resultaten van een onderzoek dat een jaar eerder, in ongeveer dezelfde tijd en bij dezelfde weersomstandigheden was gehouden. Toen werden op tien punten de passanten gesteld (observatie). Dezelfde tien punten waren in 1964 in het nummersysteem opgenomen.

Voor 100 genummerde punten werd nagegaan hoeveel passanten (inwoners van de stad Utrecht) daar op een bepaalde dag zijn langs gekomen. Het resultaat was als volgt:

Gemiddeld aantal passages per gemiddelde dag bij 100 punten

volgens de opgave voor 'gisteren'	: 6,4
volgens de opgave voor 'eergisteren'	: 5,4

De marges bij deze gemiddelden (2 s) zijn 0,7 resp. 0,6. Het verschil van 1,0 is significant.

Uit een nadere analyse van de frequentieverdelingen van het aantal gepasseerde punten blijkt dat het verschil tussen het resultaat voor 'gisteren' en dat voor 'eergisteren' geheel kan worden verklaard uit het verschil tussen de percentages van de personen die geen van de 100 onderzochte punten waren gepasseerd.

	'gisteren'	'eergisteren'
0 punten gepasseerd	28 %	39 %
1 of meer punten gepasseerd	72 %	61 %
totaal	100 %	100 %
aantal passages van hen die punten zijn gepasseerd	6,4/0,72	5,4/0,61
	8,9	8,9

Hieruit kan de conclusie worden getrokken dat bij verlenging van het tijdsinterval van één dag tot twee dagen de nauwkeurigheid van de meting teruggaat van niveau Z (100 of minder) tot ten hoogste Z - 11 (39-28 = 11). Z behoeft geen 100 te zijn, want ook bij de informaties over 'gisteren' kunnen reeds verkeerde informaties zitten.

Voorts is het goed mogelijk dat het niveau van de nauwkeurigheid lager is dan Z - 11. Het is namelijk niet uitgesloten dat ook bij hen die wel één of meer van de genummerde punten passeerden een toename van personen met onjuiste antwoorden aanwezig is. De door deze personen gedane opgaven bevatten evenwel geen systematische fout. Was dit wel het geval geweest dan was het gemiddeld aantal passages van hen die wel één of meer genummerde punten zijn gepasseerd, voor 'gisteren' en 'eergisteren' verschillend geweest.

Een analyse naar de dagen van de week toonde aan dat voor bijna elke dag een verschil tussen het aantal passages gisteren en eergisteren werd verkregen, ongeveer gelijk aan het verschil voor het totaal van alle dagen. Steeds was het percentage ondervraagden dat geen van de genummerde punten was gepasseerd voor 'eergisteren' groter dan voor 'gisteren'.

Wat is nu het werkelijk aantal passages? Het ligt voor de hand hier aan te nemen dat het kleinste tijdsinterval de werkelijkheid het meest zal doen benaderen.

Daarom werden de resultaten voor 'eergisteren' vermenigvuldigd met een

$$\text{factor } \frac{6,4}{5,4} = 1,2.$$

Na deze correlatie werd het resultaat als volgt:

	gemiddeld aantal punten dat een ondervraagde passeerde	gemiddelde x populatie ³ (x 1000)
op maandag	5,7	1.100
op dinsdag	6,1	1.200
op woensdag	7,8	1.500
op donderdag	7,7	1.500
op vrijdag	7,5	1.500
op zaterdag	6,2	1.200
op zondag	5,1	1.000
totaal per week	46,1	9.000

Bij een populatie van 196.000 personen van 15 jaar en ouder betekent dit 9.000.000 passages per week bij de 100 geselecteerde punten. Dit is 90.000 per punt per week.

De tien punten die betrokken waren in het onderzoek van 1963 bleken 12 % van de passages, zoals vastgesteld in 1964 bij de 100 punten, te omvatten.

Voorts was bij het onderzoek van 1963 geobserveerd van maandag t/m zaterdag.

Laten we voor 1964 de zondag buiten beschouwing dan komen we tot 8.000.000 passages per week. 12 % hiervan is 960.000. In 1963 werden bij deze 10 punten 1.000.000 passanten geteld (12 jaar en ouder, waarbij de leeftijd door de tellers werd geschat).

Gezien dit resultaat is er voorlopig geen reden om aan te nemen dat de keuze van u is één dag bij onderzoeken als het onderhavige, niet goed zou zijn.

Vatten we de resultaten samen, dan is het beeld als volgt:

- 1 route reconstructies met behulp van nummers voorziene, plattegronden zijn mogelijk;
- 2 bij een tijdsinterval van één dag, worden de uitkomsten verkregen die overeenkomen met uitkomsten verkregen met de observatiemethode;
- 3 bij een tijdsinterval van twee dagen daalt de nauwkeurigheid met tenminste 11 punten;
- 4 de oorzaak van de daling van de nauwkeurigheid is vooral de stijging van het percentage personen dat — volgens de routeanalyse — geen van de genummerde punten gepasseerd is; een deel van hen heeft opgegeven niet buiten de deur zijn geweest.

³ 196.000 van 15 jaar en ouder

6 Het P.K.K. onderzoek

Bij het passantenproject was sprake van een meting van feiten die, volkomen identiek, steeds meer kunnen terugkeren.

In het nu te bespreken project is de situatie anders.

Het probleem bij het pagina konfrontatie kans onderzoek was: Hoeveel pagina's van een tijdschrift worden gezien of gelezen?

In het volgende blijft het hoofdprobleem van de definitie van 'zien' of 'lezen' buiten beschouwing. Met de probleemstelling van dit artikel heeft dit hoofdprobleem weinig of niets te maken. Belangstellenden zij verwezen naar het bij NOVUM B.V. uitgegeven boek (2).

Wat is nu een verschil met het vorige project dat voor de probleemstelling belangrijk kan zijn? Teneinde hierin een inzicht te krijgen verplaatsen we ons in gedachten naar iemand waar elke week een tijdschrift in huis komt en die daarin kijkt of leest. Op een bepaalde dag komt er een enquêteur vragen naar de bladzijden van dat tijdschrift die zijn gezien of gelezen.

Pagina na pagina wordt hiervoor doorgenomen en de ondervraagde moet dan van elke pagina aangeven of die gezien of gelezen is en wanneer dat was. Het tijdvak u is daartoe verdeeld in zeven perioden van $1/3$ dag (gisteravond, gisterenmiddag, gisterenochtend enz.) en de tijd daarvoor.

Deze methode werd gevolgd aangezien de vrees bestond dat bij het vragen zonder periodenindeling de ondervraagde ongenueanceerd zou gaan antwoorden en daardoor gemakkelijk tot verkeerde antwoorden zal komen. Immers, de verleiding is groot om te zeggen: 'Ik heb alles gezien of gelezen', temeer waar menigeen daarbij nog denkt dat dit waar is.

Door het vragen naar: 'Wanneer heeft U dat gezien?' worden 'wilde' antwoorden als hiervoor hopelijk vermeden.

Bij het P.K.K. project heeft een ondervraagde dus gemakkelijk een norm voor ogen (alle pagina's); bij het passanten onderzoek daarentegen is een dergelijke norm niet aanwezig. Men kan dan de passages van een bepaalde dag opgeven zonder de behoefte te hebben aan een bepaalde norm te voldoen. Bij het P.K.K. onderzoek is een dergelijke normering wel aanwezig in de vorm van de gedachte dat men in het tijdschrift toch alles of vrijwel alles gezien of gelezen heeft. Deze situatie nu maakt dat we zeer voorzichtig moeten zijn met de keuze van het tijdsinterval.

Voor 120 pagina's van een tijdschrift werd uitgerekend aan de hand van de opgaven, hoeveel er daarvan in een bepaalde periode waren gezien of gelezen. Vervolgens werden de laatste drie perioden bijeengenomen (gisteren) en eveneens de daaraan voorafgaande drie (eergisteren). De 7e periode van $1/3$ dag daarvoor en die van het tijdvak dat daar weer voor lag werden niet gecombineerd.

Het aantal pagina's 'gezien of gelezen' werd gedeeld door 120 waarna het

volgende resultaat werd verkregen:

	kans dat een pagina gezien of gelezen is
gisteren	0,16
eergisteren	0,17
1/3 dag daarvoor	0,17
eerder	0,14

Indien we bedenken dat gisteren en eergisteren drie perioden omvatten, terwijl de derde regel (1/3 dag daarvoor) één periode omvat dan volgt onvermijdelijk de conclusie dat de feiten niet juist zijn weergegeven. Als de eerste twee regels goed zijn, had de derde regel een kans van 0,06 moeten geven. Omgekeerd, zou de derde regel goed zijn, dan had op de eerste en tweede regel een pagina konfrontatie kans van 0,51 moeten staan.

Ook de kans 0,14 van 'eerder' is niet zoals we die zouden verwachten bij een gelijkmatige verdeling van het lezen over een periode van 7 dagen. Hiervoor is echter een verklaring te geven. De ondervraging vindt uiteraard niet altijd plaats 7 dagen nadat een tijdschrift in huis is gekomen. Gemiddeld zal drie à vier dagen na ontvangst van het tijdschrift het vraaggesprek plaats vinden. De periode 'eerder' omvat dus $3\frac{1}{2} \times 3 - 7$ is $3\frac{1}{2}$ periode van 1/3 dag. Per dag zou de pagina-konfrontatie kans dan zijn: $0,14 \times 3\frac{1}{2} = 0,12$ ofwel 0,04 per periode van 1/3 dag.

Uitgaande van een konfrontatie kans per dag van 0,165 resp. 0,51 had het beeld als volgt moeten zijn:

	bij 0,165	bij 0,51
gisteren	0,165	0,51
eergisteren	0,165	0,51
1/3 dag daarvoor	0,055	0,17
eerder	0,192	0,60

Zou de tweede kolom 'waar' zijn geweest dan zou over 7 dagen genomen elke pagina een konfrontatie kans van ongeveer 3,5 hebben gekregen, hetgeen betekent dat het *gehele* tijdschrift 3,5 x gezien of gelezen is. Dit nu is zeer onwaarschijnlijk, waar nog bij komt, dat in die situatie, drie van de vier resultaten fout zouden zijn opgegeven en één resultaat goed.

7 Conclusies

Voor de probleemstelling van dit artikel kunnen we nu enige conclusies trekken:

1 – Het doorlopen van de pagina's heeft een concretiserende werking, zodanig dat het herinneringsverlies pas bij een iets groter tijdsinterval optreedt dan bij een wat meer abstracte benadering zoals met de plattegrond van Utrecht.

2 – Voor bovenstaande conclusie pleit ook de omstandigheid dat bij het passantenonderzoek het aantal gepasseerde punten, bij hen die één of meer punten gepasseerd zijn, voor gisteren even groot was als voor eergisteren. Men zou kunnen zeggen: heeft een ondervraagde eenmaal opgegeven dat hij één of meer routes heeft afgelegd, dan krijgt hij – met de kaart erbij – een concreet beeld van zijn routes.

3 – Indien de ondervraagde een verwachting heeft van een soort volledigheid (alle pagina's) dan ontstaat er een effect waarbij voor periodes die wat verder terugliggen een overschatting van de feiten ontstaat.

4 – Indien het voorgaande juist zou zijn, dan dient de onderzoeker aan het tijdsinterval waarover hij meent nog betrouwbare informatie te verkrijgen, toe te voegen een daaraan voorafgaande periode waarin de ondervraagde gelegenheid krijgt zijn volledighedsdrang tot uiting te laten komen. Laat men deze extra periode – die voor de analyse uiteraard onbruikbaar is – weg, dan loopt de onderzoeker het risico dat de ondervraagde de feiten aan het einde van de opgenomen periode gaat vertekenen.

Het bovenstaande is – zeker wat betreft de conclusies – beperkt. Het bevat evenwel elementen die, naar de mening van de schrijver, de moeite waard zijn om verder onderzocht te worden.