

NAAR EEN STANDAARD-DEFINITIE VAN DE RESPONSFRACTIE IN PERSONEN- STEEKPROEVEN *)

S.E. de Bie **)

Samenvatting

Een door Wiseman & Billington (1984) gegeven aanzet voor een standaard-definitie van responsfracties, wordt uitgewerkt voor face-to-face interviews in personensteekproeven. De kern van de voorgestelde responsfractie bestaat uit een schatting van de mate van overdekking in de groep van personen die weigerden of niet bereikbaar bleken. Rekening wordt gehouden met de hoeveelheid weigeringen waarvan wel is vastgesteld dat de betreffende personen tot de doelpopulatie horen.

1. Inleiding

Iedereen weet, dat de responspercentages in survey-onderzoek terug lopen. Niemand echter weet wat bringers van deze jobstijding daar *precies* mee bedoelen. Dat het steeds moeilijker wordt om medewerking van respondenten te krijgen, lijkt niet aan twijfel onderhevig. Met name diegenen die onderzoek doen onder (groepen van) beroepsbeoefenaren, melden een toenemende non-respons. Schattingen door medewerkers van de bij de Vereniging van Onderzoek Instituten aangesloten instituten varieerden voor de onderwijssector (leerkrachten, schoolhoofden) van 45% tot 60% respons gemiddeld. In de bedrijfssector bleek 30% geen uitzondering meer te zijn. In gevallen waarin een bedrijfstakorganisatie zelf opdrachtgever was, bleek een respons van meer dan 60% nauwelijks

* Met dank aan Aad Reeuwijk en Jelke Bethlehem voor commentaar op een eerdere versie.

** S.E. de Bie
Stuurgroep Dataverzameling
Vereniging van Onderzoek Instituten
Hooigracht 15
2312 KM Leiden
071-273625

haalbaar. Beroepsbeoefenaren zijn in toenemende mate enquête-moe.

Een opmerkelijk verschijnsel doet zich ook voor in onderzoeken onder culturele minderheden. De responspercentages lopen voor verschillende nationaliteiten flink uiteen. In een onderzoek naar het mediagebruik van diverse groepen minderheden, varieerden de responspercentages van 82% voor Turken tot 52% voor Portugezen; Marokkanen hadden 65% respons (Veldkamp, 1986, p. 14). Hier zijn echter zowel verhuizingen als niet-kloppende adressen tot de 'oneigenlijke non-respons' gerekend. Laat men deze correcties achterwege, dan zijn de cijfers respectievelijk 75%, 42% en 55%. In het CBS-onderzoek naar de leefsituatie van Turken en Marokkanen in 1984, moeten de responscijfers eveneens herberekend worden, omdat men niet-bezochte adressen tot de 'oneigenlijke non-respons' heeft gerekend (CBS, 1985, p.11). De respons was aldus 67% (i.p.v. 76) voor Turken en 54% (i.p.v. 64) voor Marokkanen. In eerder uitgevoerd onderzoek door de NOS werd onder Marokkanen 38% respons verkregen. Een vergelijking met de gerealiseerde respons voor Turken en Portugezen in dat onderzoek is niet mogelijk: "Van de gepleegde contacten met Turkse, Griekse en Portugese adressen is door de enquêteurs helaas te gebrekkig verslag gedaan, om daar een betrouwbaar inzicht in te krijgen." (Wentholt, 1983, p. 32). En dat is jammer.

Stellingen over toenemende non-respons zijn moeilijk met cijfers te onderbouwen: voorzover steekproeven al vergelijkbaar zijn, is de responsverantwoording veelal summier en/of niet uniform. Vergelijkbaarheid is wel te vinden, zij het ook niet zonder problemen, in onderzoeken die regelmatig herhaald worden. Een vergelijking van vijf CBS-onderzoeken in de periode 1972-1983 leert dat de non-respons in vier van de vijf onderzoeken in die jaren fors is toegenomen (met een factor 1.5 tot 1.75); in het vijfde onderzoek is de respons vrijwel gelijk gebleven (Bethlehem & Kersten, 1986, p. 40).

Een betere vergelijkbaarheid van non-respons percentages is een voorwaarde voor kwantificering van het genoemde probleem. Dan kunnen vragen als 'wat is een respons van 80% waard?' (Fluttert, 1980) en 'Wat niet weet, wat niet deert?' (Bethlehem & Kersten, 1982) gerelateerd worden aan de toenemende ernst van het probleem, en dus aan de toenemende risico's van dit soort onderzoek. Mijn indruk is bovendien dat een duidelijke definitie ook enige repercussies zal hebben voor de statistische verantwoording van sommig steekproef-onderzoek; ik denk in het bijzonder aan de gewoonte om niet-gebruikte (reserve-) namen en/of verkeerde adressen tot de kaderfouten te rekenen in plaats van tot de non-respons. Het loont daarom de moeite een uniforme berekening van de non-respons in survey-onderzoek na te streven. Wiseman & Billington (1984) gaven daartoe een aanzet, overgenomen door Bethlehem & Kersten (1986). Wiseman & Billington gebruiken de volgende berekeningswijze:

$$F_R = \frac{R}{R + \frac{R}{R + O_V} \cdot (C_g + W)} \quad (1)$$

waarin:

- F_R = responsfractie
- R = aantal gecompleteerde interviews
- O_V = in het veld geconstateerde overdekking
- C_g = hoeveelheid 'geen contact'
- W = aantal weigeringen

Uit deze berekeningswijze is meteen de belangrijke veronderstelling af te leiden, die ik ook zal maken: de verhouding tussen personen die wel en die niet tot de beoogde populatie behoren, is hetzelfde in de groep waarvoor deze verhouding kan worden vastgesteld ($R + O_V$) als voor de groep waarover dit gegeven niet bekend is ($C_g + W$), namelijk $R : O_V$.

Deze aanzet verdient uitwerking. De berekening kan in bovengenoemde vorm sterk vereenvoudigd worden tot

$$F_R = \frac{R + O_V}{B} \quad (2)$$

waarin:

B = bruto steekproef.

Echter, anderzijds is juist een wat gecompliceerder berekening nodig. Dat heeft te maken met het begrip 'overdekking'. Hieronder een uitwerking voor face-to-face interviews in een personensteekproef, voorafgegaan door een bespreking van het begrip 'overdekking'.

2. Overdekking

Non-respons is het verschijnsel dat van elementen die tot de steekproefpopulatie behoren en die in de steekproef zijn getrokken, in het geheel geen gegevens kunnen worden verkregen. Te onderscheiden dus van partiële non-respons, wanneer slechts een deel van de gegevens kan worden verkregen (Bethlehem & Kersten, 1986, p. 25).

In principe is het responspercentage (de responsfractie) eenvoudig te berekenen: het aantal geslaagde interviews gedeeld door de netto steekproef. In dat laatste begrip zitten echter twee problemen die te maken hebben met *overdekking*: overdekking is het verschijnsel dat het steekproefkader elementen bevat die niet tot de doelpopulatie behoren (idem, p. 17). Van sommige respondenten kan dat vastgesteld worden (vastgestelde overdekking), van veel andere echter niet (een groot gedeelte van de weigeringen en 'geen contact').

Voor *personensteekproeven* is er één duidelijke categorie die tot de overdekking gerekend moet worden, nl. wanneer vastgesteld wordt dat de getrokken persoon niet tot de bedoelde populatie behoort. Of men overleden personen wel of niet tot de overdekking rekent, is een punt van discussie. In feite is bepalend of de persoon op het gehanteerde peilmoment ook reeds overleden was. Omdat in het merendeel der gevallen het peilmoment (vrijwel) samenvalt met het moment van dataverzameling, stel ik voor om i.t.t. Bethlehem & Kersten (1986) overleden personen eveneens bij de overdekking te rekenen. Alle andere categorieën (respondent ziek, doet niet open, weigert, op vakantie, taalbarrière, vragenlijst onbruikbaar, verhuisd, etc.) horen tot de non-respons. Dat betekent dus onder meer dat foute adressen en verhuizingen tot de non-respons gerekend worden (de betreffende persoon bestaat wel, maar woont klaarblijkelijk ergens anders), tenzij de respondent elders (in dezelfde gemeente) geïnterviewd kan worden. Het beruchte '3x niet thuis' hoort uiteraard ook tot de non-respons.

Het houdt ook in, dat *alle* getrokken personen meetellen bij het bepalen van de non-respons. De veel toegepaste methode om reserve-adressen (-personen) achter de hand te houden totdat een bepaald aantal interviews is bereikt, is uit statistisch oogpunt niet verantwoord. Kan men hier om welke reden dan ook niet omheen, dan dienen de niet-gebruikte reserve-adressen in elk geval tot de non-respons gerekend te worden.

3. Aanname

Wil men een netto responsfractie berekenen ('response rate', 'uiteindelijke completion rate', etc.), dan moet een schatting gemaakt worden van de overdekking en daarmee van de omvang van de netto steekproef. De aanname van Wiseman & Billington over de verhouding tussen de hoeveelheid personen die wel en die niet tot de doelpopulatie behoren, kunnen we nu als volgt formuleren: de proportie overdekking is dezelfde in de bruto steekproef als in de groep waarvoor deze proportie kon worden waargenomen.

Die aanname leidt tot grotere nauwkeurigheid, in vergelijking met procedures waarbij alle 'onbekenden' óf wel tot de non-respons óf wel tot de kaderfouten gerekend worden. Niettemin kan ook onder deze aanname nog grote onnauwkeurigheid optreden, vooral wanneer er een verband is tussen de overdekking en verschijnselen als 'weigering' en 'geen contact'. Het is daarom van belang om zoveel mogelijk informatie over de omvang van de overdekking te krijgen.

Het is in de praktijk zeer wel mogelijk dat van sommige weigeringen wordt vastgesteld of de respondent al dan niet tot de doelpopulatie behoort. Het eenvoudigste voorbeeld is *seks*, maar ook van warme bakkers kan veelal, ondanks een weigering, vastgesteld worden dat ze tot de doelpopulatie behoren. Zelfs in steekproeven op basis van niet direct waarneembare kenmerken kan mogelijk bij weigeringen wél een antwoord op de cruciale vraag naar het lidmaatschap van de doelpopulatie verkregen worden. De gemaakte veron-

derstelling wordt acceptabeler naarmate van meer personen deze elementaire informatie wordt verzameld.

Het aantal weigeringen dient dus onderverdeeld te worden in weigeringen waarvan verder niets bekend is (en die voor ons doel bij 'geen contact' gerekend kunnen worden), weigeringen die behoren tot de doelpopulatie, en weigeringen die niet behoren tot de doelpopulatie (en die verder tot de waargenomen overdekking worden gerekend).

4. Netto steekproef

De netto steekproef (N) bestaat uit de geslaagde interviews (R), plus het aantal weigeringen e.d. waarvan bekend is dat de betreffende personen tot de doelgroep behoren (W_D), plus een schatting van het aantal potentiële respondenten in de 'onbekende weigeringen' en 'geen contact'. Die laatste schatting wordt gebaseerd op de verhouding tussen de vastgestelde doelgroep en de totale groep waarover iets bekend is:

$$N = R + W_D + \frac{R + W_D}{R + W_D + O_V} W_O \quad (3)$$

waarin:

- N = netto steekproef
- R = respons (aantal geslaagde interviews)
- O_V = vastgestelde overdekking (aantal respondenten waarvan is vastgesteld dat ze niet tot de doelpopulatie behoren)
- W_D = aantal weigeringen, 'niet in staat', etc., waarbij wél is vastgesteld dat de betreffende persoon behoort tot de doelpopulatie
- W_O = weigeringen, geen contact, etc., populatie onbekend

Voorts geldt dat de respons (R) gevonden kan worden door de bruto steekproef (B) te verminderen met de vastgestelde overdekking (O_V), de weigeringen en 'geen contact, etc.' ($W_D + W_O$):

$$R = B - O_V - W_D - W_O$$

waarin: B = bruto steekproef (alle getrokken namen)

of:

$$W_O = B - R - O_V - W_D \quad (4)$$

Substitutie van (4) in (3) levert:

$$N = \frac{R + W_D}{R + W_D + O_V} B \quad (5)$$

5. Responsfractie

De responsfractie kan nu als volgt berekend worden:

$$F_R = \frac{R}{N} = \frac{R (R + W_D + O_V)}{B (R + W_D)} = \frac{R}{B} \cdot \frac{R + W_D + O_V}{R + W_D} \quad (6)$$

of:

$$F_R = \frac{R}{B} \cdot \frac{D_V + O_V}{D_V} \quad (7)$$

waarin:

- F_R = responsfractie
- R = het aantal geslaagde interviews
- B = bruto steekproef (alle getrokken namen)
- D_V = $R + W_D$ = vastgestelde doelpopulatie
- O_V = vastgestelde overdekking
- W_D = aantal weigeringen e.d. behorend tot de doelpopulatie

Wanneer het in het veld onmogelijk is gebleken om van weigeringen en 'niet in staat' vast te stellen of bedoelde personen wel of niet tot de doelpopulatie behoren, dan staat alleen van de respondenten bij wie een geslaagd interview is afgenomen vast dat ze tot de doelpopulatie behoren: $D_V = R$. Dan gaat (7) over in (2), de vereenvoudigde vorm van de formule van Wiseman & Billington.

Voor goede steekproeven, met een geringe overdekking, zijn de verschillen tussen (7) en (2) te verwaarlozen.

6. Discussie

Het gebruik van (2) impliceert dat alle getrokken namen meetellen in de berekening van de responsfractie. Bij vaststelling van overdekking moet dit als zodanig door de interviewer geadministreerd worden; er mag geen vervanging plaatsvinden. Dat betekent enerzijds dat de bruto steekproef niet te groot genomen moet worden, en dat zo mogelijk alle personen uit de bruto steekproef worden benaderd. Laat het budget dat niet toe, dan tellen de niet benaderde personen mee in de berekening van de non-respons, hetgeen zich wrekt bij een

te grote bruto steekproef voor het beoogde aantal interviews. Anderzijds echter wreekt een te kleine bruto steekproef zich nog harder. Een proeftrekking is daarom aan te bevelen. Wordt (7) gebruikt, wat bij een vermoeden van mogelijke overdekking natuurlijk aan te bevelen is, dan moet in het veld ook van weigeringen, 'niet in staat', etc., zoveel mogelijk vastgesteld worden of de betreffende persoon wel of niet tot de doelpopulatie behoort. Ideaal zou zijn om met één vraag aan de weigeraars zowel een antwoord te krijgen op de centrale vraag van het onderzoek (Kersten en Bethlehem, 1983) als op de vraag naar het lidmaatschap van de doelpopulatie. Interviewers moet op het hart gedrukt worden veel (alles?) in het werk te stellen om deze vraag (of twee vragen) beantwoord te krijgen.

Literatuur

- Bethlehem, J.G. & Kersten, H.M.P. (1982). *Non-respons bij enquêtes of 'wat niet weet, wat niet deert?'*. Mens en Maatschappij, 57, 2, pp. 145-170.
- Bethlehem, J.G. & Kersten, H.M.P. (1986). *Werken met non-respons*. Proefschrift UvA.
- CBS (1985). *De leefsituatie van turken en marokkanen in nederland 1984 (deel 1): Eerste uitkomsten*. 's-Gravenhage: Staatsuitgeverij.
- Fluttert, P.H.M. (1980). *Wat is een response van 80% waard?: Een onderzoeksnotitie*. Sociologische Gids, 27, 2, pp. 160-165.
- Kersten, H.M.P. & Bethlehem, J.G. (1983). *Reducing the nonresponse bias by asking the basic question*. Voorburg: CBS.
- Veldkamp b.v. (1986). *Media + minderheden (deel 1): Tekstrapport*. Amsterdam: Veldkamp b.v.; Hilversum: NOS-tekstservice.
- Wentholt, H. (1983). Survey-onderzoek onder buitenlandse werknemers, kan dat wel? In: A.E. Bronner, M.T.G. Meulenberg, e.a. (1983). *Jaarboek van de Nederlandse Vereniging van Marktonderzoekers*. Amsterdam: NvM.
- Wiseman, F. & Billington, M. (1984). Comment on a standard definition of response rates. *Journal of Marketing Research*, 21, pp. 336-338.