

8. Eurobar? Onderzoekstechnische haken en ogen van de Eurobarometer

JELKE BETHLEHEM en JOOP VAN HOLSTEYN

SAMENVATTING

De Eurobarometer is een instituut. Gegevens van dit grootschalige onderzoek in alle lidstaten van de Europese Unie werken door in wetenschap, beleid en publieke opinie. Bij de datakwaliteit kunnen echter methodologische kanttekeningen worden geplaatst. Zo is onduidelijk hoe de steekproef getrokken is en wat de omvang en aard van de non-respons zijn, terwijl de vragenlijst suggestieve, op een pro-Europese houding sturende vragen kent. Veranderingen in de vragenlijst tasten daarenboven het longitudinale karakter van de Eurobarometer aan. Niet al deze methodologische problemen zijn eenvoudig oplosbaar, maar in ieder geval zou transparant moeten zijn welke onderzoekstechnische keuzen zijn gemaakt, om de beoordeling van de datakwaliteit te faciliteren..

Trefwoorden: Eurobarometer, datakwaliteit, steekproeftrekking, vragenlijst, longitudinaal onderzoek, onderzoeksverantwoording

1. INLEIDING¹

Het is een bekend gezegde: je hebt leugens, grote leugens en statistieken. In Brussel heten de statistieken eurobarometers: opiniepeilingen die worden uitgevoerd in opdracht van de Europese Commissie. (*NRC Handelsblad*, 28 januari 2008)

De Eurobarometer (EB) is een instituut. Al sinds begin jaren zeventig van de vorige eeuw worden in opdracht van de Europese Commissie grootschalige survey-onderzoeken of peilingen uitgevoerd in de lidstaten van wat nu de Europese Unie (EU) heet (voor een brede schets van achtergrond en aard van de Eurobarometer, zie Signorelli, 2012). In 2016 is al weer voor de 85ste maal een zogenoemde standaardpeiling gehouden; met in de regel twee dergelijke onderzoeken per kalenderjaar loopt de teller hard op.² Primair (vermeend) doel van de peilingen is het - vanuit longitudinaal perspectief dan wel cross-sectioneel - meten van meningen en houdingen van burgers van de EU-lidstaten over die EU en het proces van verdieping en verbreding van Europese integratie. In haar openingswoord bij een feestelijke uitgave naar aanleiding van het 35-jarig bestaan van de EB verwoordde eurocommissaris Margot Wallström een en ander als volgt: “Since 1973, the European Commission has been monitoring the evolution of public opinion in the Member States, thus helping the preparation of texts, decision-making and the evaluation of its work. (...) Over more than three decades, Eurobarometer polls have given us a solid indication of the concerns, needs and opinions of European citizens”. In zijn groet aan de lezer in diezelf-

A.E.Bronner et al. (red.), Ontwikkelingen in het marktonderzoek: Jaarboek MarktOnderzoek-Associatie, dl. 42, 2017. Haarlem: Spaar^{en}Hout.

de uitgave meldde het hoofd van het Directoraat-Generaal voor Communicatie van de Europese Commissie, Claus H. Sørensen, even bondig als onbescheiden over de EB: “A unique source of knowledge and a remarkable tool for policy advice” (beide citaten in European Commission, 2008, z.p.).

Overdrijven deed Sørensen nauwelijks. Want inderdaad, de in der loop der tijd verzamelde EB-data leveren stof voor publiek debat, vormen *input* voor bestuur en beleid, en hebben de empirische basis gelegd voor tal van wetenschappelijke publicaties en daarmee inzicht in met name de perceptie en evaluatie van de EU op het niveau van de burger (vgl. Schmitt, 2003). Wat dat laatste betreft, in een toonaangevend internationaal tijdschrift als *European Union Politics*, bijvoorbeeld, verschenen in de eerste tien jaargangen (2000-2009) in 38 afleveringen in totaal 168 oorspronkelijke wetenschappelijke bijdragen, waarvan er 28 (17 procent) geheel of grotendeels gebaseerd waren op EB-gegevens. In hun diagnose van het gebrek aan “scientific maturity in the key area of data accumulation and integration” op het terrein van vergelijkend EU-onderzoek, wijzen Gabel en de zijnen dan ook op de “value and prominence of the Eurobarometer” als uitzondering en lichtend voorbeeld (Gabel et al., 2002, p. 482 resp. p. 485). Dat voor wetenschappelijk empirisch onderzoek op massaniveau *by default* naar de mogelijkheden van de EB wordt gekeken, mag tevens blijken uit het feit dat een afwijkende aanpak lijkt te moeten worden verantwoord in het licht van die EB-mogelijkheden: “To test our hypotheses while adequately controlling for other influences, we cannot rely on the most readily available data (the Eurobarometer) because a number of the key concepts are either not included in the barometer or poorly operationalized. (...) We therefore collected new data (...)” (de Vreese et al., 2008, p.516). Het veelvuldige, welhaast vanzelfsprekende gebruik van EB-data leidt er soms zelfs toe, dat een eigenstandige verantwoording van data en dataverzameling nauwelijks nog nodig wordt geacht: “Like other students of EU public opinion, we draw our data from the Eurobarometer series”(Nelsen et al., 2001, p. 195). De EB heeft, kortom, wellicht slechts de Amerikaanse National Election Studies als serieuze concurrent als het gaat om in wetenschappelijke publicaties meest frequent gebruikte of geraadpleegde databestand (Schmitt, 2003, p.245).

Naar het woord van Sørensen is de EB echter niet enkel bron van (wetenschappelijke) kennis, maar ook anderszins van belang, tot en met het landelijke bestuur en de nationale beleidsvorming. Hij heeft gelijk, ook voor Nederland. In de vaste Kamercommissie voor Economische Zaken wenst de VVD-fractie in mei 2016 van de staatssecretaris te weten of de bewindspersoon “meer gedetailleerd inzicht [kan] geven in de Eurobarometer over dierenwelzijn”, terwijl de SP-fractie in dat overleg eveneens ten aanzien van dierenwelzijn wil weten of “de uitkomsten van de Eurobarometer tot nieuwe inzichten [hebben] geleid betreffende de richting van Europees beleid en de Nederlandse inzet daartoe” (Kamerstuk 21 501-32, nr. 925). In ander verband onderbouwt Kamerlid Pia Dijkstra (D66) in 2015 haar initiatiefwetsvoorstel tot wijziging van de Wet op de orgaandonatie onder meer met een verwijzing naar een EB uit 2009: “Daaruit komt naar voren dat 64% van de Nederlandse ingezetenen desgevraagd bereid is om na overlijden organen te doneren”(Kamerstuk 33 506, nr. 8). In weer een ander ‘dossier’ schrijft de Minister van Volksgezondheid, Welzijn en Sport aan de Tweede Kamer een brief met een geannoteerde agenda voor de Gezondheids-

raad, waarin wordt opgemerkt dat uit “de Eurobarometer valt (...) op te maken dat Nederlandse burgers gemiddeld beter op de hoogte zijn van hun rechten bij grensoverschrijdende zorg dan het EU-gemiddelde” (Kamerstuk 21 501-31, nr. 382). Dit zijn slechts enkele voorbeelden; het aantal verwijzingen naar de EB in de sfeer van politiek en beleidsvorming is aanzienlijk.

Ten slotte is er, als het gaat om reikwijdte en impact van de EB, de doorwerking naar de media en daarmee de publieke opinie en het publieke debat, als weer een peiling is afgerond. ‘Nederlanders achter EU-limaatschap’, kopt *de Volkskrant* op 7 september 2013 op basis van de toen meest recente EB. ‘Turkije ongewenst’, meldt *NRC Handelsblad* op 19 juli 2005 kort en bondig de uit de EB sprekende publieke opinie. En het *Reformatoisch Dagblad* kon op 19 december 2006 niet anders dan, waarschijnlijk met spijt en afkeer, melden dat de EB toont: ‘Nederlanders lopen voorop bij homohuwelijk’.

De EB is, kortom, een potentieel uiterst invloedrijk instrument van opinieonderzoek aangaande de EU.³ Dan geldt te meer: adels verplicht. Dat houdt hier in dat de EB vanuit methodologisch perspectief van goede, liefst superieure kwaliteit is. Het pretendeert immers te gaan om een ‘solid indication of the concerns, needs and opinions of European citizens’, in de woorden van Wallström. Over die kwaliteit van de EB-bevindingen bestaat echter toch wel enige zorg (zie o.a. Nissen 2014; Höpner & Jurczyk, 2012, 2015; zie voor een reactie van de Europese Commissie op dergelijke zorgen o.a. Bläser, 2013). Herhaaldelijk rijst bijvoorbeeld de vraag waarom uitkomsten van de EB veelal positiever zijn of lijken dan uitkomsten van andere, vergelijkbare peilingen: ‘De eurobarometer voorspelt nooit slecht weer’ (*de Volkskrant*, 19 januari 2005). Kan het zijn dat de Europese Commissie, als opdrachtgever en financier verantwoordelijk, belang heeft bij bepaalde resultaten en (on)bewust de uitkomsten stuurt, en misschien in de loop der tijd de EB meer is gaan gebruiken als beleidsdan als meetinstrument van publieke opinie (vgl. Nissen, 2014), zelfs als ‘strategische manipulatie’ van die opinie (Höpner & Jurczyk, 2015, p.5)? “In persberichten verwijst de Commissie graag naar eurobarometers. Die blijken - toeval of niet - nogal eens het beleid van Brussel te steunen” (*NRC Handelsblad*, 28 januari 2008). Werkt het Commissiebelang door in opzet en uitvoering van het onderzoek en zorgt dat er wellicht mede voor dat, in de woorden van columnist René Cuperus “de tone of voice van de Eurobarometer prettig eurofiel uitvalt” (*de Volkskrant*, 1 oktober 2012)? Is de Eurobarometer inderdaad (Euro)bar en boos? Of is hier geen sprake van? Meer in het algemeen: wat zegt een methodologische reflectie op de Eurobarometer over de ermee verkregen uitkomsten?⁴

2. ONDERZOEKSTECHNISCHE OPMERKINGEN BIJ DE EUROBAROMETER

Het is echter nog niet zo eenvoudig om scherp in beeld te krijgen hoe de Eurobarometer als onderzoeksinstrument en dataverzamelingsproject precies in elkaar zit. De constatering dat het om een *black box* gaat (Marcus, 2009, p.159), heeft niet exclusief op de Duitse situatie betrekking. De onderzoekstechnische specificaties van het veldwerk, zoals bijvoorbeeld opgenomen in een ook in omvang beknopte bijlage van

het verslag met eerste resultaten van de standaard EB van najaar 2015, geven enige maar geen volledige, afdoende informatie (European Union, 2015). Gemeld wordt in welke periode het veldwerk door welke organisatie in wiens opdracht is uitgevoerd; het aantal gerealiseerde vraaggerekken per land; het basisidee van de steekproef-trekking ('a multi-stage, random (probability) one'; alzo althans sinds 1989, zie Nissen, 2014, p.717) en de globale uitvoering daarvan tot en met de selectie van de individuele respondent; dat gewerkt is met *computer-assisted personal interviewing* (CAPI), althans "in those countries where this technique was available"; dat correctie van de data middels "marginal and intercellular" weging heeft plaatsgevonden op geslacht, leeftijd, regio en 'size of locality'; en dat het bij de resultaten om schattingen binnen een betrouwbaarheidsinterval gaat. Daarmee houdt het zo ongeveer op en aldus is het methodologische glas half leeg, niet half vol.

Het voor het EB-onderzoek verantwoordelijke DG Communicatie van de EC (b)lijkt vervolgens niet bijster toeschietelijk in het verstrekken van gedetailleerde methodologische informatie.⁵ Extra informatie over daadwerkelijk getrokken steekproeven en de uitvoering in onderscheiden landen, over non-responspercentages en aard van non-respons,⁶ over methoden van dataverzameling waar CAPI kennelijk niet beschikbaar was, en over eventuele correcties en gewichten is niet beschikbaar, ook niet bij (herhaalde) navraag (door de auteurs van deze bijdrage). De methodologische transparantie⁷ laat daarmee te wensen over, terwijl allicht de eerste aan goed onderzoek te stellen eis die van transparantie is: alle stappen van onderzoeksvraag naar bevindingen dienen navolgbaar en controleerbaar te zijn. En de EB mag dan een indrukwekkend data-instituut zijn, de "continuously growing body of data is not an indicator of data quality per se"(Nissen, 2014, p.716). Vandaar onze reflectie⁸ op enkele onderzoekstechnische keuzen en mogelijke foutenbronnen als de steekproef, de non-respons en correctie ervan, en de vragenlijst en gestelde vragen, in de hoop op een scherper beeld van de onderzoekstechnische aard van het instituut EB en de op velerlei terrein relevante uitkomsten ervan. Toon en strekking in het navolgende zijn kritisch, maar het woord van de Vlaamse auteur Louis Paul Boon voor zijn literaire kritische beschouwingen is evenzeer voor onze boodschap leidraad - 'ook de afbreker bouwt op' (Boon, 1982).

2.1. De steekproef

Om uitkomsten van een op een steekproef gebaseerde peiling te kunnen generaliseren naar de populatie, in dit geval de inwoners van de EU of afzonderlijke lidstaten ervan, moet een zogenoemde representatieve steekproef worden getrokken.⁹ Dat kan worden bereikt door die steekproef, eenvoudig gezegd, te loten uit de bevolking (zie voor klassieke studies over steekproeftheorie o.a. Kish, 1965; Cochran, 1977; zie ook Bethlehem, 2009). Elke eenheid van de tevoren gedefinieerde populatie - 'iedereen' - moet een positieve kans hebben om in de steekproef te komen en selectiekansen moeten bekend zijn. Alleen dan kunnen valide schattingen worden gemaakt. Hoe gaat dat bij de Eurobarometer? Los van het feit dat niet evident is wie tot de EU-populatie behoort (vgl. Nissen, 2014, p.717), wordt uit de informatie in EB-rapporten duidelijk, dat voor de standaard en de *flash* of flits-onderzoeken uiteenlopende steekproefontwerpen worden gebruikt. De standaard EB is zoals gezegd een CAPI-peiling. Interviewers (van TNS) bezoeken respondenten thuis en nemen daar

met behulp van een laptop de vragenlijst af. De steekproef wordt getrokken via 'standard random route'. Nergens is te vinden wat 'standard random route' betekent, maar de steekproeftrekking vindt kennelijk plaats in drie stappen:

- 1: Er wordt een aselechte steekproef van geografische gebieden getrokken. Daarbij wordt gestratificeerd naar NUTS 2 regio's; in Nederland zijn dit provincies.
- 2: In elk geselecteerd gebied wordt een aselechte startadres bepaald als eerste te bezoeken adres. Vervolgens bezoekt de interviewer volgens een vaste procedure een reeks volgende adressen. Al lopend wordt steeds een vast aantal adressen overgeslagen en komt het volgende adres in de steekproef. De interviewer gaat zo door, totdat het voorgeschreven aantal respondenten ondervraagd is.
- 3: Op het geselecteerde adres wordt uit de aldaar wonende (aanwezige?) personen (vanaf de voorgeschreven leeftijd) één persoon geloot, te weten de persoon die het eerst jarig is.

Probleem met deze aanpak is dat er de facto sprake is van een quota-steekproef, met een gerealiseerde respons die uiteindelijk gelijk is aan de gewenste absolute respons van (voor de meeste lidstaten) ongeveer 1.000 personen per land. De interviewers gaan door met hun 'jacht op de laatste respondent' (vgl. Stoop, 2005), totdat de vooraf bepaalde steekproefomvang is bereikt. Dit doet in uitvoering denken aan substitutie, waarbij niet-responderende adressen of huishoudens worden vervangen door responderende. De interviewers oogsten zo laaghangend fruit en ondervragen respondenten die gemakkelijk te bereiken zijn – wat gebeurt als iemand bij een eerste contactpoging niet thuis is of geeft? Echter, moeilijke(re) en non-respondenten hebben veelal andere karakteristieken dan zij die de meest eenvoudige prooi vormen, wat kan leiden tot vertekening van uitkomsten. Om deze reden wordt substitutie veelal afgeraden (o.a. Kohler, 2007; Vehovar, 1999).

De standaard EB is daarenboven in feite een clustersteekproef: er worden gebieden geloot en daarbinnen personen geselecteerd. In een geloot gebied wonen te ondervragen personen echter op loopafstand en treden clustereffecten op, sterker naarmate meer respondenten per cluster worden ondervraagd. Daarom wordt het maximaliseren van het aantal te selecteren clusters aangeraden (o.a. Fink, 1995b, p.16). Als personen uit eenzelfde cluster meer op elkaar lijken dan personen uit verschillende clusters, dan is de standaardfout van schattingen namelijk groter dan bij een simpele aselechte steekproef, waardoor onzekerheidsmarges groter zijn; de uitkomsten zijn minder precies of nauwkeurig. Bij de EB ontbreekt helaas in dit verband relevante informatie over het aantal ondervraagde personen per cluster. En in publicaties met EB-bevindingen is weliswaar informatie opgenomen over het berekenen van onzekerheidsmarges, maar daarbij wordt er – ten onrechte – vanuit gegaan dat sprake is van een simpele aselechte steekproef.

De flits EB is geen CAPI-peiling maar maakt gebruik van *computer-assisted telephone interviewing*: een CATI-peiling.¹⁰ In de steekproef geselecteerde personen worden gebeld. Wie de gelukkigen zijn, blijft echter in het duister gehuld. Details over de steekproeftrekking ontbreken en de website van GESIS, het instituut dat EB-datasets beschikbaar stelt, helpt de geïnteresseerde niet veel verder: "For each Flash Eurobarometer new and independent samples are drawn by random. Representativity is stated for the respective universe. Selection details (e.g. RDD, regions

quotas, ...) for the total population (15+) and the youth surveys are not published".¹¹ Dat is volstrekt duidelijk, maar weinig transparant.

Uit gepubliceerde rapporten valt af te leiden, dat veldwerkorganisatie TNS een soort Random Digit Dialing (RDD) gebruikt (zie bijvoorbeeld het rapport bij Flash Eurobarometer 432 van januari 2016, p. TS.2). Hierbij wordt uitgegaan van nummers die TNS beschikbaar heeft uit eerdere peilingen, bijvoorbeeld contactnummers van respondenten in de standaard EB - waarvan de steekproef zoals we zagen problematisch is. Van die nummers worden de laatste twee cijfers vervangen door twee willekeurige andere cijfers. De op die manier verkregen (nieuwe) nummers worden gestratificeerd naar regio om een goede spreiding over het land te krijgen. TNS geeft aan dat er zowel vaste als mobiele nummers worden gebeld. Dat roept de vraag op hoe mobiele nummers geografisch gestratificeerd kunnen worden, aangezien deze nummers, anders dan vaste nummers, geen geografische informatie bevatten. Ook wordt niet aangegeven hoe eventueel gecorrigeerd wordt voor het feit dat burgers met een mobiel én vast nummer een dubbel zo grote selectiekans hebben.

En dan is er in Nederland ook nog, als het gaat om obstakels op weg naar het telefonisch bereiken van beoogde respondenten, het Bel-me-niet-Register, een lijst van nummers die niet gebeld mogen worden voor telefonische verkoop. Nu is een telefonisch interview in het kader van een EB iets anders dan telefonische verkoop, maar toch kun je je afvragen of het verstandig en effectief is dit soort nummers te bellen. Ook het worden benaderd door een interviewer kan tenslotte leiden tot 'irritatie over ongewenste telefoontjes' waarvan het register mensen juist wenst te vrijwaren - het lijkt erop, dat vele burgers simpelweg niet door vreemden gebeld willen worden. Echter, als wordt besloten nummers van dit register niet te gebruiken, dan kan een flink deel van de populatie niet worden gebeld: in april 2013 zouden ruim 8 miljoen nummers in het Bel-me-niet-Register zijn opgenomen.¹² De specifiek voor (telefonisch) marktonderzoek ontwikkelde variant, de MOA Onderzoekfilter, bevat waarschijnlijk een veel kleiner maar desalniettemin niet te verwaarlozen aantal telefoonnummers.¹³ Het zoeken, vinden en ondervragen van aselekt gekozen respondenten via de telefoon is, kortom, een immens probleem, zoals onder anderen Britse peilers in de aanloop van de parlementsverkiezingen van 2015 moesten ervaren. Dat sprake was van een UK Polling Disaster mag niet werkelijk verbazen, gezien de kans op vertekening in de uiteindelijke steekproef als voor het realiseren van 2.000 vraaggesprekken niet minder dan 30.000 belpogingen moeten worden ondernomen...¹⁴

2.2. Non-respons en weging

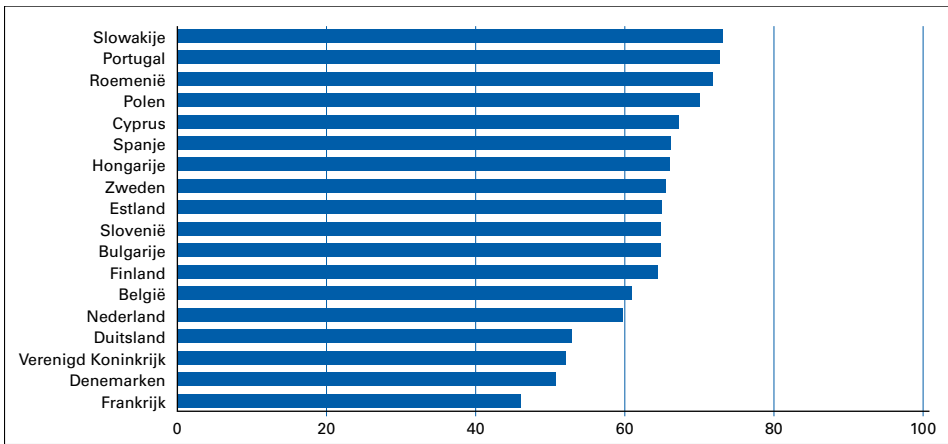
Met het moeizame trekken van een goede steekproef en bereiken van respondenten komen we bij een volgend groot en groeiend probleem: non-respons. Vooral maar niet uitsluitend telefonische peilingen, zoals bij de flits EB, kennen een problematische lage respons. In Europa is de situatie niet echt beter dan in de Verenigde Staten, waar reeds de vrees is uitgesproken dat het einde van het telefonische interview nabij is; volgens Rivers (2007) ligt bij veel Amerikaanse telefonische peilingen de respons al onder de 20 procent.

Non-respons doet zich onvermijdelijk bij via steekproeven gehouden onderzoeken of peilingen voor (zie voor een recent overzichtswerk over deze problematiek Bethlehem et al., 2011). Er zijn altijd wel geselecteerde eenheden die niet mee kunnen of willen doen aan de beoogde ondervraging. De EB wordt uiteraard eveneens geconfronteerd met dit vanuit onderzoekstechnisch oogpunt hinderlijke verschijnsel. Non-respons verkleint de effectieve steekproef en vergroot de onnauwkeurigheid van schattingen, maar een wezenlijker probleem bestaat uit de kans op vertekening, dat wil zeggen op aantasting van de validiteit of geldigheid van schattingen door de uiteenlopende mate van deelname aan het onderzoek van voor de onderzoeksvraag relevante subgroepen in de populatie. Non-respons kan leiden tot substantiële vertekeningen in bevindingen, die waarschijnlijk groter zijn naarmate het percentage non-respons groter is. Het percentage non-respons is dan ook een voorname kwaliteitsindicator, die vermeld hoort te worden in elke onderzoeksverantwoording.

In de EB-rapporten wordt (zo goed als) niets gemeld over hoogte en aard van de non-respons. Is er dan geen non-respons? Niets is minder waar. Bij de standaard en flits EB is er bijvoorbeeld non-respons doordat de interviewer iemand niet thuis treft respectievelijk geen telefonisch contact weet te leggen (geen contact), iemand weigert mee te doen (weigering), of iemand de juiste taal niet spreekt (niet in staat). Er kan in het eerste geval nog wel een poging worden gedaan de non-respons te verminderen door herhaalde contactpogingen te doen - sommige organisaties doen tot zes pogingen - maar daarvan wordt in EB-rapporten geen melding gemaakt. Hoe dan ook zal een substantiële hoeveelheid non-respons overblijven, die wat de standaard EB aan het oog onttrokken wordt door de gehanteerde pseudo-quota steekproefaanpak gecombineerd met het feit dat stevast het streefaantal van ongeveer 1.000 vraaggelassen wordt gerealiseerd. In de flits EB rapporten wordt niets gemeld over non-respons.

Bij landenvergelijkend onderzoek als de EB doet zich nog een complicerende factor voor: de hoogte van de non-respons is niet overal gelijk en eventuele gevolgen ervan kunnen per land variëren. Gegevens van de European Social Survey (ESS), die gekend is om de hardnekkige pogingen tot responsbevordering en in termen van datakwaliteit en representativiteit superieur mag worden geacht aan de Eurobarometer (vgl. Schmitt, 2003, p. 248), geven aan dat de gerealiseerde ESS-respons (ESS 3, 2006-2007) voor EU-lidstaten uiteenloopt van ongeveer 45 procent (46 procent in Frankrijk) tot dik boven de 70 procent (73 procent in Slowakije) (zie Figuur 1).

Non-respons tast hoogst waarschijnlijk de representativiteit van de respons in de EB-peilingen aan. Het is in zo'n situatie gebruikelijk om te corrigeren middels weging. Daarbij wordt alle respondenten een gewicht toegekend, waarbij ondervertegenwoordigde groepen een groter en oververtegenwoordigde groepen een kleiner gewicht krijgen. Om gewichten te kunnen berekenen zijn zogenoemde weegvariabelen nodig, die in de peiling zijn opgenomen en waarvoor de populatieverdeling bekend is. Vaak is echter die populatieverdeling onbekend, wat het aantal bruikbare weegvariabelen ernstig beperkt. Uiteindelijk betreft het bij weegvariabelen dan ook veelal demografische kenmerken als geslacht en leeftijd.



Figuur 1. Respons ESS in diverse EU-lidstaten.

Gebaseerd op Stoop et al, 2010, p.93

Wil een weging effectief zijn bij het ‘repareren’ van een gebrek aan representativiteit, dan zullen de weegvariabelen aan twee belangrijke voorwaarden moeten voldoen:

- 1: De weegvariabelen moeten sterk gecorreleerd zijn met het wel of niet meedoen aan een peiling ofwel responsgedrag en in staat zijn het responsgedrag te verklaren.
- 2: De weegvariabelen moeten sterk gecorreleerd zijn met de doelvariabelen in de peiling (te onderzoeken variabelen) en in staat zijn de doelvariabelen te verklaren.

Bij de flits EB wordt een weging uitgevoerd op basis van leeftijd, geslacht, activiteit (werkend, gepensioneerd, anders) en regio op het niveau van NUTS II (in Nederland provincies). Bij de standard EB zijn de weegvariabelen leeftijd, geslacht, regio op het niveau van NUTS II en grootte van de woonplaats. In beide gevallen betreft het aldus een beperkt aantal zeer algemene variabelen. Het is daardoor onwaarschijnlijk dat aan de genoemde voorwaarden wordt voldaan: de correlaties met responsgedrag en met de doelvariabelen zullen onvoldoende sterk zijn om vertekeningen als gevolg van non-respons effectief en afdoende weg te werken.¹⁵

2.3. De vragenlijst en vragen

De mening van EU-burgers over allerlei EU-gerelateerde zaken kan alleen op een geldige en betrouwbare manier worden vastgesteld als de vragen ernaar ten minste technisch-ambachtelijk correct zijn opgesteld (zie onder veel meer Bradburn et al., 2004; Fink, 1995a; Fowler, 1995; zie in het Nederlands o.a. Dijkstra & Smit, 1999). Vragen moeten bijvoorbeeld in beginsel op een neutrale, objectieve manier worden gesteld. Sturende, suggestieve vragen dienen te worden vermeden. Dergelijke sturende vragen kunnen ontstaan door een onhandige of ronduit slechte formulering, maar er kan ook sprake zijn van manipulatie door de vraag zo te formuleren dat respondenten met hun gedachten een bepaalde richting op worden gestuurd.¹⁶

In eerdere kritische beschouwingen over de EB neemt de vraagformulering een zeer prominente plaats in (Höpner & Jurczyk 2012, 2015; Klein, 2012; Nissen, 2014).

Höpner en Jurczyk (2012; 2015) geven diverse voorbeelden van sturende en anderszins ‘dubieuze’ vragen, waarbij het opmerkelijk was dat er een patroon leek te zitten in geconstateerde onjuistheden. Alle schendingen van richtlijnen voor vraagformulering “*systematically steer responses in a pro-European, integration-friendly direction*. In fact, we did not find a single example in which the violations steered responses inversely” (Höpner & Jurczyk, 2015, p. 18; cursief in origineel). In de Nederlandse versies komen dergelijke problemen uiteraard eveneens voor, ontsprongen als alle vragenlijsten immers zijn uit een moedervragenlijst, die primair is ingegeven door ‘the needs of European Commission’s services’. We geven enkele voorbeelden, te beginnen met een vraag uit standaard EB 44.1 uit 1995, waarin de mening werd gevraagd over mogelijke uitbreiding van de EU (zie Overzicht 1).

Overzicht 1. Voorbeeld sturende vraag (EB 44.1).

Sommige mensen zeggen dat de landen van Midden- en Oost-Europa, zoals de Tsjechische Republiek, Hongarije, Polen en Slowakije, lid zouden moeten worden van de Europese Unie.

Wat vindt u hier van? Zouden zij in de komende vijf jaar lid moeten worden, over vijf tot 10 jaar, of over meer dan tien jaar?

- In de komende vijf jaar
- Over vijf tot 10 jaar
- Over meer dan tien jaar
- Ik vind niet dat deze landen lid zouden moeten worden van de Europese Unie (Enq.: Indien spontaan gezegd)
- Weet niet

Er zijn meerdere dingen mis met deze vraag. De vraag opent met de suggestie dat sommigen het een goed idee of zelfs noodzaak (‘moeten’) vinden om de EU uit te breiden; blijkbaar is het geen slecht idee voor uitbreiding te zijn, zeker omdat kennelijk niemand anders beweert (geen ‘Andere mensen zeggen...’). Ook presenteert de interviewer slechts drie antwoordmogelijkheden, die alle impliceren dat men voorstander van uitbreiding is, al staat het moment ervan niet vast. Weliswaar kan de respondent zich uitspreken tegen uitbreiding en wordt dit antwoord alsdan geregistreerd, maar die mogelijkheid wordt niet expliciet geboden. Kortom, de vraag stuurt sterk op een mening pro-uitbreiding.

Een volgend voorbeeld betreft misschien wel één van de belangrijkste EB-vragen, die dan ook regelmatig in de peilingen is opgenomen, te weten de vraag naar mogelijk voordeel van het EU-lidmaatschap van, in ons geval, Nederland (zie Overzicht 2). Eén der vuistregels van vraagformulering is dat antwoordcategorieën evenwichtig zijn: “make sure that the answers provided for a closed question are balanced” (Bryman, 2012, p. 258). Er moeten bijvoorbeeld evenveel positieve als negatieve categorieën zijn, waarbij de negatieve bij voorkeur spiegelbeeldig zijn aan de positieve categorieën. Bij de voordeelvraag is hiervan geen sprake: ‘geen voordeel’ is immers niet het tegendeel van ‘wel voordeel’. Beter ware het, als na een geherformuleerde vraag geen twee maar drie antwoordalternatieven geboden zouden worden: ‘voordeel’, ‘geen voordeel en geen nadeel’, en ‘nadeel’. Höpner en Jurczyk (2012)

beschrijven een vraagexperiment waarin de vraagvormen met elkaar zijn vergeleken: in de EB-versie gaf 49 procent aan dat het EU-lidmaatschap een voordeel (voor Duitsland) zou zijn, in het alternatief was dit 22 procent met 46 procent die aangaf dat er geen voor- of nadelen zijn. Dat een middencategorie aanzienlijke gevolgen heeft, is trouwens genoegzaam bekend: in hun standaardwerk naar vraageffecten besteden Schuman en Presser (1981, pp.161-178) een heel hoofdstuk aan de problematiek van ‘measuring a middle position’.

Overzicht 2. Voorbeeld niet-spiegelbeeldige antwoordopties (diverse EBs).

Heeft Nederland, alles welbeschouwd, naar uw idee wel of geen voordeel bij zijn lidmaatschap van de Europese Unie?

- Wel voordeel
- Geen voordeel
- Weet niet

Overzicht 3 bevat een ander voorbeeld van onevenwichtige antwoordcategorieën bij een vraag uit standaard EB 47.2 (1997) tot en met 55.1 (2001). Wat allereerst opvalt, is dat er beduidend meer positieve (7) dan negatieve (3) opties worden geboden, wat de kans op positieve antwoorden uiteraard vergroot. De vraag dient daarenboven te worden beantwoord met behulp van een toonblad, waarop alle antwoorden in vaste volgorde vermeld staan. Er zijn echter veel opties en dat rechtvaardigt de vraag of respondenten alle alternatieven (goed) lezen. Dergelijke situaties vormen een vruchtbare voedingsbodem voor *primacy* effecten, waarbij respondenten opteren voor een antwoord vooraan in de lijst en/of voor een antwoordstrategie waarbij niet het beste maar het eerst aangetroffen acceptabele antwoord wordt gegeven (*satisficing*) (zie o.a. Krosnick, 2000; Krosnick & Alwin, 1987). Beide effecten, met een opsomming die begint met niet minder dan zeven positieve punten, resulteren in een aanzienlijk grotere kans op een positief antwoord.

Overzicht 3. Voorbeeld ongebalanceerde antwoordopties (diverse EBs).

Welke van de uitspraken die u daar [op toonblad] ziet staan, beschrijven het best wat de Europese Unie voor u persoonlijk betekent?

- Een manier om een beter toekomst voor jongeren tot stand te brengen
- Een Europese overheid
- De mogelijkheid om in Europa te gaan en te staan waar ik wil
- Een garantie voor duurzame vrede in Europa
- Een middel om de economische situatie in Europa te verbeteren
- Een manier om banen te scheppen
- Een manier om de rechten van burgers te beschermen
- Veel bureaucratie, verspilling van tijd en geld
- Slechts een droom, een utopie
- Het gevaar onze culturele verscheidenheid te verliezen
- Anders
- Weet niet

Overzicht 4. Voorbeeld vraag complex onderwerp (flits EB 211).

<p>Zou Europa naar uw mening haar eigen navigatiesysteem moeten opzetten of zou Europa moeten vertrouwen op Amerikaanse, Russische of Chinese systemen?</p> <p><input type="checkbox"/> De EU zou haar eigen onafhankelijke systeem moeten ontwikkelen.</p> <p><input type="checkbox"/> Er is geen noodzaak voor een onafhankelijk systeem.</p> <p><input type="checkbox"/> Weet niet</p>	<p>80%</p> <p>12%</p> <p>8%</p>
<p>Heeft u al iets gehoord over het Europese Galileo-project?</p> <p><input type="checkbox"/> Ja.</p> <p><input type="checkbox"/> Nee.</p> <p><input type="checkbox"/> Weet niet.</p>	<p>40%</p> <p>59%</p> <p>1%</p>
<p>Galileo is de naam van het positiebepalingsysteem dat de Europese Unie zeven jaar geleden is begonnen te ontwikkelen. Tegenwoordig lijkt het erop dat er aanvullende openbare investeringen noodzakelijk zijn om het Galileo-systeem te kunnen voltooien (circa € 2,4 miljard, de kosten voor de aanleg van een autosnelweg van circa 400 km). Waar geeft u de voorkeur aan:</p> <p><input type="checkbox"/> De EU zou de noodzakelijke investeringen veilig moeten stellen om Galileo zo spoedig mogelijk te voltooien</p> <p><input type="checkbox"/> De EU zou geen extra investeringen veilig moeten stellen, zelfs als dit betekent dat het project aanzienlijk wordt vertraagd of zelfs als het dan mislukt.</p> <p><input type="checkbox"/> Weet niet.</p>	<p>63%</p> <p>23%</p> <p>14%</p>
<p>Totaal N=</p>	<p>100% 25.664</p>

Het is eenvoudig om vragen over allerlei onderwerpen te bedenken, maar het is niet evident dat respondenten daar iets vanaf weten of voordat een vraag hun wordt voorgelegd al een mening hebben (vgl. Zaller, 1992). Vaak geven respondenten desalniettemin een ‘echt’ antwoord; de ‘weet niet’ categorie, als deze al in enige variant aanwezig is, wordt door velen liever gemeden (zie Van de Maat, te verschijnen). Een voorbeeld van een vragenblokje over een (te) complex en waarschijnlijk bij het grote publiek nauwelijks bekend onderwerp komt uit flits EB 211 (2007) en betreft het EU-navigatiesysteem Galileo (zie Overzicht 4). Uit de eerste vraag, die in de stam op onverantwoorde wijze het vertrouwen in Amerikanen, Russen en Chinezen samenvoegt, blijkt dat heel veel Europeanen (80%) het belangrijk vinden dat Europa een ‘eigen, onafhankelijk’ navigatiesysteem opzet. Vervolgens blijkt, dat slechts een minderheid (40%) gehoord heeft van het Galileo-project, waarvan daarna pas de kern wordt onthuld. De uitkomst van de derde vraag is dan dat een ruime meerderheid (63%) akkoord gaat met het veiligstellen van ‘de noodzakelijke investeringen’ - let op de dubbel positieve woordkeuze! - om het project te voltooien. In deze groep zitten aldus veel mensen die

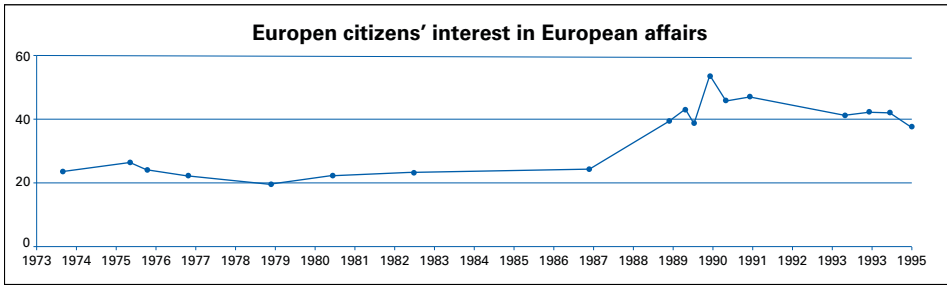
voordat de vraag werd gesteld nog nooit van het project hadden gehoord: als de antwoorden van de tweede en derde vraag worden gekruist, dan blijkt dat 59 procent van de mensen die het project niet kenden, er even zo goed 2,4 miljard in wil investeren. En ach, waar hebben we het eigenlijk over, als hiermee een mislukking wordt voorkomen en als het gaat om het equivalent van een schamele 400 kilometer snelweg...

2.4. Het longitudinale karakter

De EC is trots op de barometerreeks. “Eurobarometer has been surveying surveying the views of Europeans since 1973 and gives a unique insight into how opinions and attitudes have changed over time”, aldus de commissie in de voettekst van een publicatie naar aanleiding van veertig jaar EB (European Commission, 2014). Die reeks is zonder meer indrukwekkend, maar er vallen desondanks kanttekeningen te plaatsen bij de claim van longitudinaliteit, al was het maar vanwege de dubbelzinnigheid en daardoor potentiële verwarring van die term. Longitudinaal onderzoek kent namelijk varianten, zogenaamd trend- en cohortonderzoek waarbij opeenvolgend nieuwe steekproeven uit een en dezelfde populatie respectievelijk deelpopulatie worden getrokken en panelonderzoek, waarbij dezelfde eenheden, eenmaal getrokken, op opeenvolgende momenten worden ondervraagd (o.a. Babbie, 2007, pp.102-106).

Voor onderzoek waarbij het gaat om veranderingen van houdingen op het niveau van de individuele Europese burger is strikt genomen panelonderzoek nodig, een en dezelfde groep eenheden (mensen) die op achtereenvolgende momenten wordt ondervraagd (zie o.a. Hansen, 2008). De samenstelling van de steekproef blijft in beginsel gelijk; gemeten veranderingen in mening en houding worden toegeschreven aan ‘echte’ veranderingen.¹⁷ Voor de standard EB wordt echter geen panel in deze strikte zin gebruikt. Elke keer opnieuw wordt een steekproef getrokken, wat het moeilijk zo niet onmogelijk maakt om (individuele) verandering over tijd te meten. Gemeten verschillen kunnen zijn veroorzaakt door verschillen in steekproefsamenvatting. Alleen als ‘echte’ verschillen groter zijn dan de ruis in de steekproeven worden ze gedetecteerd, en dan op macro- en niet op microniveau.¹⁸

Bij het volgen van een verschijnsel dienen, bij panelonderzoek alsook in de minder veeleisende variant van longitudinaal onderzoek middels opeenvolgende steekproeven, vragen erover gelijk te zijn, in een in opbouw en inhoud vergelijkbare vragenlijst. In de EBs is hiervan lang niet altijd sprake; vraagteksten en antwoordmogelijkheden veranderen bijvoorbeeld. De EC zelf geeft daarvan een mooi voorbeeld in de publicatie ter gelegenheid van het 35-jarig bestaan van de EB (European Commission, 2008; zie Figuur 2). De figuur suggereert voor de periode 1973-1995 toegenomen belangstelling Europese aangelegenheden en de tekst ondersteunt die visuele boodschap: “During the first waves of the Eurobarometer, from 1973 to 1986, approximately a quarter of European citizens declared that they were very interested in Community affairs. Interest levels increased gradually after 1986. In November 1988, the proportion of respondents who expressed an interest in European affairs exceeded 30% for the first time. In November 1989, it reached the historic level of 54%” (European Commission, 2008, p.13).



Figuur 2. Belangstelling voor Europese zaken (diverse EBs).

Bron: European Commission, 2008, p.13

Overzicht 5. Voorbeeld verandering vraagtekst (diverse EBs)

Jaar	Vraagformulering
1975	In de pers (dagbladen, radio, televisie) is vaak sprake van de Europese Gemeenschap, ook wel EEG genoemd, in verband met de toekomst van Nederland of andere landen van Europa. Bent u zelf geïnteresseerd, een beetje geïnteresseerd of helemaal niet geïnteresseerd in de problemen van de Europese Gemeenschap.
1976	Bent U zelf zeer geïnteresseerd, een beetje geïnteresseerd of helemaal niet geïnteresseerd in de problemen van de EEG, de Europese Economische Gemeenschap.
1982	In de pers (dagbladen, radio, televisie) is vaak sprake van de Europese Gemeenschap, ook wel gemeenschappelijke markt genoemd. Bent u zelf geïnteresseerd, een beetje geïnteresseerd of helemaal niet geïnteresseerd in de problemen van de E.G., de Europese Gemeenschap.
1993	In hoeverre, zou u zeggen, heeft u belangstelling voor de Europese politiek, dat wil zeggen voor zaken die te maken hebben met de Europese Unie (Europese Gemeenschap). Heeft u veel belangstelling, tamelijk veel belangstelling, niet zo veel belangstelling of helemaal geen belangstelling voor Europese politiek?
2006	Zou u zeggen dat u zeer geïnteresseerd, tamelijk geïnteresseerd, niet erg geïnteresseerd of helemaal niet geïnteresseerd bent in Europese aangelegenheden.

Nadere beschouwing van de vragenlijsten leert echter dat de vraagstelling regelmatig is veranderd, zoals het EC-rapport in een kleinletterige noot aangeeft. Overzicht 5 geeft vraagteksten die in de Nederlandse versie zijn gebruikt in diverse jaren. Maar door het veranderen van die vragen wordt de vergelijkbaarheid van de ermee verkregen informatie in de tijd aangetast. Er kan niet meer onbetwistbaar worden vastgesteld of veranderingen ‘echt’ zijn of een artefact door veranderingen in de vraag. “The problem with uncovering trends over time is that not every ‘pattern’ that is

found in the data reflects a real-world change. Changes in the distribution of findings from one survey to the next may result from instrument changes” (Schmitt, 2003, p.246). Bij het maken van tijdreeksen en presenteren van longitudinale ontwikkelingen is het dan ook heel belangrijk om goed te kijken naar vragen en antwoordalternatieven die aan de trendcijfers ten grondslag liggen en is het raadzaam ten minste te wijzen op onderzoekstechnische alternatieve verklaringen voor de gesignaleerde inhoudelijke ‘ontwikkelingen’. De EC lijkt te verliefd op het longitudinale karakter van de EB en daardoor - liefde maakt ook hier blind - te eenkennig in het duiden van veranderingen als inhoudelijke wijzigingen in de onderliggende publieke opinie.

3. SLOTOPMERKING

De Eurobarometer is een instituut en blijft dat, ondanks de kritiek die erop kan, mag en wellicht moet worden geleverd - zonder te willen suggereren dat onze kritiek exclusief is voor de EB: we gaven al aan dat adel verplicht en hoge bomen nu eenmaal veel wind vangen. We hebben laten zien dat vanuit methodologisch oogpunt vraagtekens geplaatst kunnen worden bij enkele wezenlijke aspecten van dit majeure dataverzamelingsproject, dat het denken en doen aangaande de Europese Unie en het voortgaande proces van Europese integratie op tal van manieren beïnvloedt. Er zijn zeker mogelijkheden voor verbetering van de EB op het terrein van datakwaliteit en vraagformuleringen (vgl. Schmitt, 2003, p.248), maar misschien bovenal als het gaat om de verantwoording en verslaglegging van de onderzoekstechnische aanpak. Immers, in tabellen en *fancy* figuren gerapporteerde bevindingen¹⁹ zijn het voor het bredere publiek zichtbare en toegankelijke eindproduct van een onderzoeksproces waarin vele opeenvolgende en samenhangende beslissingen worden genomen en keuzen worden gemaakt. Dergelijke beslissingen en keuzen zijn noodzakelijk, maar even noodzakelijk is het daarover open en transparant te rapporteren, opdat kan worden getracht na te gaan hoe bevindingen in welke mate eventueel het gevolg zijn van een bepaalde (veranderde) aanpak of in hoeverre de gerapporteerde bevindingen de publieke opinie ten aanzien van ‘Europa’ reflecteren, waarom het tenslotte de grondleggers van de Eurobarometer allemaal te doen is geweest.

NOTEN

- 1 Met dank aan Josje den Ridder voor haar aanmerkingen op een eerdere versie van deze bijdrage.
- 2 Onder de vlag Eurobarometer varen meerdere schepen: 1) de *Standard Eurobarometer*, die onder meer de mening over Europese integratie, de Europese instituten en Europees beleid peilt en het longitudinale karakter van de EB vertegenwoordigt; 2) de *Special Eurobarometer*, die is bedoeld om de mening te peilen over een specifiek thema en waarvan de vragen, die worden ingebed in de standaard EB, in beginsel één keer worden gesteld; en 3) de *Flash Eurobarometer*, die is bedoeld om aan de hand van een korte vragenlijst snel de mening te peilen over een specifieke zaak, eventueel bij een specifieke groep burgers of andere eenheden. Daarnaast is er overigens de Qualitative Eurobarometer Survey, die in onze bijdrage buiten beschouwing blijft.
- 3 Daarnaast is de Eurobarometer voor de betrokken veldwerkorganisatie een aantrekkelijke klus van zeer aanzienlijke omvang. TNS ging er bij het binnenhalen van de opdracht in september 2013 prat op, dat het met een totaalbedrag van 85 miljoen dollar of bijna 65 miljoen euro voor vier jaar “the biggest Political & Social research contract ever commissioned” betref (<http://www.tns-opinion.com/content/tns-secures-85-million-eurobarometer-contract>).
- 4 Merk op dat een groot deel van onze kritische opmerkingen niet minder betrekking heeft of kan hebben op ander grootschalig nationaal of internationaal (vergelijkend) survey-onderzoek. Dat onze bijdrage zich exclusief richt op de Eurobarometer, laat vooral zien dat ook op onderzoeksterrein de hoogste bomen de meeste - in dit geval alle - wind vangen.
- 5 Gedurende het maken van deze bijdrage deden wij (nogmaals) een poging om de benodigde methodologische informatie te verkrijgen. In reactie op ons verzoek werd op 8 juli 2016 namens ‘Head of Unit COMM.DG.01’ expliciet toegezegd, dat een antwoord zou volgen “on the substance of those technical and methodological details form the responsible Unit, European Commission, Directorate General Communication, Unit COMM.A1 at the latest by 28 July 2016”. Op 28 juli 2016 kwam een antwoord, dat echter in nogal algemene termen terugverwees naar reeds beschikbare informatie, die in eerste instantie juist aanleiding was geweest om aanvullende, meer gedetailleerde informatie op te vragen.
- 6 Waarover wat de Nederlandse casus betreft van de kant van de veldwerkorganisatie TNS door een *research manager* het volgende wordt gemeld “De aard en omvang van de non-respons wordt niet structureel gemonitord. Doordat we van te voren telefonisch uitnodigen voor deelname, valt er - zoals te verwachten voor telefonisch onderzoek - een groot deel uit. Daarom is een structurele en gedetailleerde monitoring van de non-respons niet de moeite waard” (e-mail dd. 11 juli 2016 aan tweede auteur).
- 7 Zoals bijvoorbeeld ook voorgeschreven in relevante gedragscodes, bijvoorbeeld van ESOMAR; zie http://moa04.artoo.nl/clou-moa-web-images/images/bestanden/pdf/ICCESOMAR_CODE_NL.pdf
- 8 Waarvan niet vergeten mag worden dat deze reflectie voor ander survey-onderzoek eveneens relevant is.
- 9 We nemen deze ongelukkige maar gangbare terminologie over, terwijl eigenlijk niet van een representatieve steekproef kan worden gesproken maar van een simpele aselecte, *random* of waarschijnlijkheidssteekproef, die niet als zodanig maar naar verhoopt wat de uitkomsten betreft representatief is.
- 10 Lang niet altijd onder burgers overigens, maar bijvoorbeeld ook onder bedrijven.
- 11 Aldus <http://www.gesis.org/eurobarometer-data-service/survey-series/flash-eb/sampling-fieldwork/> (geraadpleegd 7 juli 2016).
- 12 Zie <https://www.bel-me-niet.nl/nieuws/berichten/8-miljoen-inschrijvingen-bel-me-niet-register>
- 13 Een verzoek om specifieke informatie kon vanwege de MOA Onderzoekfilter niet worden gehonoreerd, omdat “het systeem onder repair” was en de gevraagde gegevens niet kon opleveren (e-mail dd. 15 juli 2016 aan tweede auteur).
- 14 Aldus Martin Boon, directeur van ICM Research; zie <http://www.bbc.com/news/uk-politics-33228669>
- 15 Dat een adequate weging niet alleen voor de EB problematisch is maar voor elk internationaal vergelijkend survey-onderzoek is op zichzelf waar, maar doet als zodanig natuurlijk niet af aan het feit dat de bij de EB gehanteerde weegaanpak waarschijnlijk niet of nauwelijks effectief zal zijn.

- 16 Een fictief maar fraai voorbeeld van het effect van *leading questions* is onderdeel van een aflevering van Yes Prime Minister, waarin Sir Humphrey Appleby middels een korte vragenlijst toont hoe een mening als het ware wordt gevormd door een reeks opeenvolgende vragen; zie <https://www.youtube.com/watch?v=G0ZZJXw4MTA>
- 17 Een onontkoombaar nadeel van een dergelijk panel is wel, naast de hoge kosten van onderhoud ervan, dat mensen om uiteenlopende redenen of oorzaken wegvallen (paneluitval) bij opeenvolgende metingen.
- 18 Merk overigens op dat de EB ook, anders dan longitudinaal onderzoek veronderstelt, betrekking heeft op een in de loop der tijd herhaaldelijk gewijzigde populatie van *de* EU. Daarmee kan in de analyse en presentatie van bevindingen rekening worden gehouden, maar onderzoekers dienen dan wel voortdurend alert te zijn en te rapporteren op *welke* EU een en ander betrekking heeft.
- 19 Ter wille van de omvang van deze bijdrage laten we de niet altijd even geslaagde manier van presenteren van EB-bevindingen hier verder buiten beschouwing; zie over visuele presentatie van data o.a. Jones, 2007.

LITERATUUR

- Babbie, E. (2007). *The Practice of Social Research*. Belmont: Thomson Higher Education. 11th ed.
- Bethlehem, J. (2009). *Applied Survey Methods: A Statistical Perspective*. New York: John Wiley & Sons.
- Bethlehem, J.G., F. Cobben & B. Schouten (2011). *Handbook of Nonresponse in Household Surveys*. New York: John Wiley & Sons.
- Bläser, K.-A. (2013). Europa im Spiegel der öffentliche Meinung: Bilanz und Perspektiven des Eurobarometers nach 40 Jahren, *Leviathan*, 41(3), 351-357.
- Boon, L.P. (1982). *Ook de afbreker bouwt op*. Amsterdam: De Arbeiderspers.
- Bradburn, N., S. Sudman & B. Wansink (2004). *Asking Questions: The Definitive Guide to Questionnaire Design – For Market Research, Political Polls, and Social and Health Questionnaires*. San Francisco: Jossey-Bass. Rev.ed.
- Bryman, A. (2012). *Social Research Methods*. Oxford: Oxford University Press. 4th ed.
- Cochran, W.G. (1977). *Sampling Techniques*. New York: John Wiley & Sons. 3rd ed.
- Dijkstra, W., J. Smit (1999). *Onderzoek met vragenlijsten: Een praktische handleiding*. Amsterdam: VU Uitgeverij.
- European Commission (2008). *35 Years of Eurobarometer. European integration as seen by public opinion in the Member States of the European Union: 1973-2008*. http://ec.europa.eu/public_opinion/docs/35_years_en.pdf
- European Commission (2014). *Effects of the economic and financial crisis on European public opinion. 40 Years Eurobarometer*. http://ec.europa.eu/public_opinion/topics/eb40years_en.pdf
- European Union (2015). *Standard Eurobarometer 84 – Autumn 2015. Public opinion in the European Union: First results*. z.p.: European Union.
- Fink, A. (1995a). *How to Ask Survey Questions*. Thousand Oaks: SAGE. The Survey Kit Vol. 2.
- Fink, A. (1995b). *How to Sample in Surveys*. Thousand Oaks: SAGE. The Survey Kit Vol. 6.
- Fowler, F.J., Jr. (1995). *Improving Survey Questions: Design and Evaluation*. Thousand Oaks: SAGE. Applied Social Research Methods Series Vol. 38.
- Gabel, M., S. Hix & G. Schneider (2002). Who is afraid of cumulative research? Improving data on EU politics, *European Union Politics*, 3(4), 481-500.
- Hansen, J. (2008). Panel surveys, in: W. Donsbach, M.W. Traugott (eds.), *The SAGE Handbook of Public Opinion Research*. Los Angeles etc.: SAGE, 330-339.
- Höpner, M. & B. Jurczyk (2012). Kritik des Eurobarometers. Über die Verwischung der Grenze zwischen seriöser Demoskopie und interessengeleitet Propaganda, *Leviathan*, 40(3), 326-349.
- Höpner, M. & B. Jurczyk (2015). *How the Eurobarometer Blurs the Line between Research and Propaganda*. Cologne: Max-Planck-Institut für Gesellschaftsforschung. MPIfG discussion paper 15/6.
- Jones, G.E. (2007). *How to Lie with Charts*. Santa Monica: LaPuerta. Sec. ed.
- Kish, L. (1965), *Survey Sampling*. New York: John Wiley & Sons [in 1995 heruitgegeven in de Wiley Classics Library].
- Klein, M. (2012). *Europäische Meinungsmacher - Wie man Umfrageforschung für seine Zwecke missbraucht*. <http://www.sciencefiles.org/2012/03/06/europaische-meinungsmacher-wie-man-umfrageforschung-fur-seine-zwecke-missbraucht>
- Kohler, U. (2007). Surveys from inside: An assessment of unit nonresponse bias with internal criteria, *Social Research Methods*, 1(2), 55-67.
- Krosnick, J.A. (2000). The threat of satisficing in surveys: The shortcuts respondents take in answering questions, *Survey Methods Newsletter*, 20(1), 4-8.
- Krosnick, J.A., D.F. Alwin (1987). An evaluation of a cognitive theory of response order effects in survey measurement, *Public Opinion Quarterly*, 51(2), 201-219.
- Maat, J. van de (te verschijnen). *Public Opinion Without Opinion?!*. Leiden: Universiteit Leiden [proefschrift].
- Marcus, J. (2009). Der Einfluss von Erhebungsformen auf den Postmaterialismus-Index, *Methoden-Daten-Analysen*, 3(2), 137-166,
- Nelsen, B.F., J.L. Guth & C.R. Fraser (2001). Does religion matter? Christianity and public support for the European Union, *European Union Politics*, 2(2), 191-217.

- Nissen, S. (2014). The Eurobarometer and the process of European integration: Methodological foundations and weaknesses of the largest European survey, *Quality & Quantity*, 48(2), 713-727.
- Rivers, D. (2007). *Sampling for Web Surveys*. Paper presented at the Joint Statistical Meetings, Section on Survey Research Methods. Salt Lake City, Utah.
- Schmitt, H. (2003). The Eurobarometers: Their evolution, obvious merits, and ways to add value to them, *European Union Politics*, 4(2), 243-251.
- Schuman, H., S. Presser (1981). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. San Diego etc.: Academic Press.
- Signorelli, S. (2012). *The EU and Public Opinions: A Love-Hate-Relationship?* Paris: Notre Europe - Jacques Delors Institute. Studies & Reports 93.
- Stoop, I.A.L. (2005). *The Hunt for the Last Respondent: Nonresponse in sample surveys*. The Hague: Social and Cultural Planning Office of the Netherlands.
- Stoop, I., J. Billiet, A. Koch, & R. Fitzgerald (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. Chichester: John Wiley & Sons.
- Vehovar, V. (1999). Field substitution and unit nonresponse, *Journal of Official Statistics*, 15(2), 335-350.
- Vreese, C.H., H.G. Boomgaarden & H.A. Semetko (2008). Hard and soft public support for Turkish membership in the EU, *European Union Politics*, 9(4), 511-530.
- Zaller, J.R. (1992). *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.