

Over goed en slecht peilen: de casus van de koopzondag

Jelke Bethlehem

Er wordt veel gepeild in Nederland. Dat is vooral merkbaar tijdens verkiezingscampagnes, waarin marktonderzoekers soms dagelijks met nieuwe peilingen komen. Ook de landelijke overheid peilt. Die peilingen zijn vaak groter en ingewikkelder dan de kleine en snelle politieke peilingen van de marktonderzoekbureaus. Ze worden meestal enquêtes of surveys genoemd. Ze verzamelen informatie over de stand van zaken in de maatschappij. Die informatie kan de basis vormen voor nieuw beleid. Lokale overheden peilen eveneens, zij het op een wat minder grote schaal dan de landelijke overheid. Van oudsher organiseren gemeenten omnibusenquêtes. Daarin kunnen, zoals de naam al aangeeft, diverse onderwerpen aan de orde komen. Tegenwoordig zijn veel omnibusenquêtes vervangen door burgerpanels. Al die peilingen, enquêtes en surveys hebben gemeen dat op basis van een steekproef uit een bevolkingsgroep uitspraken worden gedaan over de bevolkingsgroep als geheel. Een dergelijke generalisatie is alleen mogelijk als de steekproef door loting is verkregen. Dit artikel beschrijft wat goede en wat slechte steekproeven zijn. Dit wordt geïllustreerd aan de hand van een uniek voorbeeld. Dat is het onderzoek naar de mening van de inwoners van Alphen aan den Rijn over koopzondagen. Daarvoor zijn op hetzelfde moment en met dezelfde vragenlijst drie verschillende peilingen uitgevoerd. Dit voorbeeld maakt duidelijk dat een verkeerde steekproef kan leiden tot onjuiste conclusies en dus misschien ook wel tot onjuiste beleidsbeslissingen.

1 Over peilen

1.1 Peilingen, surveys, enquêtes ...

Er wordt in Nederland veel gepeild. Vooral tijdens verkiezingscampagnes volgen opiniepeilingen elkaar in hoog tempo op. Er zijn dan verschillende marktonderzoekers actief, zoals Maurice de Hond (Peil.nl), Ipsos (Politieke Barometer), GfK Intormart (De Stemming), TNS NIPO en I&O Research. Andere peilingen, zoals de Peilingwijzer, peilen niet zelf, maar combineren bestaande peilingen tot een nieuwe peiling. Zulke peilingen worden soms 'polls of polls' genoemd. Het hoeft bij peilingen niet alleen te gaan om het meten van mogelijk stemgedrag. Ook buiten de verkiezingscampagnes wordt veel gepeild. Voor overheidsorganen als het Centraal Bureau voor de Statistiek (CBS) en het Sociaal en Cultureel Planbureau (SCP) is het 'core business' om gegevens te verzamelen ten behoeve van het beleid. Hun peilingen zijn vaak groot en complex. Ze worden meestal niet peilingen genoemd, maar surveys of enquêtes.

Gemeentelijke overheden peilen ook. Daarbij kan het gaan om het verzamelen van informatie over de sociale en economische toestand binnen de gemeente, maar ook over het inventariseren van de publieke opinie over zaken die in de gemeente spelen. Peilingen kunnen tevens een rol spelen bij New Public Management. Dat is een meer bedrijfsmatige wijze van besturen waarbij burgers en bedrijven als klanten worden gezien (zie onder meer Osborne & Gaebler, 1992: 172-180). Meetbaarheid van het functioneren van het bestuur vormt een belangrijk onderdeel van deze besturingsfilosofie. Dit kan bijvoorbeeld gestalte krijgen in de vorm van een tevredenheidspeiling bij de burgers.

Een bekend voorbeeld van een peiling in een gemeente is de *omnibusenquête*. Er zijn gemeenten waarin een dergelijke enquête al vele jaren lang als instrument bij beleidsvorming wordt gebruikt. Zoals de naam al aangeeft, kunnen in een omnibusenquête diverse onderwerpen aan de orde komen. In het verleden werden hiervoor papieren vragenlijsten gebruikt. Het kon voorkomen dat tientallen enquêteurs bij de geselecteerde adressen langs gingen om enquêteformulieren aan te reiken en ze later weer op te halen. Zie bijvoorbeeld Slot (2009) over de omnibusenquête in Amsterdam. Tegenwoordig hebben veel gemeenten een *burgerpanel*. Een dergelijk burgerpanel bestaat meestal uit een steekproef van inwoners van de gemeente die regelmatig hun mening geven over allerlei actuele lokale zaken.

Essentieel element van al die peilingen, surveys en enquêtes is dat informatie wordt verzameld bij slechts een steekproef van personen uit de doelgroep. Het raadplegen van alle inwoners zou immers kostbaar en tijdrovend zijn. Op grond van een steekproef kunnen betrouwbare en valide conclusies over de hele doelgroep worden getrokken, maar dan moet die steekproef wel aan bepaalde voorwaarden voldoen. De belangrijkste voorwaarde is dat de steekproef *representatief* moet zijn. Maar wat betekent die term 'representatief' eigenlijk? Helaas is dit niet altijd even duidelijk. Al in 1979 kwamen Kruskal & Mosteller (1979a, 1979b, 1979c) met een lijst van negen verschillende betekenissen. Om verwarring te vermijden is het dan ook onverstandig om de term 'representatief' te gebruiken zonder uit te leggen wat ermee wordt bedoeld. In dit artikel wordt een steekproef representatief genoemd als de steekproef wordt geloot en iedereen in de doelgroep van het onderzoek dezelfde kans heeft gehad om in de steekproef te komen. Soms wordt de term representatief ook wel gebruikt in de vorm van *representativiteit met betrekking tot een kenmerk*. Representatief met betrekking tot bijvoorbeeld leeftijd betekent dan dat de leeftijdsverdeling in de steekproef overeenkomt met de leeftijdsverdeling in de doelgroep. De ideale steekproef is representatief met betrekking tot alle kenmerken die in de peiling worden onderzocht.

Het idee dat een steekproef uit een doelgroep nauwkeurige informatie over de hele doelgroep kan opleveren, heeft maar moeizaam ingang gevonden. Aanvankelijk geloofde men niet dat het zo mogelijk was om zinnige conclusies te trekken over de gehele doelgroep. De discussie daarover duurde bijna veertig jaar, van 1895 tot 1934. Nu is iedereen het erover eens dat dit een goede methode van onderzoek is. Het is daarbij wel belangrijk dat de onderzoeker zich aan de regels

van goed onderzoek houdt. Wordt daarvan afgeweken, dan bestaat een ernstig risico dat verkeerde conclusies uit het onderzoek worden getrokken.

In 2015 deed zich de gelegenheid voor om verschillende manieren van peilen met elkaar te vergelijken aan de hand van een praktische casus. In de gemeente Alphen aan den Rijn was veel discussie over het wel of niet invoeren van koopzondagen. Een aantal politieke partijen besloot de mening van de inwoners te peilen. Op hetzelfde moment, en met dezelfde vragenlijst, werden drie verschillende peilingen uitgevoerd. Eén peiling voldeed aan de regels voor goede peilingen en de andere twee niet. Er kon worden vastgesteld in hoeverre de uitkomsten op elkaar leken of van elkaar afweken. Die analyse vormt een belangrijk onderdeel van dit artikel.

In de rest van dit artikel wordt in het kort beschreven hoe de theorie van het trekken van steekproeven is ontstaan en ook hoe die niet altijd correct wordt toegepast. Paragraaf 2 beschrijft hoe het kwam dat er in Alphen aan den Rijn drie peilingen tegelijkertijd werden gehouden, en wat precies de verschillen tussen die peilingen waren. In paragraaf 3 worden de uitkomsten van die peilingen geanalyseerd.

1.2 Terug in de tijd

Bestuurders hebben vanaf de vroegste tijden behoefte gehad aan informatie over hun landen of rijken. Dat was 1000 jaar voor Christus al zo bij de heersers in China en Egypte. Ook de Romeinse keizers organiseerden regelmatige tellingen van mensen en hun bezittingen. Die informatie werd vooral gebruikt om vast te stellen hoeveel inkomsten er konden worden gegenereerd uit belastingen en hoe groot de legers waren die men op de been kon brengen.

Dit was allemaal *integraal onderzoek*. Iedereen in de doelgroep moest meedoen. Het idee was nog niet opgekomen dat ook bruikbare informatie kon worden verkregen op basis van een steekproef.

Dit ging zo door tot het einde van de negentiende eeuw. In die negentiende eeuw waren het voornamelijk de nationale statistische bureaus die zich bezig hielden met het verzamelen van statistische gegevens. Voor hen was het trekken van steekproeven taboe. Waarom een steekproef trekken als het ook met een integrale telling kon? En bovendien was het ongepast om echte waarnemingen te vervangen door rekenkundige manipulaties. Het werd gezien als een vorm van discriminatie om een deel van de mensen uit te sluiten van deelname aan een onderzoek.

De doorbraak kwam in het jaar 1895. In dit jaar hield Anders Kiaer, de directeur van het Noorse Statistische Bureau, voor het eerst een pleidooi voor het gebruik van steekproeven. Hij betoogde dat je met zijn *Representatieve Methode* goede resultaten kon behalen (Kiaer, 1895). Hij construeerde een steekproef die in een aantal aspecten (geslacht, leeftijd en stad/platteland) leek op de doelgroep. Dit wordt nu een *quota-steekproef* genoemd. Hij maakte geen gebruik van loting. Kiaer kon niet aangeven hoe goed of slecht zijn schattingen waren. Andere statistici

vonden dit een ernstig nadeel. De oplossing werd aangedragen door Arthur Bowley (Bowley, 1906). Hij stelde voor om steekproeven te loten. Daardoor werd het mogelijk de theorie van de kansrekening toe te passen. Bowley toonde aan dat schattingen dan bij benadering een normale verdeling hebben. Daardoor kan de variantie van de schatter gebruikt worden als een maat voor de nauwkeurigheid. Jerzy Neyman werkte de theorie verder uit en bewees in 1934 een aantal fundamentele resultaten (Neyman, 1934). Zo introduceerde hij het *betrouwbaarheidsinterval* als instrument om de precisie van een schatting aan te geven. Dit betrouwbaarheidsinterval wordt nog steeds gebruikt bij het berekenen van de *onzekerheidsmarges* van de uitkomsten.

Deze ontwikkeling leidde tot de fundamentele principes van de steekproeftheorie. Ze staan samengevat in een artikel van Horvitz en Thompson (1952). Zij lieten zien dat je altijd valide (zuivere) schattingen van kenmerken in de doelgroep kunt maken als aan drie voorwaarden is voldaan:

- de steekproef moet zijn geloot uit de hele doelgroep;
- iedereen in de doelgroep moet een positieve kans op selectie hebben gehad;
- deze selectiekansen moeten bekend zijn.

Een dergelijke steekproef wordt een *kanssteekproef* of *aselecte steekproef* genoemd. De eenvoudigste en meest gebruikte aselecte steekproef is die waarin iedereen in de doelgroep dezelfde kans heeft om in de steekproef te komen. Dat is dus een representatieve steekproef.

Een aantrekkelijke eigenschap van kanssteekproeven is dat het altijd mogelijk is om de nauwkeurigheid van schattingen uit te rekenen. Daarmee wordt het mogelijk de ruis van de steekproef te scheiden van 'echte' effecten. Een meer gedetailleerde beschrijving van de geschiedenis van de steekproeftheorie is te vinden in Bethlehem (2009a) en Bethlehem (2013).

1.3 Nieuwe ontwikkelingen

De razendsnelle opkomst van het internet maakt het rond 1995 mogelijk om op een computer via het internet formulieren in te vullen. Zo konden dus ook vragenlijsten van enquêtes online worden aangeboden. Online enquêteren werd al snel populair. Dat is niet zo verwonderlijk, aangezien dit soort enquêtes op het eerste gezicht een aantal aantrekkelijke eigenschappen hebben:

- Het is een betrekkelijk eenvoudige manier om toegang te krijgen tot een zeer grote groep potentiële respondenten, namelijk iedereen met internet. De internet-dekking is hoog in Nederland. Volgens gegevens van Eurostat had 96% van de huishoudens toegang tot internet in 2014 (Eurostat, 2015).
- De vragenlijsten kunnen tegen zeer lage kosten worden aangeboden: er zijn geen enquêteurs nodig (zoals bij mondelinge of telefonische enquêtes) en er zijn geen drukkosten en verzendkosten (zoals bij papieren vragenlijsten).
- Een peiling kan snel worden uitgevoerd. Het hele proces van opzetten en uitvoeren van een online enquête hoeft maar weinig tijd te kosten. Het is soms

zelfs mogelijk om een peiling in één dag te doen. Zie bijvoorbeeld de ‘flitspeiling’ van het tv-programma *EenVandaag* (EenVandaag, 2015).

De relatieve eenvoud waarmee online peilingen kunnen worden opgezet, leidt tot een groot en nog steeds groeiend aanbod van dit soort peilingen. Er zijn websites (zoals bijvoorbeeld die van *SurveyMonkey*) waarmee iedereen in korte tijd een enquête in de lucht kan brengen, zonder ook maar enig verstand te hebben van de methodologische aspecten van peilingen. Veel van die enquêtes zijn niet op verantwoorde wijze opgezet. En door het grote aanbod is het moeilijk het kaf van het koren te scheiden.

Een van de lastigste problemen is het trekken van een aselechte steekproef voor een online peiling. Bij voorkeur zou er een lijst met e-mailadressen aanwezig moet zijn. Daaruit kan dan de steekproef worden geloot, waarna aan de geselecteerde personen een e-mail gezonden kan worden met daarin een link naar de vragenlijst. Helaas is een dergelijke lijst meestal niet beschikbaar. Een alternatief zou bijvoorbeeld kunnen zijn het trekken van een steekproef van namen en adressen uit het bevolkingsregister en het sturen van een brief per post met daarin een link naar de vragenlijst. Dit maakt het uitvoeren van een online peiling lastiger, duurder en tijdrovender. Daarmee gaat een aantal voordelen van de online peiling verloren. Ook is het bevolkingsregister niet voor iedereen beschikbaar. Wetgeving verbiedt veel organisaties het gebruik van informatie uit dit register, en dus ook het trekken van steekproeven hieruit.

De problemen met het loten van een steekproef voor een online peiling hebben ertoe geleid dat bij veel online peilingen het loten van de steekproef wordt omzeild. In plaats daarvan maken peilers gebruik van *zelfselectie*. Ze laten het aan de personen in de doelgroep over om zichzelf wel of niet aan te melden voor de peiling. Via reclame in de media, banners en ‘pop-up windows’ worden internetgebruikers op de peiling attent gemaakt. De peiler hoopt dan dat deze gebruikers zich daardoor laten overhalen om deel te nemen aan de peiling. Typische voorbeelden van zelfselectie zijn te vinden bij de politieke peilingen. De peiler heeft dan een panel opgezet en dat gevuld met mensen die zich daarvoor zelf hebben aangemeld. De mensen in dit panel worden regelmatig benaderd met peilingen. In zulke panels zitten vooral mensen die het leuk vinden om regelmatig aan peilingen mee te doen en geïnteresseerd zijn in politieke zaken. Onderzoek heeft aangetoond dat diverse groepen daardoor ondervertegenwoordigd zijn, zoals ouderen, laagopgeleiden en etnische minderheden. De representativiteit van dit soort panels is daarom op zijn minst twijfelachtig. Zie bijvoorbeeld Bethlehem en Biffignandi (2012) voor meer informatie hierover.

Een bijkomend probleem is dat mensen zich voor een peiling kunnen aanmelden die niet tot de doelgroep behoren. Ook kunnen personen zich soms bij peilingen meer dan één keer aanmelden, hetzij onder dezelfde identiteit hetzij onder een andere identiteit. Uiteraard kan de onderzoeker via een of meer vragen proberen dit soort zaken te ontdekken en te elimineren, maar dat is niet zo eenvoudig. Een handige respondent is in staat deze controles te omzeilen.

Een ander gevaar van zelfselectie is dat groepen personen kunnen proberen de uitkomst van de peiling te beïnvloeden. Zo'n poging tot manipulatie deed zich bijvoorbeeld voor bij een peiling voor de Tweede Kamerverkiezingen van 12 september 2012. Een groep van 2500 'infiltranten' probeerde zich aan te melden bij het panel van Maurice de Hond (Bronzwaer, 2012). Het idee van deze groep was om zich eerst te presenteren als CDA-stemmer en dan langzamerhand over te gaan naar de partij 50PLUS. De actie werd ontdekt omdat er ineens wel heel veel aanmeldingen tegelijk binnenkwamen. Het laat echter wel zien dat met een wat subtielere aanpak peilingen op basis van zelfselectie te manipuleren zijn.

Een ander voorbeeld van manipulatie deed zich voor tijdens de campagne voor de gemeenteraadsverkiezingen van 2014 in Amsterdam. Op 12 januari 2014 werd het Eerste Grote Lijsttrekkersdebat georganiseerd. Het Parool wilde weten hoe de lijsttrekkers het hadden gedaan en organiseerde een online peiling. De campagne-teams van twee partijen (SP en D66) merkten dat je na het uitzetten van cookies op je computer onbeperkt kon stemmen. De hele nacht bleven ze stemmen uitbrengen op hun lijsttrekker. Dus kregen de lijsttrekkers van deze partijen veel meer stemmen dan de lijsttrekkers van de andere partijen. Toen Het Parool merkte dat de peiling een heel erg afwijkende uitslag zou gaan opleveren, werd hij uit de lucht gehaald. De krant beschuldigde de politieke partijen van manipulatie van de peiling. Het was echter wel de krant zelf die een slechte peiling had opgezet.

Er zijn dus goede en slechte peilingen. Uit slechte peilingen kunnen wezenlijk verkeerde conclusies worden getrokken. Dat kan worden aangetoond door het toepassen van de nodige statistische theorie. Zie bijvoorbeeld Bethlehem (2009b). Soms doet zich echter ook in de praktijk de gelegenheid voor verschillende peilingen met elkaar te vergelijken. Dat was het geval bij de enquête over de koopzondagen in Alphen aan den Rijn in januari 2015. Op hetzelfde moment werden drie peilingen gehouden over hetzelfde onderwerp en met dezelfde vragenlijst, maar met verschillende manieren van peilen. In het volgende hoofdstuk worden die drie peilingen in detail met elkaar vergeleken.

2 De casus van de koopzondag

2.1 *Koopzondagen in Alphen aan den Rijn*

In de gemeenteraad van Alphen aan den Rijn was er in 2014 en 2015 discussie over koopzondagen. De meningen liepen ver uiteen. Zo wilde de VVD de winkeliers zelf laten beslissen over hun eigen openingstijden, terwijl de SGP en ChristenUnie de winkels op zondag gesloten wilden houden. Het college kwam met een voorstel om een keuze te maken uit drie varianten: twaalf koopzondagen per jaar, achttien koopzondagen per jaar of een onbeperkt aantal koopzondagen. De coalitiepartijen (CDA, Nieuw Eland, VVD, D66 en ChristenUnie) kwamen er niet uit. Daarom werd besloten om via een peiling aan de inwoners van Alphen aan den Rijn te vragen wat zij ervan vonden.

De peiling was een initiatief van de ChristenUnie. Zeven andere partijen (CDA, SP, NieuwElan, SGP, RijnGouweLokaal, D66 en GroenLinks) ondersteunden het initiatief. Vier partijen (VVD, PvdA, Beter Alphen en Groene Hart Democraat) deden niet mee.

Het oorspronkelijke idee van de politici was om inwoners op straat in de winkelcentra te gaan interviewen. Toen duidelijk werd dat dit geen representatieve steekproef zou opleveren, besloten de politici om ook een peiling te doen met het AlphenPanel, het online burgerpanel van de gemeente Alphen aan den Rijn. Om alle inwoners van Alphen aan den Rijn de gelegenheid te geven hun mening kenbaar te maken, werd ook nog eens een online peiling met zelfselectie georganiseerd. En zo waren er drie peilingen tegelijkertijd. In dit artikel worden de drie verschillende aanpakken met elkaar vergeleken.

2.2 Drie manieren van peilen

De eerste manier van peilen was het interviewen van de bezoekers in de winkelcentra in de gemeente. Op één dag, zaterdag 17 januari 2015, tussen 11.00 en 15.00 uur, konden de winkelende mensen een vragenlijst invullen. Ze werden daarvoor aangesproken door vrijwilligers van de deelnemende politieke partijen. Die stelden zich niet op als neutrale enquêteurs, maar waren duidelijk herkenbaar als leden van een politieke partij. Het is de vraag of dit geen negatieve invloed heeft gehad op de objectiviteit en representativiteit van de enquête. De enquêteurs bleken geen ervaring te hebben met enquêteren. Ze hadden geen gedetailleerde instructie gekregen. Zo bleken ze niet zo goed te weten wat ze moesten doen met winkelende personen uit andere gemeenten. Ze hadden daarvoor geen aanwijzingen gekregen. Door sommige enquêteurs werden ze wel geïnterviewd en door andere niet.

De personen die op deze wijze werden geïnterviewd, vormden geen representatieve steekproef uit de Alphenese bevolking. De conclusie kan hooguit luiden dat ze representatief waren voor het winkelende publiek op zaterdagmiddag. Daarmee werden groepen genegeerd, zoals bijvoorbeeld mensen die, om wat voor reden dan ook, niet op zaterdag winkelen.

Er zijn andere voorbeelden van peilingen die door deze wijze van gegevensverzameling tot verkeerde conclusies leidden. Een aansprekend voorbeeld is de lokale omroep die ooit op zaterdagmiddag een luisteronderzoek deed in het plaatselijke winkelcentrum en tot de niet zo verrassende, maar wel onjuiste, conclusie kwam dat vrijwel niemand luisterde naar het sportprogramma op zaterdagmiddag.

De tweede manier van peilen door de acht politieke partijen in Alphen aan den Rijn was het inschakelen van het *AlphenPanel*. Dit burgerpanel bestaat uit een steekproef van inwoners van de gemeente Alphen aan den Rijn die toegezegd hebben om op gezette tijden een vragenlijst in te vullen. De deugdelijkheid van deze peilingen staat of valt met de representativiteit van het panel. Het panel was voor het grootste deel gevuld met behulp van een aselechte steekproef uit het bevolkingsregister. Het bestond aanvankelijk uit 450 personen die zichzelf hadden

aangemeld. Het is echter later uitgebreid tot 1600 personen die via loting uit het bevolkingsregister zijn verkregen. Jongeren zijn enigszins ondervertegenwoordigd en ouderen zijn wat oververtegenwoordigd. En er zitten iets meer mannen dan vrouwen in. Al met al is dit panel redelijk representatief en kan het dus worden gebruikt om te generaliseren naar de gehele Alphense bevolking.

De derde manier van peilen was een enquête via het internet. Daarvoor werd geen aselechte steekproef getrokken. Iedereen die dat wilde, kon de vragenlijst invullen. Omdat iedereen voor zichzelf kon besluiten om wel of niet mee te doen, is dit een geval van *zelfselectie*. Door de politici werd dit de *open enquête* genoemd. De vragenlijst werd aangeboden via een link op de website van de gemeente Alphen aan den Rijn. Ook werd de link vermeld in berichtgeving in de media.

Zoals al eerder aangegeven, heeft een online peiling met zelfselectie grote nadelen. Een eerste nadeel is dat iedereen hem kan invullen, dus ook personen van buiten de doelgroep. Zo konden bij de Alphense peiling ook inwoners uit andere gemeenten de vragenlijst invullen. Uit verhalen blijkt dat dit bij deze peiling ook is gebeurd.

Een tweede nadeel is dat deelnemers aan zelfselectie-peilingen geen goede afspiegeling zijn van de bevolking. Het zijn vooral mensen die het leuk vinden aan dit soort dingen mee te doen of zich betrokken voelen bij het onderwerp (koopzondagen). Bij de open enquête bleken vooral de stemmers op de kleine christelijke partijen oververtegenwoordigd te zijn (zie paragraaf 2.3).

Een derde nadeel is dat het mogelijk is de uitslag te manipuleren. Het was bij de open enquête mogelijk om de vragenlijst meer dan één keer in te vullen. Niets stond iemand in de weg om in korte tijd de vragenlijst een keer of tien in te vullen. Verhalen wijzen erop dat dit hier ook is gebeurd.

Bij de Alphense koopzondag-enquête waren er oproepen aan leden van specifieke groepen om toch vooral de vragenlijst in te vullen. Zo schreef de kerkenraad van de Gereformeerde Kerk in Boskoop (een van de dorpen in de gemeente) op haar website:

‘Graag roepen wij u op aan deze enquête mee te doen en uw mening kenbaar te maken. De tendens in de samenleving is dat de zondag steeds meer gezien wordt als een gewone dag in plaats van een geheiligde dag, een dag van afzondering en rust. Laten we daarom met elkaar onze verantwoordelijkheid nemen, onder biddend opzien tot de Heere die alles regeert en bestuurt.’

De Hervormde gemeente in Benthuizen (ook een dorp in de gemeente) riep op 3 februari 2015 haar leden op om te gaan stemmen:

‘Op de website van de Gemeente Alphen aan den Rijn kunt u en kan jij een enquête invullen over het onderwerp koopzondagen. Het moderamen beveelt invulling van deze enquête bij u en jou van harte aan.’

En dan was er ook nog de oproep van de gemeente Alphen aan den Rijn op haar website om de enquête meer dan één keer in te vullen:

‘Ook u bent van harte uitgenodigd om aan deze enquête deel te nemen, ook al heeft u ook al eerder bij een van de winkelcentra in de gemeente meegedaan.’

Dit soort oproepen draagt zeker niet bij tot de representativiteit van de steekproef. Integendeel, de representativiteit wordt er juist door aangetast.

De drie manieren van peilen overziend komt de enquête via het panel nog het dichtst in de buurt van een representatieve steekproef. Daarom kunnen deze gegevens het best worden gebruikt om een indruk te krijgen van de mening van de Alpenaren. Het is verstandig om de twee andere aanpakken te negeren.

Er is een hardnekkig misverstand bij veel mensen dat problemen met peilingen vanzelf verdwijnen als de steekproef maar groot genoeg is. Helaas is dit niet zo. Vertekeningen ten gevolge van mankementen in de peiling blijven bestaan, hoeveel gegevens er ook worden verzameld. Daarom is het niet verstandig om de gegevens van de winkelcentrum-enquête en de open enquête te voegen bij die van de panel-enquête. Dit zou een redelijk goed gegevensbestand alleen maar vervuilen. De acht politieke partijen namen uiteindelijk ook het verstandige besluit de panel-enquête leidend te laten zijn bij de politieke besluitvorming.

2.3 De uitkomsten

Het AlphenPanel is ooit opgezet op initiatief van de gemeente Alphen aan den Rijn. Het praktische beheer van het panel en ook het uitvoeren van de enquêtes was in handen van I&O Research. Ook de open enquête werd uitgevoerd door I&O Research. Het veldwerk voor de enquête in de winkelcentra werd georganiseerd en uitgevoerd door de politieke partijen zelf.

Op 3 maart 2015 maakten de acht politieke partijen de uitkomsten van de drie enquêtes bekend. Voor de enquête in de winkelcentra werden 754 vragenlijsten ingevuld. De open enquête leverde 1550 ingevulde formulieren op. Voor de panel-enquête werden 1600 leden van het panel uitgenodigd om de vragenlijst in te vullen. Uiteindelijk deden 857 leden dat. Dit komt neer op een respons van 54%. Dit lijkt een redelijk responspercentage. De respons had hoger kunnen zijn omdat de leden van het panel ermee hebben ingestemd om af en toe een vragenlijst in te vullen. Bovendien was het een aansprekend onderwerp.

De gemeente Alphen aan den Rijn heeft 107.000 inwoners. Zo’n 66% woont in de stedelijke kern Alphen aan den Rijn. De overige 34% woont in de dorpen daaromheen. Bij de panel-enquête en de open enquête is geregistreerd waar de respondenten woonden. Daardoor kon de verdeling van de respons over de kernen in deze enquêtes worden vergeleken met de verdeling in de populatie. Zie hiervoor tabel 1.

Tabel 1 *Verdeling van de respons over de kernen*

	Panel	Open enquête	Populatie
Aarlanderveen	2%	1%	1%
Alphen aan den Rijn	70%	55%	66%
Benthuizen	3%	13%	3%
Boskoop	13%	18%	14%
Hazerswoude-Dorp	4%	8%	5%
Hazerswoude-Rijndijk	4%	2%	5%
Koudekerk aan den Rijn	3%	2%	4%
Zwammerdam	1%	1%	2%
Totaal	100%	100%	100%

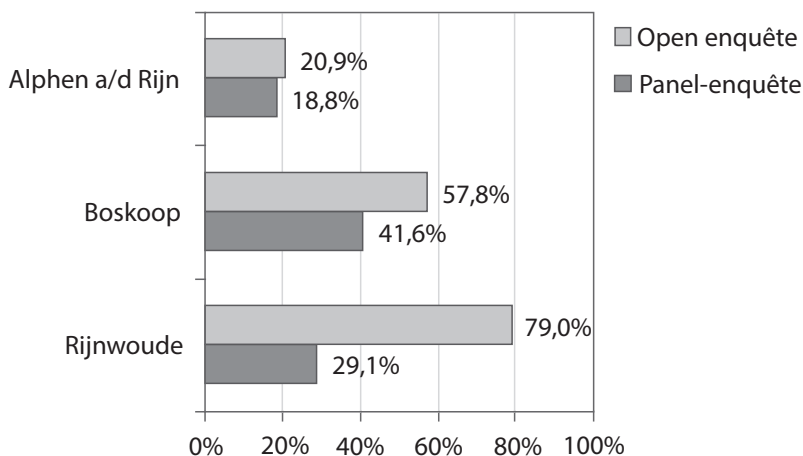
De responsverdeling van de panel-enquête lijkt veel op de verdeling in de populatie. Er zitten wat meer respondenten uit de stedelijke kern Alphen aan den Rijn in de enquête dan verwacht zou kunnen worden op grond van de verdeling van de bevolking over de kernen: 70% in plaats van 66%. Voor de overige kernen bedraagt het verschil hooguit 1%. De conclusie kan zijn dat de panel-enquête in hoge mate representatief is met betrekking tot de verdeling over de kernen.

Bij de open enquête is de representativiteit behoorlijk aangetast. Er zijn flinke verschillen tussen de verdeling in de respons en de verdeling in de populatie. Zo woont slechts 3% van de populatie in Benthuizen, maar in de enquête is 13% afkomstig uit dit dorp. De Benthuizenaren zijn duidelijk oververtegenwoordigd in de open enquête. Hetzelfde, zij het in wat mindere mate, is te zien bij Boskoop (18% in plaats van 14%) en Hazerswoude-Dorp (8% in plaats van 5%). De oververtegenwoordiging van bepaalde dorpen moet ertoe leiden dat een of meer andere dorpen of kernen ondervertegenwoordigd zijn. Dat is in de open enquête vooral de stedelijke kern Alphen aan den Rijn. In de enquête is slechts 55% afkomstig uit Alphen aan den Rijn, terwijl het 66% had moeten zijn.

Waarom zijn de dorpen Benthuizen, Hazerswoude-Dorp en Boskoop oververtegenwoordigd in de open enquête? Een aannemelijke verklaring is dat in deze dorpen veel aanhangers van de christelijke partijen wonen. Benthuizen is een dorp waarin de SGP en ChristenUnie altijd goed scoren. Zo kreeg de combinatie SGP/ChristenUnie bij de gemeenteraadsverkiezingen van 2010 zelfs 46% van de stemmen. Als dan een kerk in Benthuizen haar leden oproept om toch vooral mee te doen aan de open enquête, dan is die oververtegenwoordiging verklaarbaar. Ook in Boskoop en Hazerswoude-Dorp trekken de christelijke partijen (CDA, SGP, ChristenUnie) traditioneel veel kiezers. En ook in Boskoop deed een kerk een oproep om de vragenlijst van de open enquête in te vullen.

Om meer inzicht te krijgen in het gebrek aan representativiteit bij de open enquête is het percentage tegenstanders van koopzondagen berekend in de verschillende delen van de gemeente. Die percentages zijn vergeleken met die van de panel-enquête. De resultaten staan in figuur 1.

Figuur 1 *Het percentage tegenstanders van koopzondagen in de verschillende delen van de gemeente*



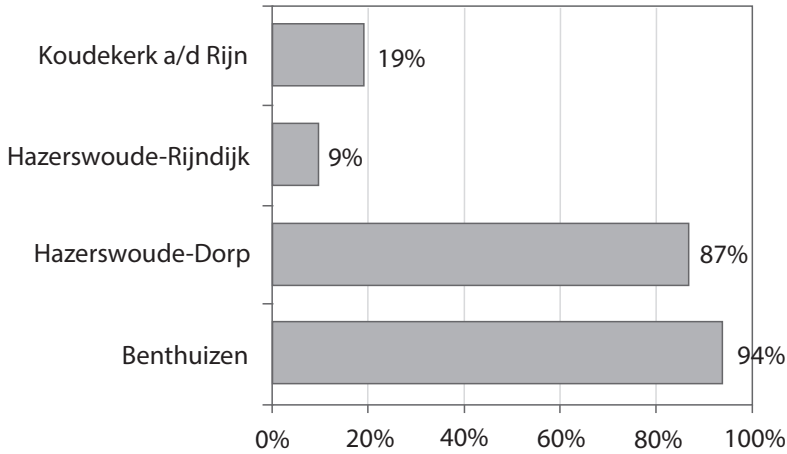
De drie delen (Alphen aan den Rijn, Boskoop en Rijnwoude) corresponderen met de vroegere gemeenten die op 1 januari 2014 zijn samengevoegd tot de nieuwe gemeente Alphen aan den Rijn. Het 'oude' Alphen aan den Rijn bestond uit de stedelijke kern Alphen aan den Rijn en de dorpen Aarlanderveen en Zwammerdam. Rijnwoude bestond uit de vier dorpen Benthuisen, Hazerswoude-Dorp, Hazerswoude-Rijndijk en Koudekerk aan den Rijn. Boskoop bestond uit het dorp Boskoop. Het percentage tegenstanders is in Rijnwoude slechts 29,1% in de panel-enquête, terwijl het in de open enquête heel hoog is (79,0%). De conclusie moet luiden dat de tegenstanders wel heel erg oververtegenwoordigd zijn in de open enquête. Er zijn ook verschillen tussen de panel-enquête en de open enquête in Boskoop. In de open enquête is 57,8% tegen koopzondagen en in de panel-enquête is dit maar 41,6%.

In figuur 2 worden de uitkomsten van de open enquête in Rijnwoude nog wat nader bekeken door ze uit te splitsen naar dorp. Dan blijkt dat bijna iedereen in Benthuisen (94%) tegen koopzondagen is. Het percentage tegenstanders is ook erg hoog in Hazerswoude-Dorp (87%). Er is een groot contrast met de andere twee dorpen in Rijnwoude. Daar is het percentage tegenstanders juist erg laag: 9% in Hazerswoude-Rijndijk en 19% in Koudekerk aan den Rijn. De tegenstanders zitten dus vooral in Benthuisen en Hazerswoude-Dorp.

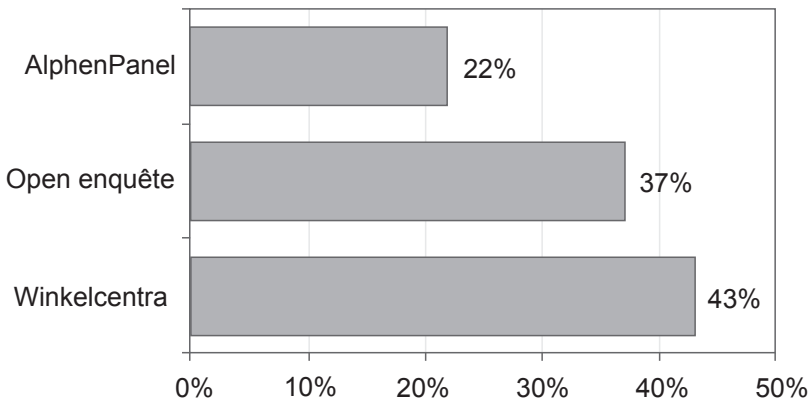
Deze analyse van de representativiteit laat duidelijk de risico's zien van het werken met een peiling die is gebaseerd op zelfselectie. Een dergelijke peiling is niet representatief en kan daardoor makkelijk leiden tot verkeerde conclusies. Daarom moeten zelfselectie-peilingen in de praktijk worden vermeden.

Er is gepoogd de respons van de panel-enquête en de open enquête te corrigeren voor een eventueel gebrek aan representativiteit. Er is een weging uitgevoerd

Figuur 2 *Het percentage tegenstanders van koopzondagen in de dorpen van de voormalige gemeente Rijnwoude*



Figuur 3 *De percentages tegenstanders van de koopzondag*



waarbij de respons representatief is gemaakt met betrekking tot de verdeling over drie delen (de voormalige gemeenten Alphen aan den Rijn, Boskoop en Rijnwoude) aangaande leeftijd en geslacht. Merk op dat dit niet wil zeggen dat de respons daarmee ook automatisch representatief is met betrekking tot andere kenmerken in de enquêtes.

Doel van het onderzoek van de acht politieke partijen was het krijgen van meer inzicht in de mening van de inwoners van de gemeente over koopzondagen. Dus werd de vraag gesteld of men voor of tegen koopzondagen was. Figuur 3 toont de uitkomsten voor de drie manieren van peilen. Merk op dat de uitkomsten van de panel-enquête en de open enquête zijn gewogen om te corrigeren voor een mogelijk gebrek aan representativiteit.

Het percentage tegenstanders varieert nogal van enquête tot enquête. Ze variëren tussen de 22% en de 43%. De uitkomst van de panel-enquête zal het dichtst liggen bij het werkelijke percentage in de populatie. Daarom lijkt 22% een redelijke goede schatting van het aantal tegenstanders. Er moet wel rekening mee worden gehouden dat schattingen op basis van een steekproef altijd een onzekerheidsmarge hebben. Voor die 22% tegenstanders is die marge gelijk aan bijna 3%. Dit betekent dat het exacte percentage tegenstanders in de gehele bevolking met grote waarschijnlijk zal liggen tussen de 19% en 25%.

Bij de open enquête is de schatting voor het percentage tegenstanders (na weging) aanzienlijk hoger: maar liefst 37%. Dat is dus 15 procentpunten meer dan de 22% van de panel-peiling. Hiermee zit de open enquête er behoorlijk naast. Een mogelijke verklaring van deze afwijkende schatting is een oververtegenwoordiging van personen die vanuit hun geloof tegen de koopzondag zijn. Weliswaar wordt in de weging de oververtegenwoordiging van de inwoners van Rijnwoude gecorrigeerd, maar dan kunnen nog steeds stemmers op de christelijke partijen in Benthuizen en Hazerswoude-Dorp oververtegenwoordigd zijn. Wegen naar deel van de gemeente heft deze oververtegenwoordiging kennelijk niet op.

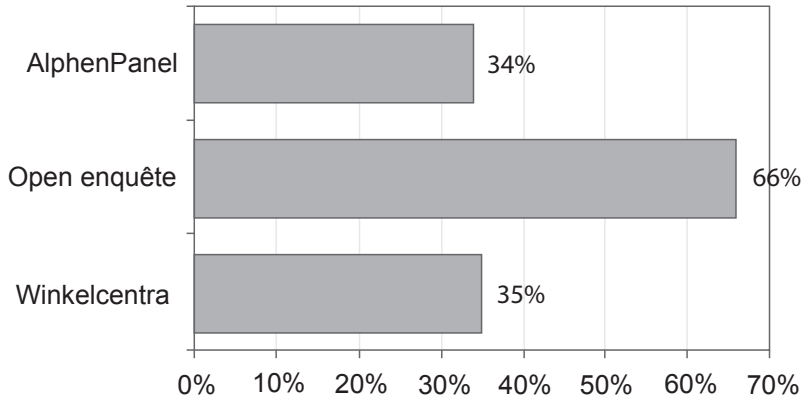
Bij de enquête in de winkelcentra is de schatting van het aantal tegenstanders van de koopzondag nog groter: 43%. Dat is bijna dubbel zoveel als bij de panel-enquête. Een verklaring hiervoor valt lastig te geven zonder verder onderzoek. Misschien is het wel zo dat juist mensen die op zaterdag winkelen geen behoefte hebben aan winkelen op zondag. En mensen die graag op zondag willen winkelen omdat ze op zaterdag niet kunnen, zitten niet in de enquête.

De verschillen tussen de drie peilingen zijn te groot om aan toevallige afwijkingen in de steekproef te kunnen toeschrijven. Er is hier sprake van grote systematische verschillen. De conclusie moet zijn dat de uitkomsten van de open enquête en de enquête in de winkelcentra fout zijn. Het lijkt erop dat de enquête in de winkelcentra de grootste afwijkingen had. Dat die enquête het zo slecht deed, ligt niet aan het feit dat het een mondelinge (face-to-face) enquête was. Over het algemeen zijn mondelinge enquêtes kwalitatief juist goed. De reden is dat er enquêteurs worden ingezet. Die kunnen mensen overtuigen om mee te werken en respondenten helpen bij het beantwoorden van de vragen. Dat levert betere gegevens op. Dat de enquête in de winkelcentra het zo slecht deed, had vooral te maken met het selectiemechanisme van de steekproef. Dat leverde een steekproef op die verre van representatief was.

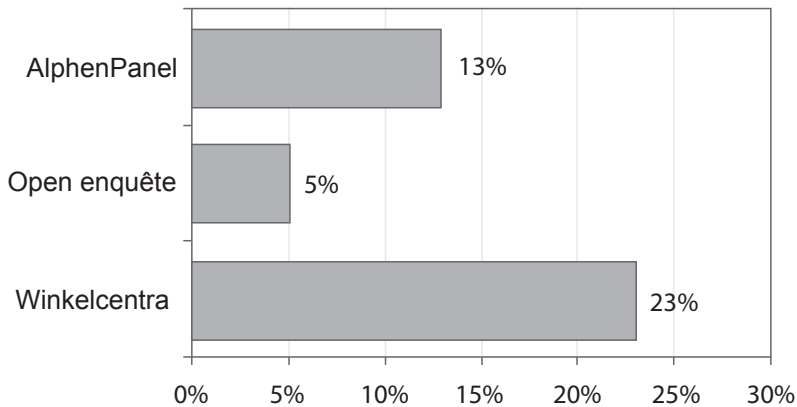
Er zijn in het onderzoek nog diverse andere vragen gesteld. En ook voor die vragen zijn er grote verschillen te zien tussen de verschillende enquêtes. Hieronder volgen nog twee voorbeelden.

Aan de tegenstanders van de koopzondag is gevraagd waarom ze tegen koopzondagen zijn. Een voor de hand liggende reden is dat iemand vanuit zijn of haar geloof tegen koopzondagen is. In figuur 4 staan de percentages tegenstanders die vanwege hun geloof tegen koopzondagen zijn.

Figuur 4 De percentages tegenstanders vanuit het geloof



Figuur 5 Koopzondagen moeten alleen gelden voor supermarkten, bouwmarkten en tuinen



De schatting van de panel-enquête is waarschijnlijk de beste van de drie. Dat zou betekenen dat ongeveer een derde van de tegenstanders (34%) tegen koopzondagen is vanwege het geloof. In de open enquête is dat percentage bijna twee keer zo groot: 66%. Dit is weer een aanwijzing dat mensen met een christelijke achtergrond oververtegenwoordigd zijn in deze enquête. Het percentage in de winkelcentra (35%) ligt (toevalligerwijs?) dicht in de buurt van dat van de panel-enquête.

Ten slotte staan in figuur 5 de uitkomsten van de vraag of de koopzondag alleen voor supermarkten, bouwmarkten en tuincentra moet gelden of dat die voor alle winkels moet gelden. Deze vraag is alleen gesteld aan de voorstanders van de koopzondag.

De percentages schieten alle kanten op. De beste schatting is waarschijnlijk de 13% van de panel-enquête. Het percentage van 5% in de open enquête is veel te

laag. En het percentage van 23% in de winkelcentra-enquête is veel te hoog. Het is lastig om zonder verder onderzoek deze afwijkingen te verklaren.

3 Conclusie

Het onderzoek naar de koopzondagen in Alphen aan den Rijn bood de gelegenheid om drie verschillende manieren van peilen met elkaar te vergelijken. De verschillen bleken behoorlijk groot te zijn.

Als een gemeente een burgerpanel heeft, dan kan dit heel goed voor dit soort peilingen worden gebruikt. Een goed panel voldoet aan twee voorwaarden:

- Het burgerpanel moet zijn gevuld door het trekken van een steekproef van inwoners die is geloot uit de hele bevolking.
- Elke inwoner van de gemeente moet dezelfde kans hebben gehad om in de steekproef te komen.

De geselecteerde personen worden uitgenodigd lid te worden van het panel. Zo ontstaat een panel dat representatief is voor alle inwoners van de gemeente. Het vormt een goede afspiegeling van de bevolking. Het ligt voor de hand om de leden van het panel te loten uit de bevolkingsadministratie.

Er moet voor worden gewaakt dat niet te veel mensen weigeren, aangezien dit de representativiteit kan aantasten. En ook bij de peilingen vanuit het panel moet de respons bij voorkeur hoog zijn. Het is verstandig om een wegging uit te voeren om een eventuele verstoring van de representativiteit te corrigeren.

Een online peiling op basis van zelfselectie is een slecht meetinstrument. De representativiteit kan op allerlei manieren zijn aangetast. Zo kunnen ook personen van buiten de doelgroep de vragenlijst invullen, en daarmee de uitkomsten 'vervuilen'. Verder komt het vaak voor dat een vragenlijst meer dan één keer kan worden ingevuld. Ook is het mogelijk de uitkomsten te manipuleren door aandrang op bepaalde groepen om toch maar de vragenlijst in te vullen.

Ook een peiling op straat is een slechte peiling. Immers, alleen mensen die op dat moment op straat zijn, kunnen worden geïnterviewd. Iedereen die, om wat voor reden dan ook, niet op straat is, blijft buiten het onderzoek. Daarmee wordt de doelgroep van het onderzoek in feite ingeperkt tot de mensen op straat op dat tijdstip. En die groep hoeft niet representatief te zijn voor alle inwoners.

Het zal duidelijk zijn dat het belangrijk is om peilingen op een zorgvuldige wijze uit te voeren. Een verkeerde aanpak kan makkelijk tot afwijkende uitkomsten leiden. Op de website www.peilingpraktijken.nl zijn daarvan nog talloze andere voorbeelden te vinden.

Daar waar het gaat om het beoordelen van de kwaliteit van een door derden uitgevoerde peiling, is het nuttig om gebruik te maken van de Checklist voor Peilingen (Bethlehem, 2012). Deze checklist helpt om het kaf van het koren te scheiden.

Het was een initiatief van het Centraal Bureau voor de Statistiek (CBS), de Vereniging voor Onderzoeksjournalisten (VVOJ) en het Nederlandstalig Platform voor Survey-Onderzoek (NPSO).

Voor wat betreft de koopzondagen was het een lastige situatie in Alphen aan den Rijn: de stedelijke kern was voor en de dorpen eromheen waren tegen. Welke besluit ook werd genomen, het zou altijd teleurstellend zijn voor een deel van de inwoners. Uiteindelijk besloot de gemeenteraad in maart 2015 met een grote meerderheid (28 tegen 10 stemmen) om koopzondagen ongelimiteerd toe te laten. Wel werd een motie van het CDA en de lokale partij Nieuw Elan aanvaard, waarin vastgelegd wordt dat de winkels op zondag pas om 13.00 uur open mogen.

Er zijn ook andere benaderingen mogelijk bij de besluitvorming. Een voorbeeld is de gemeente Ede. Op 10 juni 2015 werd een referendum gehouden, waarin de inwoners zich konden uitspreken voor of tegen koopzondagen. Dit referendum kan worden opgevat als een peiling waarbij alle inwoners in de steekproef zaten. De uitslag was dat een krappe meerderheid van 57% tegen koopzondagen was. Dat weerhield de Edese gemeenteraad er niet van toch akkoord te gaan met ongelimiteerde koopzondagen in de stedelijke kern Ede. In de dorpen rondom Ede blijven de winkels gesloten op zondag.

De ervaringen met de peilingen in de gemeente Alphen aan den Rijn leiden tot enkele conclusies. In de eerste plaats is het goed mogelijk om een peiling te gebruiken als instrument bij het vormen van beleid. Het kan gaan om het meten van sociaaleconomische stand van zaken in de gemeente of ook het peilen van de mening van de burgers over allerlei lokale zaken. Als in een gemeente een meer bedrijfsmatige wijze van besturen wordt gehanteerd (New Public Management), dan kan een peiling ook worden gebruikt voor het meten van de tevredenheid bij de klanten (burgers en bedrijven).

Een tweede conclusie is dat de politici in Alphen aan den Rijn besloten tot het houden van een peiling zonder zich te realiseren wat er komt kijken bij een goede peiling. Sommigen waren zich zelfs niet bewust van het bestaan van een burgerpanel in de gemeente. Daarom kwamen ze in eerste instantie niet op het idee om de peiling uit te voeren met dit panel, maar op straat te gaan enquêteren. Het lijkt daarom zinvol de kennis van de gemeenten op dit terrein te verbeteren.

Literatuur

- Bethlehem, J.G., *The rise of survey sampling*, Discussion Paper 09015, Den Haag/Heerlen: 2009a.
- Bethlehem, J.G., *Applied survey methods. A statistical perspective*, Hoboken, NJ: 2009b.
- Bethlehem, J.G., *Peilingen beoordelen. Een checklist*, www.peilingpraktijken.nl/publicaties: 2012.
- Bethlehem, J.G., *Deugdelijke peilingen*, Discussion Paper 2013 | 10, Den Haag/Heerlen: 2013.

- Bethlehem, J.G., *Slechte peiling, gemanipuleerde uitslag*, www.peilingpraktijken.nl: 13 januari 2014.
- Bethlehem, J.G. & S. Biffignandi, *Handbook of web surveys*. Hoboken, NJ: 2012.
- Bowley, A.L., 'Address to the Economic Science and Statistics Section of the British Association for the Advancement of Science', *Journal of the Royal Statistical Society* 69, 1906, p. 548-557.
- Bowley, A.L., 'Measurement of the precision attained in sampling', *Bulletin of the International Statistical Institute, XII, Book 1*, 1926, p. 6-62.
- Bronzwaer, S., 'Infiltranten probeerden de peilingen van Maurice de Hond te manipuleren', *NRC Handelsblad*, 13 september 2012.
- EenVandaag, *Opiniepanel: achter de schermen bij een flitspeiling!*, http://opiniepanel.eenvandaag.nl/uitslagen/60679/opiniepanel_achter_de_schermen_bij_een_flitspeiling?autoplay=1.
- Eurostat, *Level of internet access – households*, Luxemburg: 2015. Retrieved from <http://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tin00134>.
- Horvitz, D.G. & D.J. Thompson, 'A generalization of sampling without replacement from a finite universe', *Journal of the American Statistical Association* 47, 1952, p. 663-685.
- Kiaer, A.N., 'Observations et expériences concernant des dénombrements représentatives', *Bulletin of the International Statistical Institute, IX, Book 2*, 1895, p. 176-183.
- Kruskal, W. & F. Mosteller, 'Representative sampling I: Non-scientific literature', *International Statistical Review* 47, 1979a, p. 13-24.
- Kruskal, W. & F. Mosteller, 'Representative sampling II: Scientific literature, excluding statistics', *International Statistical Review* 47, 1979b, p. 111-127.
- Kruskal, W. & F. Mosteller, 'Representative sampling III: The current statistical literature', *International Statistical Review* 47, 1979c, p. 245-265.
- Neyman, J., 'On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection', *Journal of the Royal Statistical Society* 97, 1934, p. 558-606.
- Osborne, D. & T. Gaebler, *Reinventing government. How the entrepreneurial spirit is transforming the public sector*, New York: 1992.
- Slot, J., 'Vragen? Geen vragen!', *Ons Amsterdam*, 2009.