

## Poging een krachtige en unieke persoonlijke ervaring samen te vatten

Deze blog gaat over vragenlijstonderzoek. De statisticus David Spiegelhalter definieerde getallen uit vragenlijstonderzoek als onvolkomen pogingen om een verzameling van krachtige en unieke persoonlijke ervaringen samen te vatten. Laat ik beginnen met het persoonlijke.

Onlangs nam ik, samen met zo'n 60 leden van [Scientist Rebellion](#) en ruim 1500 andere bezorgde Nederlanders, deel aan de veertigste blokkade van de A12 in Den Haag. Zoals



steeds, wilden we burgerlijk ongehoorzaam, maar vreedzaam, protesteren tegen het dagelijks weggeven van €110 miljoen van 'hardwerkende' Nederlanders in de vorm van belastingvoordelen aan bijvoorbeeld Shell Nederland, Dow Benelux, Yara Sluiskil, ExxonMobil Benelux, BP Nederland en Tata Steel. Allemaal bedrijven die forse klimaatschade

veroorzaken en blijven lobbyen voor beleid dat de noodzakelijke energietransitie vertraagt. Burgemeester van Zanen maakte er een nodeloze principekwestie van maar schoot daarmee zichzelf lelijk in de voet. De toegang tot de A12 vanaf het Malieveld was hermetisch afgesloten met hoge hekken, die bedekt waren met zwarte schermen om de demonstratie van sympathisanten te ontmoedigen. Veel activisten, waaronder mijn groep, konden de A12 niet op en, plan B volgend, bezetten we enkele kruispunten naast het Malieveld. Er ontstond heel wat overlast. Wanneer Van Zanen de ratio had laten winnen van zijn emotie en ons had toegelaten tot het symbolische stukje A12 tussen Tweede Kamer en Ministerie van Economische Zaken & Klimaat was de overlast voorspelbaar en minder geweest. De extra overlast leidde tot heel wat media-aandacht, een cruciaal doel van burgerlijke ongehoorzaamheid. Om klimaatactivisten in een kwaad daglicht te stellen, berichtte de politie dat ambulanceverkeer werd gehinderd. Een leugen omdat de activisten altijd ogenblikkelijk een rijstrook vrijmaken voor ambulances die een blokkade willen passeren, om daarna de rijen weer te sluiten.

De volgende dag luisterde ik naar Gijs Rademaker die bijna triomfantelijk vertelde dat uit opinie-onderzoek in zijn [RTL-panel](#) blijkt dat Nederlanders het klimaat op plaats 10 van

hun prioriteiten hebben staan. “Achter immigratie van Nigerianen, immigratie van Afghanen, (immigr ..), immigratie van Soedanezen en woningschaarste (door immigratie)”, dacht ik bij mezelf. Dat bleek iets genuanceerder te liggen, maar ter zake. In hoeverre kun je resultaten van opinie-onderzoek via zelfselectie-panels vertrouwen? Zo’n panel bestaat uit mensen die zichzelf aanmelden. Dat is nooit een willekeurige steekproef van alle Nederlanders, want de hardwerkende hebben er geen tijd voor.

“In hoeverre kun je resultaten van opinie-onderzoek via zelfselectie-panels vertrouwen?”

Ik deed snel een experiment dat ethische toetsing vooraf nooit had overleefd. Ik meldde me aan als RTL-panellid. Tien minuten later was ik lid als transgender, met 3 jaar havo, wonend in Noord Holland (maar niet in groot Amsterdam), gescheiden samenwonend met kinderen, werkend in loondienst en bij de laatste verkiezingen stemmend op de PVV. Ik werk inderdaad in loondienst. Ik had verwacht dat men mij zou laten verklaren dat ik alles naar waarheid had ingevuld, [want zoiets helpt](#), maar dat hoefde niet. Of dat ik een week moest wachten op werkelijk lidmaatschap zodat Gijs mijn nepprofiel m.b.v. kunstmatige intelligentie direct kon ontmaskeren. Niets ervan. Iets later die avond maakte ik een tweede profiel vanaf hetzelfde IP-adres. Ik ben daarin een vrouw in Limburg met enkele interessante kenmerken. Dat die kenmerken kloppen is cruciaal, want RTL weet dat zelfselectie in deelname aan het panel leidt tot niet-representativiteit. Door [enquête-resultaten uit het panel te wege](#)n op leeftijd, geslacht, opleiding, huishoudsamenstelling en politieke voorkeur probeert RTL representativiteit op die vijf kenmerken te herstellen. Dat kan echter alleen voor die vijf. Het kan niet voor kenmerken waarvan men de percentages in het panel weliswaar kent, maar niet die in de algemene bevolking. In feite kan RTL dus alleen enigszins geldige vragenlijsten uitzetten over onderwerpen waarvan de resultaten zeer sterk samenhangen met deze vijf kenmerken. Deze vijf kenmerken (kunnen) veranderen in de tijd (ja, zelfs geslacht) en ze moeten dus regelmatig worden geüpdatet. Anders klopt de weging niet meer. Updaten panelleden regelmatig? De RTL website vertelt het ons niet.

RTL is summier over haar methodologie. [Jelke Bethlehem](#), voormalig survey-specialist van Statistics Netherlands (het CBS) gedurende meer dan 35 jaar, liet me weten dat hij resultaten uit [panel-surveys wantrouwt \(zie het artikel op blz.32 in de link\)](#), zelfs na weging.

Zo, dat was het persoonlijke. Nu de onderzoeksmethodologie. De complexiteit van vragenlijstonderzoek (‘polls’) wordt onderschat. Veel van onze studenten doen vragenlijstonderzoek. Ik hoor wel eens dat het zo vaak gebeurt dat Amsterdammers rond

“Veel van onze studenten doen vragenlijstonderzoek. Hoeveel leren onze studenten over de methodologie van polls?”

living labs verlangen naar een vragenlijstvrije zone rond living labs. Hoeveel leren onze studenten over de methodologie van polls? Een heel interessante bron om spelenderwijs meer te leren over methodologie voor dit type onderzoek is het boek [Sex by Numbers](#) van David Spiegelhalter. Mensen bevragen over hun seksualiteit ligt gevoelig. Juist

daarom hebben [sommige onderzoekers](#) in dit veld diep nagedacht over hoe je vragenlijstonderzoek zorgvuldig doet. Spiegelhalter onderscheidt vijf niveaus van betrouwbaarheid (validiteit): niveau 0 betekent ‘borrelpraat’ zonder onderliggende evidence. Niveau 1 betekent ‘onbetrouwbare cijfers’, met als voorbeeld [Shere Hite’s Report](#) on Female Sexuality (1976) waarvoor ze 100 000 vragenlijsten verspreidde onder specifieke vrouwengroepen (zelfselectie) en er 3019 retour kreeg. Niveau 2 betekent cijfers uit niet-willekeurige steekproeven, die er een factor 2 op de odds-schaal naast kunnen zitten (d.w.z. een schatting van bijvoorbeeld 50% kan een waarheid van 33% of 66% betekenen). Niveau 3 betekent cijfers uit willekeurige steekproeven met herhaalde pogingen om non-responders te bereiken en vraagmethodes die eerlijke antwoorden uitlokken; cijfers die er relatief niet meer dan 25% naast zitten (een schatting van bijvoorbeeld 12% betekent dan een waarheid van tussen de 9% en 15%). Niveau 4 tenslotte zijn officiële statistieken, zoals geboortecijfers.

Tenslotte, beste Gijs, ik beloof je dat ik in de komende weken mijn twee RTL-nepprofielen zal opheffen. Als me dat niet lukt, volg jij dan jouw plan B: “Als de panelredactie op basis van deelname vermoedt dat panelleden niet oprecht zijn in hun deelname, kunnen de antwoorden van deze panelleden uit de analyse verwijderd worden.”

Gerben ter Riet

13 januari 2025