

NPSO bijeenkomst ‘paradata’ 11-3-2010 te Brussel

Verslag: Peter Lugtig p.lugtig@uu.nl

De zesde lezingen middag van het Nederlandstalig Platform voor Survey Onderzoek ging over paradata. Paradata zijn data die voortkomen uit een interview die iets vertellen over de context en proces van het interviewen. De presentaties van iedere spreker zijn te downloaden via www.npsso.net. Hier worden ook volgende bijeenkomsten van het NPSO op aangekondigd:

Programma:

- 13:00-13:15 – Inleiding ‘wat zijn paradata en welke data voor welk gebruik verzamelen’
– *dr. Ann Carton, studiedienst van de Vlaamse regering*
- 13:30 – 14:15 - Het verbeteren van web en elektronische vragenlijsten door het gebruik van ‘audit trail’.
- *dr. Mattijn Morren en prof. Dr. Ger Snijkers, Centraal Bureau voor de Statistiek, Nederland*
- 14:15:15:00 - Concrete voorbeelden van het gebruik van paradata bij het management van (face-to-face) dataverzameling
- *Luc Schulpen, managing director bij TNS Dimarso*
- 15:30-16:30 - Vaststellen en remediëren van non-response bias met behulp non-response surveys in ESS ronde 3 (2006-2007) met data uit België, Noorwegen, Polen en Zwitserland.
- *Hideko Matsuo, prof. em. dr. Jaak Billiet en prof. dr. Geert Loosveldt, Centrum voor Sociologisch Onderzoek (CeSO), K.U.Leuven*

Korte samenvattingen

1. Ann Carton – inleiding ‘wat zijn paradata en welke data voor welk gebruik verzamelen’

Ann Carton begint met een inleiding over paradata (procesdata), een term geïntroduceerd door Mick Couper eind jaren '90. Paradata zouden idealiter iets zeggen of kunnen helpen bij het analyseren van de kwaliteit van het surveyproces. De introductie van Computer gestuurd interviewen (CAPI/CATI/CASI) heeft het mogelijk gemaakt paradata eenvoudig te verkrijgen. Potentieel zouden zij tijdens en na het veldwerk, veel functies kunnen hebben, maar zijn paradata in de praktijk ook bruikbaar? Ann bespreekt een aantal mogelijkheden en onmogelijkheden in het verzamelen en gebruik van paradata.

Voor het corrigeren van data (wegen) is het tevens nodig dat paradata samenhangen met afhankelijke variabelen in het onderzoek. Is het gebruik van paradata daarom wel zinvol als het gaat om het verbeteren van datakwaliteit? Wat te doen met privacy-issues door het gebruik van paradata?

2. Ger Snijkers – Het verbeteren van web en elektronische vragenlijsten door het gebruik van ‘audit trail’.

Ger Snijkers vertelt over het gebruik van paradata in audit trails voor bedrijvenquêtes bij het Centraal Bureau voor de Statistiek (CBS). Paradata zijn naar zijn idee met name zinvol in het ontwikkelen van de vragenlijsten. Je zou kunnen analyseren welke acties een respondent achtereenvolgens neemt bij het beantwoorden van een vraag, en bijvoorbeeld ook de tijd coderen die het duurt om een vraag of sectie te doorlopen. Hiermee krijgt de onderzoeker inzicht in waar de problematische secties van de vragenlijsten zich bevinden.

In de audit trails bij het CBS wordt gecodeerd hoe lang respondenten over een vraag doen, of ze de vragenlijst printen, hem later invullen, gegevens versturen, in hoeveel sessies ze werken etc. Ook presenteert Ger data over de tijd tussen respons en het versturen van uitnodigingsbrieven en reminders. Hieruit heeft het CBS geleerd over de datum van het versturen van uitnodigingsbrieven en de tijdspanne tussen het versturen van uitnodigingen en herinneringsbrieven aan bedrijven.

3. Luc Schulpen – Paradata bij TNS Dimarso

Luc vertelt over gebruik van paradata bij TNS Dimarso. TNS gebruikt veel paradata over contactproces. Bij elke contactpoging noteert interviewer de aard en het resultaat van de poging. Daarnaast wordt in elektronische surveys de interviewduur gecodeerd. De contactdata worden gebruikt tijdens het gehele veldwerkproces – met interviewers kan contact worden opgenomen als blijkt dat het hele project of het individuele contact met respondenten niet goed gaat. Ook kan worden geanalyseerd of per project of per regio de enquête bijsturing behoeft.

4. Hideko Matsuo – Universiteit Leuven

Binnen de European Social Survey (ESS) zijn er veel verschillen in niveaus van Nonresponse. Dit kan leiden tot verschillen in nonresponse bias tussen landen. Om zicht te krijgen op mogelijke verschillen in bias, is in Polen, Noorwegen, Zwitserland en België een survey gehouden onder nonrespondenten in de derde ronde van de ESS. In België had deze de vorm van een aantal centrale vragen die door de interviewer meteen na een ‘zachte weigering’ aan de deur werden gesteld. In de andere landen werden nonrespondenten op een later moment herbenadert.

Door de antwoorden van de respondenten binnen het ESS te vergelijken met de respondenten op de Nonresponse-survey kan inzicht worden verkregen in een deel van de bias die er bestaat in de verschillende landen. Met behulp van propensity-scores worden de twee samples naar elkaar toe gewogen. Op de meeste variabelen verdwijnen in alle landen de verschillen, behalve op de variabele politieke interesse. Er blijven op andere variabelen in sommige landen wel en andere landen geen verschillen bestaan, wat kan duiden op een verschil in nonresponse bias. De verschillende interviewstrategieën en responspercentages in de verschillende maken het moeilijk om definitieve conclusies te trekken.

Verslag discussie
Jan Pickery – Studiedienst van de Vlaamse Regering
jan.pickery@dar.vlaanderen.be

Presentatie 1

Het verbeteren van web en elektronische vragenlijsten door het gebruik van ‘audit trail’
Mattijn Morren & Ger Snijkers

Vraag

Mijn ervaring is dat er op paradata zeer veel ruis zit. Ook in jouw presentatie waren er voorbeelden van uitzonderlijk lange (onmogelijke) invultijden. Ben je wel zeker van de interpretatie die je geeft aan de data?

Antwoord

Eigenlijk is er geen echte ruis. Het gaat telkens over acties van de respondent die geregistreerd worden. Wel is het zo dat je lang moet nadenken over de interpretatie.

Vraag

Is jullie interpretatie van die profielen wel correct? Hebben jullie geprobeerd om profielen (“zorgvuldig”, “quick and dirty”, ...) te valideren? Je zou kunnen teruggaan naar bedrijven om na te gaan hoe ze die vragenlijsten daadwerkelijk hebben ingevuld.

Antwoord

Interessante suggestie, dat is vooralsnog niet gebeurd.

Vraag

Kunnen jullie een aanwijzing geven over de frequentie van bepaalde profielen? Bvb. hoeveel mensen doorlopen eerst de vragenlijst? Hoeveel mensen beginnen er gewoon direct aan (“Kick and rush”)?

Antwoord

We hebben wel enkele profielen opgesteld, maar het is niet evident om te tellen hoe vaak die voorkomen. We hebben wel nog de intentie om dat te doen, maar voorlopig kunnen we deze vraag dus niet beantwoorden.

Presentatie 2

Concrete voorbeelden van het gebruik van paradata bij het management van (face-to-face) dataverzameling

Luc Schulpen

Vraag

Hoever gaan jullie in het managen van de financiële implicaties? Bvb. welke contactpogingen kosten meer geld? Op welk moment? Bij een groter aantal adressen per cluster zijn contactpogingen doorgaans goedkoper. Hoe gebruiken jullie dat?

Antwoord

Het is natuurlijk altijd onze doelstelling om het rendement van steekproefdesigns te proberen maximaliseren. Wij vinden wel dat opdrachtgevers te weinig rekening houden met kostoverwegingen bij het opstellen van hun design of bij het bepalen van contactprocedures. Naast de kostoverwegingen kunnen er ook implicaties zijn voor de motivaties van interviewers waar opdrachtgevers eveneens te weinig oog voor hebben.

Vraag

Leggen jullie linken tussen de verschillende monitoring- en knipperlichtsystemen? Wij hebben bvb. de indruk dat er ruwweg verschillende types interviewers zijn, zoals bvb. interviewers die een hoge respons behalen maar slechte interviewkwaliteit afleveren en vice versa. Kunnen jullie de verschillende alarmsystemen globaal overschouwen?

Antwoord

Dat gebeurt al. Een combinatie van alle elementen van het interviewergedrag komt samen in de interviewerrapporten. Die maken bijsturing mogelijk. Het is dus zeker geen zwart/wit verhaal van goede en slechte interviewers. Soms moeten wij echter ook wel toegeven dat er bij bepaalde interviewers niets meer bij te sturen valt. Die zijn dan voor bepaalde projecten niet meer selecteerbaar.

Vraag

Kunnen jullie jullie systeem dat zeker een meerkost betekent ook verkopen aan commerciële partners?

Antwoord

Dat varieert natuurlijk. Maar het is zeker niet zo dat die commerciële partners er helemaal niet in geïnteresseerd zijn. Zij willen in eerste instantie wel weten, wat leren die paradata mij over mijn data. Als wij de impact duidelijk kunnen maken en kunnen aangeven wat ze meer weten kan je hun interesse wel opwekken. Ook dit is zeker geen zwart/wit verhaal.

Vraag

We vragen meer en meer van interviewers: goed afnemen van vragenlijsten, aangepaste contactstrategie, correcte registratie van paradata, ... Blijft dat allemaal haalbaar?

Antwoord

Goede vraag. Kijk ook eens naar de vragenlijsten die almaar langer worden. Ik denk inderdaad dat een zekere selectie zich opdringt. Wat hebben we echt nodig aan paradata? Wat levert het op? Een grondige selectie op basis van die afwegingen is aangewezen.

Presentatie 3

Vaststellen en remediëren van non-response bias met behulp nonrespons surveys in ESS ronde 3 (2006-2007) met data uit België, Noorwegen, Polen en Zwitserland.

Hideko Matsuo, Jaak Billiet en Geert Loosveldt

Vraag

Je vermeldde de attitude t.o.v. surveys, maar hebt die variabele later niet meer gebruikt.

Waarom niet?

Antwoord

Deze vraag werd alleen gesteld aan de nonrespondenten.

Vraag

Een DSQ is dat eigenlijk toegelaten, gegeven de richtlijnen van de privacycommissie?

Antwoord

Strikt genomen niet, maar een enquête, zelfs een korte, enkele maanden later is natuurlijk veel duurder.

Vraag

Hebben jullie op alle opgesomde variabelen gewogen?

Antwoord

Inderdaad, op alle variabelen die zowel voor de respondenten als de nonrespondenten van de hoofdsurvey beschikbaar waren, 6 variabelen dus.

Vraag

Is de benaderde groep van nonrespondenten in België een goede afspiegeling van de totale groep nonrespondenten?

Antwoord

Waarschijnlijk niet. Ten eerste werd de DSQ alleen afgenomen bij weigeraars en die verschillen natuurlijk van de niet-gecontacteerde respondenten en ten tweede heeft die DSQ natuurlijk geen 100% respons bij de weigeraars. Als we kijken naar bvb. buurtkenmerken, dan kunnen we ervan uitgaan dat ook die respons selectief was.

Vraag

Hoe verloopt eigenlijk de praktische organisatie van zo'n DSQ?

Antwoord

De enquêteur moet na de weigering nog een korte vragenlijst afnemen en wordt daar ook voor betaald, zij het natuurlijk minder dan voor een volledig afgenomen interview. Goede enquêteurs kunnen die DSQ waarschijnlijk verkopen als een soort trade-off: "OK, u wilt niet deelnemen aan het uitgebreide onderzoek, maar misschien kan ik u toch enkele korte vraagjes voorleggen".

Vraag

Wat is eigenlijk de procedure voor de berekening van de normale ESS-gewichten?

Antwoord

Dat is niet mijn domein, maar het gaat om een combinatie van designgewichten, nonresponsgewichten en poststratificatiegewichten.

Algemene discussie achteraf

In de algemene discussie achteraf was er veel aandacht voor het aspect *privacy*.

Vraag

Wordt er in de introductie aan de respondent meegedeeld dat er paradata worden geregistreerd? Hoe doe je dat zonder argwaan op te wekken?

Antwoorden en reacties

- Zijn paradata gevoeliger dan de andere data die je verzamelt?
- Misschien niet, maar de respondent is er zich bij die andere data wel van bewust en kan weigeren om die te verstrekken. Je kan als respondent niet kiezen om paradata niet te verstrekken.
- De perceptie is ook belangrijker (“alles wat ik zeg wordt op een bepaalde manier geregistreerd”) dan welke data je juist verzamelt.
- Is er wel een probleem? We verzamelen steeds meer paradata en beginnen die nu zelf privacygevoelig te noemen, maar buiten wijzelf ziet er misschien niemand een probleem.
- Er is in samenwerking met Mick Couper een onderzoek gaande naar de gevoeligheid voor de registratie van paradata. Resultaten zouden binnenkort beschikbaar moeten zijn.

Ander discussiepunt, de *crossnationale vergelijkbaarheid* van paradata

Vragen en opmerkingen

- Bij contactpogingen is de definitie van 's avonds misschien wel anders in Spanje dan in de lage landen?
- Zijn achtergestelde buurten in Napels dezelfde als in Zürich?
- We mogen ons bijgevolg niet blindstaren op de paradata op zich, maar moeten aan veldwerkmensen uitleg vragen over hoe die paradata juist geregistreerd worden.
- Paradata maken idealiter ook deel uit van de contextdata (zie bvb. ESS), want zij kunnen verantwoordelijk zijn voor verschillen tussen landen.