

Nonrespons en representativiteit in panel studies: Een korte notitie*

Toon W. Taris

Kurt Lewin Instituut/Vrije Universiteit Amsterdam

Vakgroep Sociale Psychologie

De Boelelaan 1081

1081 HV Amsterdam

Telefoon: 020-4448865

E-mail: V73UTOON @ HASARA11.BITNET

* Dit onderzoek werd gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk onderzoek, nummer PGS 56-381. De auteur dankt de anonieme referee van *Kwantitatieve Methoden* voor diens opbouwende kritiek.

INTRODUCTIE

Uit overzichtsstudies (Goyder, 1987; Hox & De Leeuw, 1993) blijkt dat de nonrespons in survey-onderzoek vaak 30-40 procent bedraagt. Dit is een groot probleem omdat het gevaar dat deze nonrespons *niet random* ofwel *selectief* is levensgroot aanwezig is: bepaalde groepen in de doelpopulatie zijn dan onder- of oververtegenwoordigd in de steekproef. Van de Pol (1989) laat in een overzicht zien dat inwoners van grote steden, mannen, en ouderen vaak zijn ondervertegenwoordigd. Conclusies die zijn gebaseerd op een dergelijke vertekende steekproef zijn niet generaliseerbaar naar de populatie, waardoor nonrespons kan leiden tot grote problemen.

Dit geldt nog sterker voor longitudinaal onderzoek, waar nonrespons vaak cumulatief werkt (personen die in een bepaalde meting niet participeren doen meestal ook niet mee in latere metingen). Als de responswaarschijnlijkheid R de kans is dat een subject participeert in het onderzoek en w het aantal metingen is, participeert -- als R constant is -- na w metingen nog de fractie R^w . Als R bijvoorbeeld .8 is en $w = 4$, participeert na 4 metingen nog 41 procent in de studie. Een dergelijke uitval is niet ongevoel (zie Van de Pol, 1989).

R is niet noodzakelijkerwijs constant voor verschillende metingen. Vaak wordt geconstateerd dat de nonrespons lager wordt voor elke opeenvolgende meting. Taris, Van der Vaart en Dijkstra (1993) rapporteren voor een panelstudie met drie metingen bijvoorbeeld een uitval van respectievelijk 37, 20 en 11% voor elke volgende meting. Een dergelijke dalende nonrespons is echter *niet* het goede teken waarvoor het soms wordt gehouden; het is juist een aanwijzing dat de steekproef minder representatief wordt voor de doelpopulatie vanwege selectieve nonrespons. Hieronder wordt kort ingegaan op dit fenomeen. Aan de hand van een voorbeeld worden de implicaties van selectieve nonrespons besproken, en wordt ingegaan op de consequenties hiervan in (de rapportage van) onderzoek.

VOORBEELD

Een populatie bestaat uit twee even grote groepen die verschillen wat betreft kans R dat zij zullen meedoen in onderzoek. Groep A bestaat uit oudere, weinig extraverte inwoners van grote steden (kans R_a dat zij zullen meedoen is .6), terwijl groep B juist bestaat uit jongere, sociaal vaardige bewoners van voorsteden (responswaarschijnlijkheid R_b is .9). Neem aan dat beide groepen even vaak benaderd worden, en dat de steekproef moet bestaan uit 1000 personen. Hoeveel personen worden dan benaderd, en hoe zal -- als R_a en R_b constant zijn -- de steekproef na een aantal metingen zijn samengesteld?

Het aantal personen dat benaderd moet worden om een steekproef van 1000 personen te verkrijgen kan berekend worden door het vergelijkingstelsel $R_a A + R_b B = 1000$, en $A = B$ op te lossen; in totaal zullen dus 1334 personen benaderd worden. Van de 667 benaderde leden

van groep B zullen 600 personen meedoen; van de 667 personen in groep A echter slechts 400 personen. Bij de eerste meting van onze studie is de A/B-ratio dus .67: voor elk lid van groep A doet 1.5 lid van groep B mee. Tabel 1 laat zien hoe deze getallen veranderen voor een panelstudie met 4 metingen. Groep B is met elke volgende meting meer dominant aanwezig.

Tabel 1: Hoe verschillende responsewaarschijnlijkheden de samenstelling van de steekproef beïnvloeden, 4 metingen, $R_a = .6$ en $R_b = .9$.

	meting 1	meting 2	meting 3	meting 4
groep A	400	240	144	87
groep B	600	540	486	438
totale steekproef	1000	780	630	525
ratio B/A	.67	.44	.30	.20

Het interessante gegeven is echter dat de nonrespons voor elke meting, vergeleken met de voorgaande meting, afneemt. De nonrespons op tijdstip 1 is 25%; voor de volgende metingen is dit respectievelijk .22, .19, en .17. Dit is te danken aan het feit dat in groep B -- die immers een steeds groter deel van de steekproef vormt -- de uitval lager is.

Vergelijk nu de trend in tabel 1 met de situatie die zou ontstaan als de responswaarschijnlijkheid voor alle subjecten gelijk zou zijn, stel, $R = .75$. Om een steekproef van 1000 personen te verkrijgen zouden we ook hier 1334 personen moeten benaderen. Bij de tweede meting participeren nog 750 personen, op tijdstip 3 563, en op tijdstip 4 doen nog 423 personen mee. Vergelijking van deze resultaten met tabel 1 laat zien dat de totale nonrespons in de heterogene steekproef afneemt, dat dit leidt tot een grotere steekproef, die echter minder representatief is voor de populatie dan de homogene steekproef. Met andere woorden, een afnemende nonrespons is een teken dat de representativiteit van de steekproef geschonden is: de paradox is hier dat een kleinere nonrespons leidt tot een *minder* representatieve steekproef, met alle problemen van dien!

De ernst van dit probleem is afhankelijk van het verschil tussen R_a en R_b . Echter, ook bij een gering verschil in antwoordwaarschijnlijkheden van .1 (bijvoorbeeld $R_a = .7$, $R_b = .8$) leidt selectieve nonrespons tot een oververtegenwoordiging van groep B ten opzichte van groep A van 3.16 op tijdstip 4. De percentages nonrespons zijn dan respectievelijk .43, .29, .27 en .26. Een klein verschil in responswaarschijnlijkheden kan dus grote gevolgen hebben.

DISCUSSIE

Bovenstaande resultaten geven aan dat het bestuderen van de nonrespons per meting de onderzoeker in staat stelt een ruwe indicatie te verkrijgen van de mate waarin selectieve

nonrespons is opgetreden. Een *dalende* nonrespons betekent dat de conclusies zeer voorzichtig moeten zijn, omdat generaliseerbaarheid naar de populatie niet gewaarborgd is; verder onderzoek naar de selectiviteit van de nonrespons is aan te raden. Een *constante* of *stijgende* nonrespons geeft daarentegen weinig reden tot zorg. Gegeven de ernstige gevolgen van selectieve nonrespons lijkt het niet teveel gevraagd om onderzoekers te vragen om bij de rapportage van hun onderzoek aan te geven wat de nonrespons was, uitgesplitst per meting.

Indien er sprake is van dalende nonrespons, is het mogelijk om na te gaan of er variabelen gevonden kunnen worden die deze daling verklaren. Stel dat de steekproef gestratificeerd wordt naar sekse. Als de nonrespons voor mannen en vrouwen *afzonderlijk* voor elke opvolgende meting gelijk is, heeft men het mechanisme achter de dalende nonrespons ontdekt. Bij vervolganalyses kan dan worden volstaan met het controleren voor sekse, om vertekende parameterschattingen te voorkomen. Als aangenomen kan worden dat de multivariate verdeling van de doelvariabelen voor participanten en niet-participanten gelijk is, kan volstaan worden met weging van de steekproef (Bethlehem & kersten, 1985).

Afsluitend: onze resultaten illustreren eens te meer dat het van het grootste belang is om nonrespons zoveel mogelijk te voorkomen. Dit geldt in nog versterkte mate voor longitudinaal onderzoek, omdat de gevolgen van selectieve nonrespons cumulatief werken en een steekproef al snel niet meer als representatief voor de doelpopulatie gezien kan worden. Daarnaast is het van belang dat bij heterogene groepen selectieve uitval tegengegaan wordt. Wanneer er aanwijzingen zijn dat bepaalde groepen eerder uitvallen dan anderen, dan kan men wellicht de veldwerkstrategieën zo aanpassen, dat er voor die specifieke groepen extra moeite gedaan wordt (bijvoorbeeld een speciale benadering toepassen).

REFERENTIES

- Bethlehem, J.G. & Kersten, H.P.M. (1985). On the treatment of nonresponse in sample studies. *Journal of Official Statistics*, 1, 287-300.
- Goyder, J. (1987). *The silent minority*. Cambridge: Blackwell University Press.
- Hox, J.J. & De Leeuw, E.D. (1993). Nonrespons bij schriftelijke vragenlijsten, telefonische interviews en face-to-face interviews: een vergelijkende meta-analyse. *Kwantitatieve Methoden*, 14, 59-78.
- Taris, T.W., Van der Vaart, W. & Dijkstra, W.D. (1993). Nonrespons en de representativiteit van de follow-up metingen. In W.D. Dijkstra (ed.), *Het proces van sociale integratie van jong-volwassenen* (pp. 65-76). Amsterdam: VU-Uitgeverij.
- Van de Pol, F.J.R. (1989). *Issues of design and analysis of panels*. Amsterdam: Sociometric Research Foundation.