

# *Nauwkeurigheid van gerapporteerde giften aan goededoelenorganisaties*

René Bekkers (VU)<sup>1)</sup>

## *1. Het probleem van onnauwkeurigheid*

Vanzelfsprekend is de waarde van onderzoek groter naarmate het nauwkeuriger schattingen oplevert. Dit geldt ook voor onderzoek op basis van enquêtes, dat in de sociale wetenschappen in de afgelopen decennia de standaard is geworden. Een probleem dat al snel optreedt wanneer respondenten in enquêtes niet nauwkeurig rapporteren is dat generalisatie naar de doelpopulatie systematische fouten bevat. Een bekende bron van fouten betreft de steekproefsamenstelling: wanneer de steekproef onevenredig veel respondenten bevat die systematisch te hoog (of te laag) scoren op het verschijnsel waarvan de populatiewaarde geschat moet worden, zal die schatting systematisch te hoog (of te laag) uitvallen.

Veel onderzoek naar nauwkeurigheid van rapportages gaat over onwenselijk gedrag, zoals seksueel misbruik en kindermishandeling (Hardt & Rutter, 2004), het gebruik van alcohol (Midanik, 1982) en drugs (Harrell, 1997) en over criminaliteit (Thornberry & Krohn, 2004). Er is ook veel onderzoek naar gezondheidsklachten (zie bijvoorbeeld Hahn et al., 1997). In deze onderzoeksgebieden is een nauwkeurige rapportage van belang voor de optimale behandeling van slachtoffers en/of daders. Minder onderzoek is er gedaan naar de nauwkeurigheid van sociaalwenselijk gedrag, zoals het stemmen bij verkiezingen. Dit is begrijpelijk omdat er door ongedige conclusies over sociaalwenselijk gedrag als gevolg van onnauwkeurigheid geen mensen overlijden en er geen hoge kosten mee gemoeid zijn.

Maar er is nog een andere reden om bezorgd te zijn over onnauwkeurigheid in rapportages door respondenten in enquêtes. Als enquêtegegevens gebruikt worden om theorieën te toetsen over verbanden tussen verschillende sociale verschijnselen kunnen onnauwkeurige rapportages tot ongedige conclusies over deze verbanden leiden. Deze conclusies kunnen net als de rapportages systematische fouten bevatten. In het onderzoek naar stemgedrag is gesteld dat het gebruik van zelfrapportages 'common sense' theorieën over de samenleving ondersteunt (Presser & Traugott, 1992).

De algemene veronderstelling in de literatuur over de effecten van onnauwkeurigheid op de geldigheid van conclusies luidt dat vertekening in zelfrapportages problematischer is naarmate de vertekening een sterkere samenhang vertoont met het verschijnsel dat wordt bestudeerd (Katosh & Traugott, 1981; Presser, 1984;

Phillips & Clancy, 1972; Bernstein, Chadha, & Montjoy, 2001; Presser & Traugott, 1992). Een voorbeeld op het terrein waarop ook de analyses later in dit artikel zich begeven maakt deze veronderstelling duidelijk. De hoogte van bestedingen door huishoudens neemt toe met het inkomen. Maar wanneer leden van huishoudens met een hoger inkomen in enquêtes minder nauwkeurig zijn in hun rapportage van bestedingen en de nauwkeurigheid afneemt naarmate het inkomen stijgt, zal het verband tussen het huishoudinkomen en de uitgaven onderschat worden. In ons geval zou het verband tussen het inkomen en giften aan goededoelenorganisaties op basis van enquêtes zwakker lijken dan het in werkelijkheid is.

## 2. *Bronnen van vertekening*

De wijze waarop respondenten in enquêtes rapporteren over verschijnselen waarover zij informatie hebben kan op een vruchtbare manier worden bestudeerd vanuit sociaalpsychologische theorieën over sociaal gedrag (Tourangeau, Rips & Rasinski, 2000). In deze theorieën wordt menselijk gedrag gezien als een resultaat van het nastreven van doelen, gegeven de moeite die dat kost en de mogelijkheden die de situatie daarvoor biedt. In deze bijdrage ontbreekt de ruimte om een algemeen theoretisch kader voor rapportages in enquêtes verder uit te werken. Ik volsta daarom met een kort overzicht van de literatuur op de meest relevante onderzoeksgebieden.

Respondenten geven in een interviewsituatie meer nauwkeurige antwoorden naarmate zij daartoe sterker gemotiveerd zijn, bijvoorbeeld omdat zij het onderwerp interessanter vinden. Respondenten kunnen in de interviewsituatie echter ook andere doelen nastreven. Respondenten vinden het ook belangrijk dat zij door anderen positief beoordeeld worden en stemmen hun antwoorden af op wat zij verwachten dat anderen positief waarderen (wat de sociale norm is). Tenslotte vinden respondenten het prettig een positief beeld van zichzelf te geven omdat dit bijdraagt aan hun psychologisch welbevinden.

Respondenten geven in een enquête meer nauwkeurige antwoorden naarmate de interviewsituatie dat gemakkelijker maakt. Wanneer respondenten door de manier waarop de vragen gesteld worden meer geholpen worden nauwkeurige antwoorden te geven zullen zij nauwkeuriger zijn. De doelen die respondenten in interviewsituaties hebben komen in veel gevallen tot uiting in sociaalwenselijke antwoorden: zij doen zichzelf bijvoorbeeld voor als minder crimineel, gezonder, sterker betrokken bij de politiek en vrijgevinger dan zij in werkelijkheid zijn. Sociaalwenselijke antwoorden worden daardoor minder vaak gegeven naarmate het onderwerp van de enquête meer normatief geladen is (Tourangeau, Rips & Rasinski, 2000).

Sociaalwenselijke antwoorden worden eveneens minder vaak gegeven wanneer de respondenten die niet voldoen aan de norm zonder gezichtsverlies een antwoord-

categorie kunnen kiezen waardoor ze niet hoeven te liegen (Belli, Traugott, Young & McGonagle, 1999), en er geen mensen bij het invullen van de enquête aanwezig zijn die sociaalwenselijke antwoorden kunnen goedkeuren (Tourangeau & Smith, 1996; Tourangeau, Jobe, Pratt, and Rasinski, 1997; Stocké, 2007a). Ook geven respondenten vaker nauwkeurige antwoorden wanneer het om saillante gebeurtenissen gaat (Sudman & Bradburn, 1973).

Naar aanleiding van het klassieke werk van Crowne & Marlowe (1960, 1964) is er veel onderzoek gedaan naar de mogelijkheid dat de tendens om sociaalwenselijke antwoorden te geven een stabiele eigenschap van respondenten is. De consensus is intussen echter dat het geven van sociaalwenselijke antwoorden sterk afhankelijk is van de situatie. Hoewel sociaalwenselijk gedrag wel een zekere mate van consistentie vertoont (zie bijvoorbeeld Wilhelm & Bekkers, 2010), hangt het geven van sociaalwenselijke antwoorden op verschillende soorten vragen nauwelijks met elkaar samen (Presser, 1984; Clark & Tiff, 1966). De conventionele meetinstrumenten voor sociale wenselijkheid zijn eerder een maat voor daadwerkelijk sociaalwenselijk gedrag (Barger, 2002; McCrae & Costa, 1983; Mills & Kroner, 2005) en hangen niet samen met de nauwkeurigheid van rapportages in enquêtes (Johnson, O'Rourke, Burris & Warnecke, 2005). Overigens bleek in een Nederlandse studie waarin giften aan KWF Kankerbestrijding zoals gerapporteerd in een enquête werden vergeleken met de organisatie geregistreerde giften dat respondenten die het vaker eens waren met de stelling 'Ik doe alles om het anderen naar de zin te maken' hun giften sterker overdreven (Bekkers & Wiepking, 2010).

De zoektocht naar andere kenmerken van respondenten die samenhangen met de kans op sociaalwenselijke antwoorden is tot nu toe eveneens vrij onvruchtbaar gebleken (Presser & Stinson, 1998). Sommige studies vinden dat het opleidingsniveau positief samenhangt met de nauwkeurigheid van rapportages. Dit betreft gezondheidsgelateerde verschijnselen (zie bijvoorbeeld Hahn et al., 1997) en de recentere studies naar stemgedrag (Silver, Anderson & Abrahamson, 1996). In studies naar stemgedrag van voor 1980 wordt nauwelijks een verband met het opleidingsniveau aangetroffen (Kathosh & Traugott, 1981). In de klassieke Denver-validatiestudie uit de jaren '40 werd een positief verband gevonden tussen het opleidingsniveau en de nauwkeurigheid van gerapporteerd geefgedrag (Cahalan 1968). In de meer recente KWF-validatiestudie werd echter een negatief verband gevonden (Bekkers & Wiepking, 2010): hoger opgeleiden overdreven hun giften meer dan lager opgeleiden.

In het empirische gedeelte van dit artikel onderzoek ik de nauwkeurigheid van door respondenten in een enquête gerapporteerde giften. Onderzoekers in het vakgebied van filantropische studies (voor een overzicht van de literatuur zie Bekkers & Wiepking, 2007) gebruiken in de regel enquêtes om giften vast te stellen bij gebrek aan registerdata op dit gebied. Sinds het midden van de jaren '90 wordt het volume en de aard van de filantropische geldstromen in Nederland vastgesteld met gegevens uit 'Geven in Nederland' enquêtes die om het jaar georganiseerd worden door de Vrije Universiteit. Op basis van deze gegevens

wordt geschat dat het totale volume aan donaties aan goededoelenorganisaties in Nederland in 2007 € 1 945 miljard euro bedraagt (Schuyt, Gouwenberg & Bekkers, 2009).

Twee studies zijn in het buitenland verricht naar de nauwkeurigheid van gerapporteerde giften. In de Denver-validiteitstudie bleek dat 34% van de gerapporteerde giften niet daadwerkelijk waren ontvangen (Parry & Crossley, 1950). Burt en Popple (1998) vonden in een onderzoek onder studenten in Australië een overschatting van giften met 30% wanneer de studenten niet wisten dat hun antwoorden gecontroleerd zouden worden. Wanneer ze dat wel wisten was de overschatting veel kleiner.

In Nederland is eerder met twee registraties nagegaan wat de nauwkeurigheid is van rapportages in enquêtes over geefgedrag. In beide studies betreft het enquêtegegevens uit het Geven in Nederland-onderzoek. Een vergelijking van giften aan de R.K.-Kerk op macro-niveau heeft laten zien dat in het Geven in Nederland-onderzoek de giften met zo'n 22% worden onderschat (Schuyt, Gouwenberg & Bekkers, 2009, p. 206). Een vergelijking van giften aan KWF Kankerbestrijding op individueel niveau heeft laten zien dat de giften met ongeveer 30% werden overdreven door respondenten die in werkelijkheid giften hebben gedaan en die ook hadden gerapporteerd (*true positives*). *False positives* (gerapporteerde giften die in werkelijkheid niet zijn gedaan) en *false negatives* (giften die zijn gedaan maar niet zijn gerapporteerd) konden door de opzet van het onderzoek niet worden waargenomen.

### 3. *Data en methoden*

#### 3.1 *Schatting van vertekening*

Om de mate van vertekening te schatten is een vergelijking van diverse bronnen van gegevens nodig. In dit artikel vergelijk ik de giften aan Greenpeace Nederland zoals gerapporteerd door respondenten in een enquête met de giften afkomstig van de huisadressen waarop deze respondenten woonachtig zijn zoals die zijn geregistreerd door Greenpeace. Uitgedrukt in de totale baten uit eigen fondsenwerving is Greenpeace een van de tien grootste goededoelenorganisaties in Nederland (CBF, 2009). Deze organisatie is gekozen omdat alle giften ook geregistreerd worden in een database. Een probleem in de eerdere KWF-validatiestudie (Bekkers & Wiepking, 2010) is dat niet alle giften aan KWF ook geregistreerd worden, waardoor het aantal *false positives* en de mate van vertekening niet kan worden vastgesteld. KWF krijgt veel kleine giften binnen via de huis-aan-huis-collecte. Deze giften worden in de regel niet afzonderlijk op huisadres geregistreerd. Greenpeace houdt geen huis-aan-huis-collecte maar krijgt uitsluitend giften per bank/giro binnen. Van de giften is daardoor wel bekend van welk huisadres ze afkomstig zijn. Hierdoor kunnen ook *false positives* worden vastgesteld.

### 3.2 *Enquêtegegevens*

De enquêtegegevens die in dit artikel worden gebruikt zijn afkomstig uit het Nederlands Donor Panel (NDP), een online marktonderzoek naar vertrouwen in goededoelenorganisaties en geefgedrag dat elk kwartaal wordt gehouden door WWAV, een bureau voor non-profit marketing en communicatie (WWAV, 2009). NDP-respondenten worden geworven via e-mail uit de e-mailadressen van consumenten die eerder hebben deelgenomen aan de 'Grote Consumenten Enquête' (GCE). De GCE is een tweejaarlijks *mixed-mode* consumentenonderzoek naar de bestedingen en voorkeuren van huishoudens. De vragenlijst wordt schriftelijk verspreid onder alle huishoudens in Nederland, maar kan ook *online* worden ingevuld. Aan elke editie van het NDP nemen zo'n 200 000 respondenten deel. Uitnodigingen voor het NDP worden gestuurd naar toevallige selecties van GCE-respondenten van wie een geldig e-mail adres bekend is.

In de hierna volgende analyses maak ik gebruik van de gegevens uit de NDP editie van mei 2007. In totaal werden uitnodigingen om 'deel te nemen aan een onderzoek over goede doelen' verstuurd naar 60 000 e-mailadressen. Zo'n 2 300 respondenten vulden de enquête volledig in. De respons van 3,8% is laag vergeleken met typische respons percentages voor persoonlijke interviews of schriftelijke vragenlijsten die via reguliere post worden verstuurd (Stoop, 2005). Ook is de respons niet representatief voor de gehele Nederlandse bevolking. De respondenten waren vaker vrouw (57%), hadden minder vaak betaald werk (27% tegenover 74% volgens CBS Statline), hadden minder vaak een HBO- of universitair diploma (8% tegenover 24% volgens CBS Statline), en zeiden vaker dat ze bloed hadden gegeven (8% tegenover 3%) en gestemd hadden bij de afgelopen Tweede Kamerverkiezingen (88% tegenover 80%). Het doel van dit artikel is echter om meetfouten te schatten, en niet om steekproeffouten te onderzoeken. Het is de vraag of de respondenten in de enquête nauwkeurig hebben gerapporteerd, en niet of de steekproef representatief is. Bij de interpretatie van de resultaten is het wel van belang om de samenstelling van de steekproef in het achterhoofd te houden. Het is mogelijk dat juist de meer nauwkeurigere respondenten vaker aan het onderzoek hebben meegedaan.

In de NDP-enquête zijn vragen opgenomen over giften aan goededoelenorganisaties, die zijn overgenomen uit de 'Geven in Nederland' (GIN) vragenlijst (Bekkers & Wiepking, 2010). In de NDP-enquête zeiden 1,463 respondenten (64%) dat (iemand in) het huishouden in het afgelopen jaar geld heeft gegeven aan goededoelenorganisaties via bank/giro of internet. Deze respondenten kregen vervolgens een lijst te zien met organisaties met de vraag om aan te geven aan welke van deze organisaties (iemand in) het huishouden had gegeven. Wanneer respondenten een organisatie aankruisten kregen zij de vervolgvraag hoeveel ze hadden gegeven in het afgelopen jaar. In de NDP-enquête werd in tegenstelling tot in het GIN-onderzoek geen 'weet niet' antwoordcategorie aangeboden bij de vraag naar het bedrag. Omdat 'weet niet' een antwoord is waarmee respondenten die onterecht hebben beweerd gegeven te hebben niet hoeven te liegen kan de mate van sociale wense-

lijkheid in het NDP hoger zijn dan in GIN. Het NDP is net als GIN een *online* vragenlijst, waardoor sociaalwenselijke antwoorden verminderd zouden moeten worden omdat er geen interviewer aanwezig is.

### 3.3 *Koppeling*

In beide datasets waren gegevens opgenomen over huisadressen in de vorm van de postcode (4 cijfers en 2 letters) en het huisnummer. Aan Greenpeace Nederland werden alle postcodehuisnummercombinaties van respondenten in het NDP ter beschikking gesteld. De datamanager van Greenpeace verrichte een koppeling op basis van de unieke postcodehuisnummercombinaties, die 322 geslaagde koppelingen opleverde. Greenpeace zette alle giften vanuit deze adressen uit de periode van 1 mei 2006 tot en met 31 mei 2007 uit de database in een apart bestand, dat de auteur vervolgens heeft gekoppeld aan de NDP-data.

## 4. *Resultaten*

### 4.1 *Totale nauwkeurigheid*

In tabel 1 is de vertekening te zien in de rapportages over de giften door de NDP-respondenten. Door 85% van de respondenten die aan de enquête deelnamen is geen gift gerapporteerd en ook geen gift gedaan. Van iets meer dan 7% van de respondenten werd een gift ontvangen die bovendien werd gerapporteerd. Iets meer dan 4% van de respondenten rapporteert giften die niet zijn ontvangen (*false positives*). Van iets minder dan 4% zijn wel giften ontvangen die niet zijn gerapporteerd. Dit betekent dat 35% van alle geregistreerde giften niet werd gerapporteerd door de respondenten. Van de giften die werden gerapporteerd is 38% niet geregistreerd.

Opvallend genoeg rapporteren respondenten van wie geen giften zijn geregistreerd (*false positives*) lagere giften (€ 21,01) dan respondenten die giften rapporteren die wel in de database zijn opgenomen (*true positives*, € 28,22). Dit verschil is sterk significant ( $F=10\ 471$ ,  $p<.001$ ). Met andere woorden: als respondenten giften rapporteren die in werkelijkheid niet zijn gedaan zijn dat relatief kleine leugens, wellicht als gevolg van de behoefte om sociaalwenselijk te antwoorden (Presser & Traugott, 1992).

Donaties die wel zijn geregistreerd maar niet gerapporteerd werden (*false negatives*: € 25,68) zijn wat lager dan de *true positives*. Dit verschil ondersteunt de saillantie hypothese die voorspelt dat kleinere giften gemakkelijker worden vergeten (Bekkers & Wiepking, 2010). Het verschil is echter niet erg groot en niet significant ( $F=1\ 773$ ,  $p<.183$ ). Onder de respondenten die daadwerkelijk giften hadden gedaan en dat ook gerapporteerd hadden waren de geregistreerde giften zo'n 14% lager dan de gerapporteerde giften. Dit verschil is significant in een gepaarde t-toets ( $t=2\ 181$ ,  $p<.031$ ).

**Tabel 1**  
Vertekening in gerapporteerde giften

	Gift niet gerapporteerd	Gift wel gerapporteerd
Gift niet geregistreerd	'True negatives' 84,6% (n=1945)	'False positives' 4,4% € 21,01 gerapporteerd (n=101)
Gift wel geregistreerd	'False negatives' 3,9% (n=89) € 25,68 geregistreerd	'True positives' 7,2% € 32,88 gerapporteerd (n=165) € 28,22 geregistreerd

Wanneer alle positieve waarnemingen van zowel gerapporteerde als geregistreerde giften worden gebruikt (alle gevallen behalve de *true negatives*) blijkt het grootste negatieve verschil tussen gerapporteerde en geregistreerde giften een onderrapportage van € 145 te zijn. Het grootste positieve verschil is een overrapportage van € 155. Het gemiddelde verschil is € 0,26. De mediaan en de modus van het verschil tussen gerapporteerde en geregistreerde giften zijn allebei € 0.

#### 4.2 *Wie overdrijft het geefgedrag?*

Welke kenmerken van respondenten gaan samen met verschillen tussen de gerapporteerde en geregistreerde giften? Er werden drie groepen onderscheiden van ongeveer gelijke omvang: respondenten die giften rapporteerden die meer dan € 5 onder de geregistreerde giften lagen (de onderrapporteerders; 33,2%), respondenten die giften rapporteerden die meer dan € 5 boven de geregistreerde giften lagen (de overrapporteerders, 32,1%), en de respondenten die een bedrag rapporteerden dat minder dan € 5 van het geregistreerde bedrag lag (33,5%). Zijn onderrapporteerders het spiegelbeeld van overrapporteerders of zijn zij een wezenlijk andere categorie van respondenten, met geheel eigen kenmerken? Deze vraag wordt beantwoord in tabel 2, waarin een multinomiale logistische regressie-analyse van de richting van vertekening wordt gepresenteerd. De referentiecategorie wordt gevormd door de respondenten die het bedrag ongeveer goed (+/- € 5) hadden gerapporteerd. De resultaten in tabel 2 laten zien dat de kenmerken van overrapporteerders en onderrapporteerders nogal verschillend zijn. Voor zes kenmerken geldt dat het teken van de variabele in de regressievergelijking verschilt tussen de twee contrasten. Voor geen enkele van deze kenmerken zijn de parameters echter significant. voor zeven variabelen is het teken wel hetzelfde in beide contrasten. Twee hiervan zijn significant (in beide contrasten). Dit zijn het stemmen bij de vorige verkiezingen en het vertrouwen in goededoelenorganisaties.

Respondenten die ouder zijn dan 65 of van wie de leeftijd onbekend is rapporteren vaker lager dan de geregistreerde bedragen, terwijl tweeverdieners juist vaker het bedrag goed hebben in plaats van onderrapporteren. Kerkgangers en responden-

**Tabel 2**  
**Multinomiale logistische regressie van de richting van vertekening in gerapporteerde giften**

	onderrapportage		overrapportage	
	<i>odds ratio</i>			
Vrouw	0,90	0,89	1,21	1,18
Ouder dan 65	2,15	2,51	1,28	1,48
Leeftijd onbekend	2,54*	2,64*	1,02	1,13
Tweeverdiener	0,51(*)	0,55(*)	1,02	1,09
Geen betaald werk	0,62	0,54	0,91	0,75
Opleidingsniveau	1,06	1,11	1,08	1,13
Partner in huishouden	0,69	0,74	1,67	1,71
Inkomen (log)	0,71	0,74	1,47	1,10
Eigen huis	1,33	1,35	0,90	0,92
Kerkgang	0,76	0,84	2,02*	2,30*
Gestemd bij laatste verkiezingen		0,24*		0,30(*)
Bloeddonor		0,70		0,58
Vrijwilliger		1,21		0,74
Totaal gegeven bedrag (log)''		1,12		1,80***
Vertrouwen in goededoelenorganisaties		0,58**		0,63*

\*\*\* p<.001; \*\* p<.01; \* p<.05; (\*) p<.10.

Referentie categorie: rapportage minder dan € 5 van geregistreerde bedrag.

ten die ook hogere giften aan andere goededoelenorganisaties rapporteerden gaven vaker te hoge giften aan Greenpeace op. Respondenten die rapporteren dat ze hebben gestemd bij de afgelopen verkiezingen en respondenten die een hoger vertrouwen in goededoelenorganisaties zeggen te hebben rapporteren vaker hun giften accuraat in plaats van dat ze de giften over- of onderdrijven. Dit geldt ook voor respondenten die zeggen geen betaald werk te hebben en respondenten die zeggen bloed te geven, hoewel deze variabelen de conventionele significantieniveaus (net) niet bereiken. Opvallend is dat geslacht, het opleidingsniveau, inkomen en eigen huisbezit geen significant verband vertonen met de richting van vertekening. In tabel 3 wordt de analyse herhaald voor de hoogte van het verschil tussen gerapporteerde en geregistreerde giften. De afhankelijke variabele in deze analyse is het gerapporteerde bedrag verminderd met het geregistreerde bedrag. Hoe hoger de waarde van deze variabele, hoe hoger de overdrijving is. Negatieve waarden op de variabele duiden juist op een grotere onderrapportage. De resultaten laten zien dat kerkelijke respondenten, in deeltijd werkenden en tweeverdieners grotere verschillen laten zien, terwijl respondenten die zeggen vrijwilligerswerk te doen en respondenten van wie de leeftijd onbekend is juist sterker onderrapporteren. Leeftijd, geslacht, opleiding, inkomen en huisbezit vertonen geen significante verbanden met de mate van overdrijving.

## 5. Conclusie en discussie

De vergelijking van giften aan Greenpeace zoals gerapporteerd in een consumentenenquête en geregistreerd door de organisatie laat zien dat er zowel respondenten zijn die giften noemen die niet zijn geregistreerd als respondenten die giften die



**Tabel 3**  
**Regressie van het verschil tussen gerapporteerde en geregistreerde giften**

	B	B
Vrouw	-1,0	-2,0
Ouder dan 65	-2,7	-2,3
Leeftijd onbekend	-10,0(*)	-9,5(*)
Tweeverdiener	7,9(*)	7,4(*)
Parttime werk	11,8*	12,0*
Opleidingsniveau	-0,5	-0,6
Partner in huishouden	7,9	6,3
Inkomen (log)	5,1	1,0
Eigen huis	-4,3	-4,4
Kerkgang	9,9*	10,3*
For-profit dienstensector	5,1	5,9
Gestemd bij laatste verkiezingen		-8,2
Bloeddonor		1,8
Vrijwilliger		-8,0*
Totaal gegeven bedrag (log)		5,7**
Vertrouwen in goededoelenorganisaties		2,3
Constante	-18,9	-23,1
Adj. R <sup>2</sup>	,028	,058

\*\*\* p<.001; \*\* p<.01; \* p<.05; (\*) p<.10.

zijn geregistreerd verzuimen te rapporteren. Sommige respondenten die giften rapporteren overdrijven terwijl andere respondenten juist onderrapporteren. De totale vertekening is nagenoeg nul doordat vergeten en onderrapporteren en het onterecht claimen en overdrijven elkaar nagenoeg geheel opheffen. Dit resultaat is opmerkelijk gezien de resultaten van eerdere studies naar de nauwkeurigheid van gerapporteerde giften en ander sociaalwenselijk gedrag, waarin 'virtually all the error is in the socially desirable direction' (Presser & Traugott, 1992).

Het is moeilijk een definitieve verklaring te geven voor het lage niveau van sociaalwenselijke antwoorden in de huidige analyses. In dit onderzoek gaat het over een andere organisatie en een andere bron van enquêtegegevens, met een andere steekproef dan in eerder onderzoek. De steekproefsamenstelling kan gedeeltelijk een verklaring bieden. De categorieën van respondenten die in de enquête die hier is gebruikt oververtegenwoordigd zijn (vrouwen, lager- en middelbaar opgeleiden, mensen zonder betaald werk) zijn volgens de resultaten van de analyses iets beter in het rapporteren van giften dan de categorieën die ondervertegenwoordigd zijn. De verbanden zijn echter niet groot en evenmin significant. Een andere gedeeltelijke verklaring is dat in voorgaande studies niet altijd de *false negatives* zijn meegeteld. In studies waarin dat wel is gedaan (bijvoorbeeld Burt & Popple, 1998) wordt sociaalwenselijk gedrag echter doorgaans nog steeds overdreven. Een volledige verklaring vormen de steekproefsamenstelling en het meenemen van de *false negatives* dus niet. Hoe dan ook, het lijkt wel veilig om te concluderen dat de uitgebreide *online* vragenlijst het niveau van sociaalwenselijke antwoorden in ieder geval niet heeft vergroot.

Natuurlijk kent ook de opzet van het hier gepresenteerde onderzoek problemen. De respons is erg laag door het gebruik van uitnodigingen per e-mail. Ook is de respons erg selectief, vooral voor wat betreft de arbeidssituatie en het opleidings-

niveau. Wellicht ook hebben respondenten die beter op de hoogte zijn van hun giften aan goededoelenorganisaties de enquête vaker ingevuld. Een ander probleem is dat respondenten 'het afgelopen jaar' geïnterpreteerd kunnen hebben als 'het afgelopen kalenderjaar'. Hierdoor kan de totale nauwkeurigheid onderschat zijn. Het is ook mogelijk dat respondenten sinds de giften zijn gedaan zijn verhuisd, waardoor de nauwkeurigheid eveneens wordt onderschat.

Het resultaat dat de gerapporteerde donaties lager zijn dan geregistreerde donaties onder de respondenten die zowel in de enquête als in de registratiedata voorkomen komt wel overeen met de resultaten van eerder onderzoek (Bekkers & Wiepking, 2010; Burt & Popple, 1998). De overdrijving die nu gevonden is (14%) ligt echter lager dan de overdrijving in eerdere studies (30%). Hoewel onduidelijk is waarom de overdrijving nu kleiner is, is het een geruststellende bevinding. Eveneens in overeenstemming met eerder onderzoek is de bevinding dat kleinere giften vaker worden vergeten dan grotere. Deze bevinding ondersteunt de hypothese dat saillante gebeurtenissen nauwkeuriger worden gerapporteerd.

De bevindingen over kenmerken van respondenten die samenhangen met discrepanties tussen gerapporteerde en geregistreerde giften wijken af van eerdere bevindingen van Bekkers & Wiepking (2010). Nu worden geen significante verbanden gevonden met geslacht, opleiding of inkomen. Wel vinden we een positief verband met kerkgang, terwijl eerder een negatief verband werd gevonden met kerkelijkheid. Het is onduidelijk hoe de resultaten verklaard kunnen worden. Kerkgangers zijn doorgaans vrijgevinger dan onkerkelijken (Bekkers & Wiepking, 2007).

De resultaten bieden steun voor de algemene stelling dat geheugenproblemen en een gebrek aan informatie een belangrijke bron van vertekening zijn in enquêtes. Oudere respondenten gaven minder vaak nauwkeurige antwoorden. Dit kan verklaard worden door afnemend cognitief functioneren. Ook tweeverdieners gaven minder vaak nauwkeurige antwoorden. Deze respondenten zijn waarschijnlijk niet altijd goed op de hoogte van het geefgedrag van de partner.

Een opvallende bevinding is dat de nauwkeurigheid van rapportages over geefgedrag hoger is onder respondenten die ook ander sociaalwenselijk gedrag rapporteren zoals het stemmen bij verkiezingen en het geven van bloed. In enquêtes ligt het percentage respondenten dat zegt gestemd te hebben bij de verkiezingen en het percentage bloed donoren doorgaans hoger dan het populatiegemiddelde. Ook in de huidige gegevens is dat het geval. Vaak worden deze discrepanties geïnterpreteerd als sociaalwenselijk gedrag in enquêtes. Nu blijkt dat de respondenten die zeggen gestemd te hebben en bloed te geven juist de meer nauwkeurige respondenten zijn en minder vaak sociaalwenselijk te antwoorden. Voor respondenten die zeggen vrijwilligerswerk te doen geldt zelfs dat zij de giften onderdrijven. Deze respondenten zijn dus te bescheiden. Het ligt voor de hand dat de oververtegenwoordiging van stemmers en bloeddonors (en waarschijnlijk ook vrijwilligers, hoewel daar geen populatiecijfers voor bekend zijn) onder de NDP-

respondenten veroorzaakt wordt door hun grotere bereidheid aan enquêtes mee te werken (Abraham, Helms & Presser, 2009). Geruststellend is ook het verband tussen nauwkeurigheid en het vertrouwen in goededoelenorganisaties. Het vertrouwen dat deze respondenten zeggen te hebben is geen sociaalwenselijk antwoord. Aanvullende analyses op basis van de geregistreerde giften (hier niet gerapporteerd) laten bovendien zien dat stemmers, bloeddonors, vrijwilligers en respondenten met meer vertrouwen in goededoelenorganisaties ook daadwerkelijk vaker aan Greenpeace geven. Het zijn de betrokken burgers in Nederland die vaker geven, vaker meedoen aan enquêtes en daarin ook nog de meest nauwkeurige antwoorden geven.

## Referenties

Abraham, K.G., Helms, S. and Presser, S. (2009). How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work in the United States. *American Journal of Sociology*, 114: 1129–1165.

Bekkers, R., and Wiepking, P. (2007). Generosity and Philanthropy: A Literature Review (October 28, 2007). Available at SSRN: <http://ssrn.com/abstract=1015507>.

Bekkers, R., and Wiepking, P. (2010). 'Accuracy of Self-reports on Donations to Charitable Organizations'. *Quality & Quantity*, 44.

Belli, R.F., Traugott, M.W., Young, M., and McGonagle, K.A. (1999). Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring. *Public Opinion Quarterly*, 63: 90–108.

Bernstein, R., Chadha, A., and R. Montjoy (2001). Overreporting Voting: Why It Happens and Why It Matters. *Public Opinion Quarterly*, 65: 22–44.

Burt, C.D.B., and Poppo, J.S. (1998). Memorial Distortions in Donation Data. *Journal of Social Psychology*, 138: 724–733.

Cahalan, D. (1968). Correlates of Respondent Accuracy in the Denver Validity Survey. *Public Opinion Quarterly*, 32: 607–621.

CBF (2009). *Financiële resultaten van goededoelenorganisaties in Nederland*. Verslag Fondsenwerving 2008. Amsterdam: CBF.

Crowne, D.P., and Marlowe, D. (1960). A New Scale of Social Desirability Independent of Psychopathology. *Journal of Consulting Psychology*, 24: 349–354.

Crowne, D.P., and Marlowe, D. (1964). *The Approval Motive: Studies in Evaluative Dependence*. New York: John Wiley & Sons.

Hahn, R.A., Eaker, E., and Rolka, H. (1997). Reliability of reported age at menopause. *American Journal of Epidemiology*, 146: 771–775.

Hardt, J., and Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *Journal of Child Psychology and Psychiatry*, 45 (2): 260–273.

Harrell, A. (1997). The validity of self-reported drug use data: the accuracy of responses on confidential self-administered answered sheets. *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, ed. Harrison, L., and Hughes, A., Rockville: U.S. Department of Health and Human Services.

Johnson, T.P., O'Rourke, D.P., Burris, J.E., and Warnecke, R.B. (2005). An investigation of the effects of social desirability on the validity of self-reports of cancer screening behaviors. *Medical Care*, 43 (6): 565–573.

Katosh, J.P., and Traugott, M.W. (1981). The Consequences of Validated and Self-Reported Voting Measures. *Public Opinion Quarterly*, 45: 519–535.

Maxfield, M.G., Weiler, B.L., and Widom, C.S. (2000). Comparing self-reports and official records of arrests. *Journal of Quantitative Criminology*, 16: 87–110.

Midanik, L. (1982). The Validity of Self-Reported Alcohol Consumption and Alcohol Problems: A Literature Review. *British Journal of Addiction*, 77: 357–382.

Mills, J.F., and Kroner, D.G. (2005). An Investigation Into the Relationship Between Socially Desirable Responding and Offender Self-Report. *Psychological Services*, 2: 70–80.

Parry, H.J., and Crossley, H.M. (1950). Validity of Response to Survey Questions. *Public Opinion Quarterly*, 14, 61–80.

Phillips, D.L., and K.J. Clancy (1972). Some Effects of “Social Desirability” in Survey Studies. *American Journal of Sociology*, 77: 921–940.

Presser, S. (1984). Is Inaccuracy on Factual Survey Items Item-Specific or Respondent-Specific? *Public Opinion Quarterly*, 48: 344–355.

Presser, S., and Stinson, L. (1998). Data Collection Mode and Social Desirability Bias in Self-Reported Religious Attendance. *American Sociological Review*, 63: 137–145.

- Presser, S., and Traugott, M. (1992). Little White Lies and Social Science Models: Correlated Response Errors in a Panel Study of Voting. *Public Opinion Quarterly*, 56: 77–86.
- Schuyt, T.N.M., Gouwenberg, B.M., and Bekkers, R. (2009). *Geven in Nederland 2009: Giften, Sponsoring, Legaten en Vrijwilligerswerk*. Amsterdam: Reed Business.
- Silver, B.D., Anderson, B.A., Abramson, P.R. (1986). Who Overreports Voting? *American Political Science Review*, 80: 613–624.
- Stocké, V. (2007a). Response Privacy and Elapsed Time Since Election Day as Determinants for Vote Overreporting. *International Journal of Public Opinion Research*, 19: 237–246.
- Stocké, V. (2007b). Determinants and Consequences of Survey Respondents' Social Desirability Beliefs about Racial Attitudes. *Methodology*, 3: 125–138.
- Stoop, I.A.L. (2005). *The Hunt for the Last Respondent*. The Hague: SCP.
- Sudman, S., and Bradburn, N.M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 73: 805–815.
- Thornberry, T.P., and Krohn, M.D. (2000). *The Self-Report Method for Measuring Delinquency and Crime*, in *CJ 2000: Innovations in Measurement and Analysis*, ed.
- Duffee, D., Crutchfield, R.D., Mastrofski, S., Mazerolle, L., McDowall, D., and Ostrom, B., Washington, DC: National Institute of Justice.
- Tourangeau, R., Jobe, J.B., Pratt, W.F., and Rasinski, K. (1997). Design and Results of the Women's Health Study. In *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, ed. Harrison, L., and Hughes, A., Rockville: U.S. Department of Health and Human Services.
- Tourangeau, R. Rips, L.J., and Rasinski, K.A. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Tourangeau, R., and Smith, T. (1996). Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly*, 60: 275–304.
- Traugott, M.W., and Katosh, J.P. (1979). Response validity in surveys of voting behavior. *Public Opinion Quarterly*, 43: 359–377.

Udry, J.R., Gaughan, M., Schwingl, P.J., and Van den Berg, B. (1996). A Medical Record Linkage Analysis of Abortion Underreporting. *Family Planning Perspectives*, 28: 228–231.

Wilhelm, M.O., and Bekkers, R. (2010). Helping Behavior, Dispositional Empathic Concern, and the Principle of Care. *Social Psychology Quarterly*, 73: 1–22.

WWAV (2009). *Onderzoeksrapportage Het Nederlandse Donateurspanel*, December 2009. Woerden: WWAV/NETQ/WDM Nederland.

## Noten

- 
- <sup>1)</sup> René Bekkers, Center for Philanthropic Studies, Faculty of Social Sciences, VU University Amsterdam. De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands. r.bekkers@vu.nl  
Dit artikel zou niet mogelijk zijn geweest zonder de inzet van Jan van Berkel, Baukje Stam, Dana Henny en Remco Meyaard van WWAV en door Kees Zitman en Willem van den Berg van Greenpeace. Dit artikel is een korte versie van het paper 'Accuracy of Self-Reported Donations to Charitable Organizations', gepresenteerd op het congres 'Registers in sociaalwetenschappelijk onderzoek: mogelijkheden en valkuilen', 24 juni, Leiden.