

Accuracy of self-reports on donations to charitable organizations

René Bekkers · Pamala Wiepking

Published online: 12 June 2010
© Springer Science+Business Media B.V. 2010

Abstract In this paper we assess overall accuracy in survey self-reports on giving to charitable organizations, direction of bias in self-reports, and the influence of this bias on relationships. We compare donations to one specific health charity reported in the Giving in the Netherlands Panel Study 2003 with donations recorded in the database ($n = 191$). We find that (a) reported donations are significantly higher than recorded donations; (b) reported amounts contributed are correlated very strongly with recorded contributions; (c) differences between amounts reported and amounts recorded are positively related to education, religious affiliation, and the tendency to social desirability, and negatively to household income. This suggests that effects of education are overestimated and effects of income and religious affiliation are underestimated using self-reports on donations rather than archival records.

Keywords Accuracy · Self-reports · Survey research · Philanthropy · Donations

1 Introduction

The purpose of this paper is to assess (1) the overall accuracy in survey self-reports on charitable giving; (2) the direction of bias in self-reports; and (3) the influence of this bias on relationships between socio-demographic variables and giving to charitable organizations.

These are important issues because numerous studies on charitable giving rely on data from self-reports obtained through survey questionnaires (e.g., Brooks 2005; Brown and Ferris 2007; Rooney et al. 2005; Schervish and Havens 1997). Several researchers have expressed concerns about the potential social desirability bias in national surveys asking for

R. Bekkers
Sociology, Utrecht University, Utrecht, The Netherlands

R. Bekkers (✉) · P. Wiepking
Department of Philanthropic Studies, Faculty of Social Sciences, VU University Amsterdam,
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
e-mail: R.Bekkers@fsw.vu.nl

self-reports on charitable donations (Hall 2001; Lee et al. 1995; Wilhelm 2007). However, hardly anything is known about the validity of these self-reported contributions.

To our knowledge, only three studies have examined accuracy of self-reports on charitable giving. In the Denver Validity Survey, Parry and Crossley (1950) found that 34% of reported contributions to the Community Chest were in fact not received. Using the same data, Cahalan (1968) reports how overreporting is related to respondent characteristics. Burt and Popple (1998) found that students in Australia tended to overestimate the amount they had donated five weeks earlier, but less so if they knew their answers would be checked for accuracy. These little known results are an important reason of concern for the validity of survey research on charitable giving, relying on self-reported amounts. We need to know much more about the accuracy of self-reports on philanthropic donations.

Previous studies on the accuracy of self-reports have mainly examined other phenomena, such as the difference between reported and recorded police arrests (Maxfield et al. 2000), drug use (Harrell 1997), household finances (Warriner 1991), corporate administration (Stray 2007), voting (Bernstein et al. 2001), mammography screening (Holt et al. 2006), and hospital visits (Ayhan and Isiksal 2004). Results from these studies are in line with the classic hypothesis that behaviors that are more socially desirable are more likely to be overreported (Parry and Crossley 1950). Thus crime and drug use are underreported, there is no overall over- or underreporting for hospital visits, household finances and corporate administration, while voting and mammography screening are overreported. Because reporting donations is the socially desirable thing to do, we expect that respondents are motivated to exaggerate the number and value of charitable contributions. Respondents might even report contributions that they have not made at all.

However, social desirability is not the only factor influencing the accuracy of self-reports. Even truthful respondents are likely to be imperfect reporters on charitable contributions, especially if they are made by other members of the household. Respondents may simply not have complete knowledge about the contributions made by other members of the household, or may have forgotten about these donations at the time of the survey.

We explore these issues of validity with a unique dataset created by matching self-reports in the Giving in the Netherlands Panel Study 2003 (GINPS03 2003) on contributions to the most popular charitable organization in the Netherlands, KWF Kankerbestrijding (Dutch Cancer Association; henceforth: KWF), with donations as recorded in the charities' own database.

2 Determinants of inaccuracy

Broadly speaking, respondents may give inaccurate reports on household contributions to charitable causes for two reasons: (1) because they do not have accurate information on total household contributions, and (2) because they are tempted to or prefer to present this information in a less accurate manner.

2.1 Inaccuracy caused by limited information

Inaccuracy in self-reports may be the result of limited information. In this case, respondents may (adequately) respond 'don't know', or give an estimate of their household's likely contribution. The psychology of memory accuracy is complex: memory may be inaccurate for many reasons (Koriat et al. 2000). We will discuss three reasons. One obvious factor

in accuracy is self-performance: memory for events performed by oneself is more accurate than for events performed by others (Hornstein and Mulligan 2004). Assuming that decision making responsibility includes performing the act of donating, one would expect that respondents who are not engaged in decisions on charitable contributions within the household have limited information, and are more likely to give an inaccurate or ‘don’t know’ answer.

Finally, personal salience of the event is also an important factor in accuracy of self-reporting. More salient events are remembered more accurately. Salience is higher for (a) unusual events, (b) events involving greater social or economic costs and (c) consequential events (Sudman and Bradburn 1982, p. 42). Charitable gifts are made frequently, usually involve small amounts of money, and often have no important consequences in the lives of donors. One would therefore expect the overall accuracy of self-reports on donations to be rather low. The salience hypothesis implies that more salient donations should be reported more accurately. We assume that the salience of a charitable donation is higher if the contribution is either large or unusual. Consider a household that regularly donates money to a multitude of charitable organizations and a household that seldom donates money. We expect that a donation to a randomly chosen organization is more likely to be reported accurately by the latter household than by the former. The larger the number of donation acts performed by a household, the lower the accuracy of reports by members of this household on each individual donation act.

We also assume that the salience of a charitable donation to a specific organization is higher if the contribution constitutes a larger part of the total amount donated by the household. Thus we expect that the contribution to a randomly chosen organization is more likely to be reported accurately if that contribution constitutes a larger part of the total amount donated by the household.

Finally, we hypothesize that the salience of a contribution to the organization that we are studying in the present paper—KWF, a cancer research charity—is higher in the lives of people who are in close contact with cancer patients. We therefore expect that people who are in close contact with cancer patients report donations to KWF more accurately.

In addition to situational sources of accuracy, there are also individual sources of accuracy. Because accuracy of recall depends on the quality of cognitive functioning, we expect that older and lower educated respondents are more likely to have forgotten about their contributions or to remember them inaccurately—assuming that older and lower educated persons have lower levels of cognitive functioning. Previous studies on accuracy of self-reports on medical events have found that accuracy of self-reports increases with education (Ayhan and Isiksal 2004; Hahn et al. 1997). In previous research it has been found that more extensive survey questionnaires increase recall of charitable contributions among older and lower educated respondents (Bekkers and Wiepking 2006).

2.2 Inaccuracy caused by motivated responding

Inaccuracy of self-reports may not only be the result of incomplete information, but can also be the result of motivated responding. In the case of motivated responding, respondents knowingly report an inaccurate amount that they anticipate will generate positive rewards (or avoid reporting accurately because they anticipate negative rewards). The most well known example of motivated responding is the socially desirable answer. Responses may have a ‘social desirability bias’ when they present a more favorable picture of the household. In our case, a social desirability bias would result in exaggerating charitable contributions.

There are both individual and situational sources of socially desirable responding. The main situational source of socially desirable responding is the presence of others who may approve of socially desirable answers. In general, social desirability increases with a higher level of personal interaction between interviewer and respondent (Schwartz et al. 1991). The GINPS03 survey employs Computer Assisted Self-Administered procedures, which is likely to limit the social desirability bias. Because there is no interaction with an interviewer, the only reasonable source of approval for a social desirable answer in an online survey is self-approval. Only those respondents who would like to appear more generous to themselves than they actually are should give a socially desirable answer.

The response categories offered in a survey also affect social desirability bias. Offering face-saving response options may lower the level of social desirability (Belli et al. 1999). In our survey, respondents were first asked whether their household had made a donation to KWF, and if so, they were asked what the amount donated was. In the amount question, a 'don't know' option was available. The 'don't know' answer to the amount question is a face-saving option for respondents who first claimed to have made a donation while in fact none was made. The possibility of giving a 'don't know' answer is likely to have reduced the tendency to give socially desirable answers.

The well known 'SD-scale' developed by Crowne and Marlowe (1964) is an attempt to tap the individual tendency to give socially desirable answers. It is assumed that some respondents are more concerned with social evaluation and are therefore more likely to give socially desirable answers than other respondents. It has been argued that social desirability is not just a response bias in surveys but also a fundamental psychological process that motivates behavior (Barger 2002; McCrae and Costa 1983). Because many charitable gifts are made in public, people with a stronger desire for positive social evaluation may in fact be speaking the truth when they claim to give more to charities than those who claim they do not care about their social reputation. It is well known that making the act of giving publicly observable increases giving (Eckel and Grossman 1996; Hoffman et al. 1996).

A less well-known example of motivated responding is the case of the modest answer. Respondents with a 'modesty bias' may wish to present a less favorable picture of their household. As in the case of social desirability, there are both situational and individual sources of modesty. Previous research has shown that people are likely to underestimate their future performance when they anticipate feedback (e.g. a critical evaluation of their functioning), and when the outcome is important for their sense of self (Carroll et al. 2006; Sweeny et al. 2006; Sweeny and Shepperd 2007). In contrast, if feedback is not anticipated, people are inclined to present themselves as better than the average person. For moral behaviors this phenomenon is called the 'holier than thou' effect (Epley and Dunning 2000). Respondents in the GINPS are unlikely to anticipate feedback about their reported level of donations. They may therefore feel 'holier than thou' and report higher donations than they have made.

Individual sources of modesty originate in personality as well as religious and social norms encountered in one's social network. Respondents with an inclination to modesty are likely to underestimate their household's contributions to KWF. Modesty and related virtues such as humility and temperance are important in various religious traditions. In the gospel of Matthew, Jesus warns not to boast of one's own good works so as not to risk being deprived of the heavenly reward: "Do not let your left hand know what your right hand is doing" (Mt 6,1-4). Therefore we expect religious individuals to report lower than recorded amounts donated.

3 Data and methods

To assess the magnitude and direction of bias in self-reports on amount donated to charitable organizations in household surveys and its correlates, we matched data from the second wave of the Giving in the Netherlands Panel Study (GINPS03 2003) with data on donations provided by a major fundraising charity, KWF Kankerbestrijding (KWF). In both datasets we target donations in the calendar year 2003. KWF is a health charity aimed at reducing the incidence of cancer, improving the chances of curing the disease, and creating a better quality of life for cancer patients and their families. Almost two thirds (64.8%) of the households reported having made a contribution to KWF in the calendar year 2003.

GINPS03 was conducted in May 2004. Invitations for the survey were sent out to 1,557 persons by e-mail. 1,316 respondents (85%) completed the questionnaire. Respondents were questioned using an adaptation of the 'IU-Method+Area' module, in which first questions are asked about methods of donating followed by questions about donations to different charitable subsectors (Rooney et al. 2001). After the questions about total donations in different areas, respondents in GINPS03 were also prompted about their household's donations to 64 specific charitable organizations, including KWF (Wiepking 2008). Respondents were asked to first select those organizations to which their household had donated in 2003. After that, for all positive responses the exact amount donated and method of giving was asked. Finally, at the end of the GINPS03 survey respondents were asked whether they objected against matching their survey responses to other databases for research purposes. Only 5% objected; we discarded the data for these individuals in the present study. We provided the names and addresses of the 95% respondents who did not object against matching their data to KWF. Unfortunately, a large proportion of all donations made to KWF are not recorded in the 2003 KWF database. Only those donations for which a home address is known are recorded. Because we had no way of knowing whether donations reported from addresses that were not located in the KWF database were actually made or not, we excluded those households from our analyses. A large group of donations not recorded in the KWF database are donations made in door-to-door collections.¹ Almost two thirds (65.4%) of GINPS03 respondents who reported a household contribution to the KWF in 2003 reported a donation in the annual door-to-door collection. These donations are usually small. 69.2% of the reports by GINPS03 respondents on amounts donated to KWF in door-to-door collections are €10 or less. 7.9% reports a contribution of just €1; 17.1% reports €2. These small donations are unlikely to be registered by KWF.

As a result, only a sub sample of all donations that have actually been made to KWF is available in the KWF database for matching with GINPS03. GINPS respondents were matched using the ZIP code and first four letters of the respondents' last name. The addresses of 191 GINPS03 respondents were located in the KWF database spanning the years 1999–2004. For a subset of 139 respondents donations were recorded in the calendar year 2003. Because donations in door-to-door collections are not recorded and are typically small, it should be kept in mind that our conclusions are based on a subset of respondents who tend to make larger donations.

¹ Only door-to-door donations made using a credit slip (indicating name, address and bank number) are recorded in the KWF database. These contributions are likely larger than the average cash door-to-door contribution. 75.1% of the 723 donations reported from addresses not located in the KWF database are reported as being made in a door-to-door collection and are therefore unlikely to have been recorded. 76.3% of the donations reported from addresses not located in the KWF database are €10 or less.

4 Measures

4.1 Dependent variables: reported and recorded amount donated

The dependent variable in this study is the difference between *reported* and *recorded* amounts donated. We refer to this variable as the ‘degree of bias’. The *recorded* amount refers to the amount donated by the respondents recorded in the KWF database for the calendar year 2003. Recorded amounts ranged from €1.13 to €200. The mean amount recorded was €23.17, the median was €15.00.

The *reported* amount refers to the amount donated in the calendar year 2003 by the respondents reported in the GINPS03. 105 of the 139 respondents who were located in the KWF database reported an amount donated to KWF by their household in GINPS03.

The 34 respondents who did not report an amount consist of two groups. The first group contains 16 respondents belonging to households that actually made contributions (as recorded in the KWF database) but failed to report them in GINPS03. These contributions range from €4.50 to €28. The mean value of recorded but non-reported contributions (€13.57) is lower than the mean value of recorded and reported contributions (€24.41), but the difference is not significant in an independent samples t-test for equality of means (F -value = 2.350, $df = 137$, $p < .153$). The lower mean donations among those respondents who failed to recall them is in line with the hypothesis that less salient events are reported less accurately.

The second group contains 18 respondents, whose households were recorded in the KWF database to have made contributions, but reported in GINPS03 that they did not remember the amount donated to KWF. The recorded amounts donated by these respondents ranged from €2.50 to €35. The mean amount donated by these households as recorded in the KWF database (€17.35) was lower than the mean amount donated by those that did report an amount (€24.04), but the difference is not significant in an independent samples t-test for equality of means (F -value = 1.616, $df = 137$, $p < .355$). The lower amount donated by respondents who reported ‘don’t know’ is in line with the hypothesis that more salient events are reported less accurately. Thus, a ‘don’t know’ response or a failure to report a contribution does not necessarily reflect a small contribution. In the analyses we exclude both groups of respondents, because we lack information to determine the level of accuracy of their reports.

4.2 Independent variables

We examine the degree and direction of bias in reported contributions in relation to variables measuring decision-making authority regarding charitable contributions within the household, the salience of contributions to KWF, cognitive functioning, social desirability, modesty, and socio-demographic background characteristics.

4.2.1 Decision making

To test the hypothesis that self-performed events are measured more accurately we investigate donations reported by respondents who made the decisions on charitable contributions and by respondents who did not. GINPS03 respondents provided information on how decisions on charitable contributions above €10 are usually made within the household. Respondents could answer that these decisions are usually made by their partner, by themselves, or in consultation between themselves and their partner. We created a dummy variable contrasting

the middle category with the other two categories to test whether respondents who decide and report on contributions themselves report contributions more accurately than respondents reporting on contributions decided upon by their partner or on contributions decided upon jointly.

4.2.2 *Salience*

To test the hypothesis that the salience of an event is related to accuracy, we include a number of variables indicating the salience of contributions to KWF. First of all, we include the natural log of the total amount donated in 2003 to causes other than health reported in GINPS03. We exclude reported donations to health causes to avoid including donations to KWF (incorporated in our dependent variable) in a predictor variable. If the contribution to KWF is small relative to other contributions, it is unlikely to be remembered accurately. To test whether accuracy of reporting decreases when KWF is one of many organizations that the household donates to, we included the number of different charitable organizations that the respondent reported having made donations to in the calendar year 2003, using the list of 64 organizations. Finally, we included a variable indicating whether the respondent knows anyone in their close personal environment who has suffered from cancer in the past twelve months. 18.1% of the respondents said 'yes' to this question.

4.2.3 *Cognitive functioning*

We include dummy variables for age (between 18–34 and 35–54 years, 55 and over is the reference category) and education (primary education vs secondary and tertiary education) to examine whether older and lower educated respondents are less accurate respondents, as suggested by previous research (Bekkers and Wiepking 2006). We also include a brief measure of verbal intelligence. The measure is based on a vocabulary test asking for the correct meaning of 12 'difficult' words (Gesthuizen and Kraaykamp 2002), modeled after the WORDSUM variable included in the General Social Survey (Alwin 1991). We created dummy variables for a very low score (0–5 correct answers; 9.5%) and a very high score (11 or 12 correct answers; 8.5%).

4.2.4 *Social desirability*

Unfortunately, the GINPS03 did not include a full social desirability scale. However, among the items of a scale designed to measure selection of situations matching the respondent's personality and values we did find an item that comes close to measuring the tendency to give socially desirable answers. The item read "I do everything to make others feel comfortable". Response categories ranged from disagree completely to agree completely. We assume that respondents who agree more strongly with this statement are more likely to give socially desirable answers in surveys. Note that agreement with this item may represent not only a response bias but also one's true concern with social evaluation.

4.2.5 *Modesty*

We include dummy variables for religious affiliation (Catholic, Reformed Protestant, Rereformed Protestant, and other religion; non-religious is the reference category) to test

the hypothesis that religiously affiliated people have a stronger inclination toward modesty. We also include a more direct indicator of the tendency to give modest answers: the (standardized) difference between the agreement with two statements on the helpfulness of ‘most people’ and one’s own helpfulness. In response to the statement “I would sacrifice little to help others”, 63.8% disagreed and 18.1% disagreed completely with the statement. However, when the statement pertained to other people (“Most others would sacrifice little to help others”), only 23.8% disagreed and 1.9% disagreed completely. Likewise, 51.4% disagreed or completely disagreed with the statement “I rather work for my own welfare than for that of others”, but 60.4% agreed or completely agreed with the statement “Most people rather work for their own welfare than for that of others”. Thus most respondents felt ‘holier than thou’ with regard to their helpfulness. Subtracting the agreement with the statements on one’s own helpfulness from the agreement with the statements on the helpfulness of others we obtain two indicators of modesty. We computed the mean of these two indicators to form a scale for modesty (Cronbach’s $\alpha = .547$).

We control for gender of the respondent and a household’s financial resources with the standardized score of annual after-tax household income.

5 Results

5.1 Reported amounts donated are significantly higher than recorded amounts

The recorded amounts donated by the 105 respondents for whom we also have non-missing data on reported contributions ranged from €1.13 to €200. The mean recorded amount was €25.63, the median was €15 and the mode was €10. Among these 105 respondents, reported amounts ranged from €2 to €250. The mean amount was €33.45, the median and the mode were €25. Thus, on average, respondents overestimated donations with €7.82. This is an overestimation of the recorded contributions with 30.5%. The overestimation is somewhat smaller than the overestimation obtained in a previous study (Burt and Popple 1998). The distribution of the difference between reported and recorded donations is shown in Fig. 1.

Figure 1 shows that the distribution of the difference between reported and recorded donations is somewhat skewed. The largest underestimation is €42, and the largest overestimation is €90. The median difference between reported and recorded amounts donated was (an overestimation of) €2.32; the mode was €0.00. About a quarter reported lower than recorded amounts, and thus underestimated their household’s donations. 17.1% reported the exact recorded amount. The remaining 57.1% overestimated the amount donated. If we take a cut-off point of €5.00, about half of the sample reports accurately ($n = 51$, 48.6%), 14.3% report more than €5.00 less than their household’s recorded donations ($n = 15$), and 37.1% report contributions more than €5 above their recorded contribution ($n = 39$, 37.1%). A paired samples t-test shows that the difference between the exact reported and recorded amounts is significant (t -value 3.98, $df = 104$, $p \leq .001$).

5.2 Very strong correlations between recorded and reported contributions

Despite the fact that the reported amounts are significantly higher than the recorded amounts, the correlation between reported and recorded contributions is very strong: .853 ($p \leq .001$). When we consider the correlation between reported and recorded contributions for households with different patterns of decision-making on charitable contributions, we see that

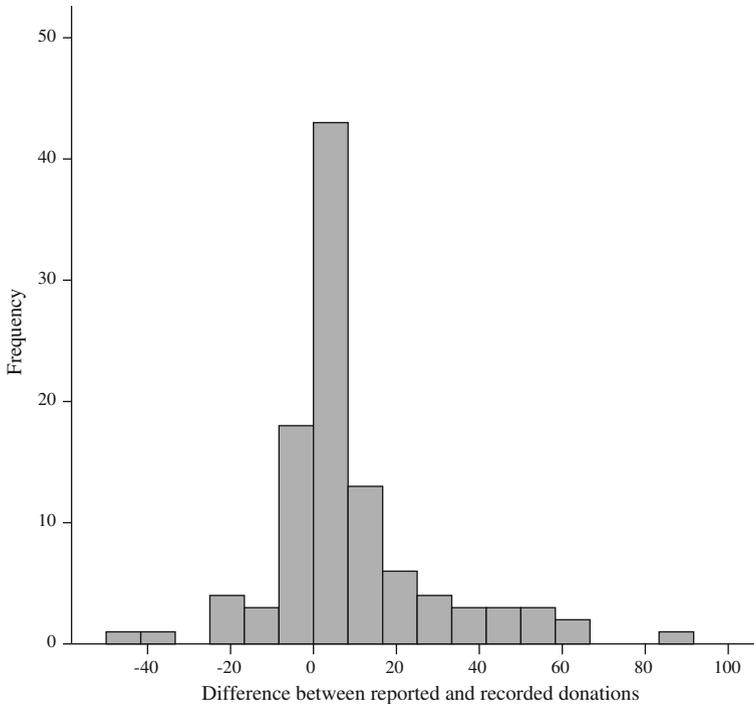


Fig. 1 Histogram of difference between reported and recorded donations to KWF in 2003

Table 1 Correlations between recorded and reported contributions

Decision-making	Reporter	N	Correlation
Female in couple	Female	12	.400
Female in couple	Male	5	-.439
Male in couple	Male	14	.920***
Male in couple	Female	8	.770 * *
Joint	Female	18	.409*
Joint	Male	27	.743***
Single female	Female	8	.885***
Single male	Male	13	.923***

* Significant at 10%;
 ** significant at 5%;
 *** significant at 1% (two-sided)

the correlation strongly depends on both who in the household decides and who reports on donations (see Table 1). The lowest level of accuracy is found for two-person households in which decisions about donations were made by the female and amounts were reported by the male. The correlation between reported and recorded contributions is even negative among this group of cases. It should be noted, however, that this finding concerns a very small group of respondents ($n = 5$). Hence, the correlation is not significant. All the other correlations are positive and exceed .400.

A few systematic patterns emerge from Table 1. In households where decision-making on charitable contributions was a joint task, accuracy of reported contributions was lower ($r = .614$; $p \leq .001$; $n = 48$) than in households where a specific person had the responsibility

for decisions on charitable contributions ($r = .842$; $p \leq .001$; $n = 45$). Female respondents in general were slightly less accurate ($r = .666$; $p \leq .001$; $n = 54$) than male respondents ($r = .854$; $p \leq .001$; $n = 52$).

5.3 Correlates of the degree of bias in reported amounts donated

In Table 2 we analyze the degree of bias in reported amounts donated—that is: the difference between reported and recorded amounts donated, among the 105 recorded donations in the KWF database. In this analysis we first include socio-economic background variables (Model 1). In Model 2 we add variables measuring the salience of contributions to KWF, and in Model 3 we add variables that may mediate the effects of these variables. Note that negative values of the dependent variable indicate that reported amounts were lower than recorded amounts, and positive values indicate that reported amounts were higher than recorded amounts. Hence positive parameter estimates indicate that the estimates were less positive in the analysis of archival records than in the analysis of self-reports, suggesting an overestimation of the parameters using self-report data.

The results in Model 1 in Table 2 reveal that the degree of bias is significantly related to education, religious affiliation, and income. The standardized effect sizes for these variables vary between .15 (for Reformed Protestant religious affiliation) and .25 (for tertiary education). Lower educated respondents, respondents with a (Protestant) religious affiliation and respondents in higher income households tend to report lower than recorded amounts donated.

The results in Model 2 reject the salience hypothesis. We find that deciding on charitable contributions oneself is associated with a higher amount reported than recorded. Respondents reporting they are the decision maker on charitable contributions within the household tend to overestimate their donations. This result stands in flat contrast with the expectation of the salience hypothesis, which predicted that the degree of bias would be smallest for respondents who made decisions on contributions, and thus has most information on the donations. With respect to the salience hypothesis, we furthermore expected that the larger the number of donation acts performed by a household, the lower the salience of a donation to a single organization such as KWF, decreasing the accuracy of reporting. We also assumed that the salience of a charitable donation to a specific organization is higher if the contribution constitutes a larger part of the total amount donated by the household. Thus we expected that the contribution to KWF is more likely to be reported accurately if that contribution constitutes a larger part of the total amount donated by the household. We find no effect of both the (self-reported) number of charitable organizations supported and the total amount donated to charitable organizations. Finally, we find that knowing a cancer patient is associated with a higher amount reported than recorded. We assumed that knowing a cancer patient increases salience of donations to KWF, and would lead to more accurate reporting. Hence, our results show clearly no support for the salience hypothesis.

The results in Model 3 support the hypothesis on social desirability as a source of motivated responding, but do not support the hypothesis on modesty. Agreement with the item tapping social desirability is associated with reporting higher than recorded amounts donated. The standardized β coefficient (.240) is among the strongest in the regression model. We find no relationship between the degree of bias and the modesty measure. The standardized β coefficient for a high score on the vocabulary test is of similar magnitude as the coefficient for social desirability (.241). People with a high verbal intelligence tend to overestimate their household's donations to KWF with about €11. In Model 3, knowing a cancer patient

Table 2 OLS regression of differences between reported and recorded donations to KWF in the calendar year 2003

	Model 1			Model 2			Model 3					
	B	SE	β	p	B	SE	β	p	B	SE	β	p
Female	-2.039	4.038	-.051		-3.403	4.064	-.084		-4.118	4.264	-.102	
Age under 35	1.219	8.220	.016		.881	8.174	.012		5.388	8.158	.071	
Age 35-55	4.128	4.317	.096		4.858	4.301	.113		6.410	4.344	.149	
Secondary education	7.232	4.814	.160		8.317	4.770	.184	*	7.351	4.759	.163	
Tertiary education	11.577	5.227	.247	**	9.898	5.292	.211	*	9.194	5.314	.196	*
Catholic	-4.902	5.532	-.106		-5.014	5.625	-.108		-6.606	5.760	-.143	
Reformed Protestant	-6.966	5.290	-.152	*	-10.984	5.908	-.240	*	-10.461	5.895	-.229	*
Rereformed Protestant	-11.588	6.146	-.213		-12.630	7.063	-.233	*	-9.984	6.984	-.184	
Other religion	-12.543	8.863	-.146		-15.599	10.098	-.181		-13.687	9.961	-.159	
Household income (z)	-4.229	2.052	-.224	**	-4.742	2.038	-.252	**	-5.795	2.027	-.307	***
Number of charities supported					.382	.381	.108		.485	.382	.137	
Total donations except health					.327	1.547	.028		-.183	1.523	-.016	
Knows cancer patient					8.240	5.225	.159		10.282	5.219	.198	**
Respondent decides					7.482	4.508	.161	*	4.131	4.844	.089	
Low vocabulary test score									.662	7.136	.009	
High vocabulary test score									11.154	4.825	.241	**
Social desirability									4.853	2.180	.240	**
Modesty									1.303	2.152	.065	
Constant	9.063	4.823		*	1.179	7.656			.498	7.841		
Adjusted R ²	.072				.097				.143			

Reference category male, age over 55, primary education, non-religious, medium score on vocabulary test

* Significant at 10%; ** significant at 5%; *** significant at 1%

also has a sizeable relationship (€10) with the difference between reported and recorded amounts donated. Finally, household income turns out to have the strongest relationship with the degree of bias in reported amounts donated.

6 Conclusion and discussion

In this paper we assessed overall accuracy in survey self-reports, direction of bias, and influence on relationships in the case of giving to charitable organizations. We compared amounts donated to one specific health charity (KWF) reported in the Giving in the Netherlands Panel Study 2003 with the amounts recorded in the KWF database. Reported amounts were significantly higher (€7.82) than recorded amounts. This is an overestimation of 30.5%. The correlation between reported and recorded contributions is .853 ($p \leq 0.001$). Finally, we found that respondents with a lower level of education, respondents with higher household income and religious respondents reported amounts lower than recorded. Respondents with a stronger verbal ability, who know a cancer patient, and who have a stronger tendency to give socially desirable answers reported higher than recorded amounts.

We argued that respondents may give inaccurate reports on amounts donated by their household because of limited information or because of motivated responding. Research in cognitive psychology shows that memory in general is often inaccurate (Koriat et al. 2000). This is especially the case for events that are observed rather than self-performed and for events that are not salient. However, we found only limited support for predictions from this body of research.

Households in which salience of giving to charitable causes was lower did not over- or underreport donations to KWF. In this specific case of accuracy of reporting on charitable donations, salience is not associated with the accuracy of self-reports. However, there is one exception. Knowing a cancer patient and deciding on giving in a household is associated with a larger reported than recorded amount donated. In these instances, higher salience was associated with a higher degree of bias in self-reports. It is difficult to explain this finding. One possibility is that respondents who overreported their household's donations legitimated these exaggerated reports with a 'yes' to the question on knowing a cancer patient. Two other possible explanations of this finding are (1) that these respondents have in fact donated to another cancer charity than KWF (there are four other national Cancer charities in the Netherlands), and (2) that they have donated to KWF in ways not registered in the database. This could be the case for donors who gave money to friends or acquaintances who raised money for KWF in their personal network and donated this money to KWF on behalf of others. With the data at hand we cannot determine which of these explanations is (most) valid.

We did find support for the case of motivated responding. First, we found that respondents who say they do everything to make others feel more comfortable have a higher probability of overreporting. It is likely that these people suffer from the classic 'social desirability bias', and want to present a more favorable picture of their household. This is a disconcerting result because we only have one item available to measure social desirability. If we would have had a more reliable measure of social desirability it is likely that the effect of social desirability would have been stronger.

Second, Protestant respondents are less likely to report higher than recorded amounts donated than the non-religious. Donations by Rereformed Protestants to secular organizations like KWF are significantly motivated by altruistic and religious motives (Bekkers and Schuyt 2008), and this might induce a modesty bias. However, our measure of the tendency to give modest answers was not related to the degree of bias. Additional analyses reveal that

survey items measuring modesty were positively related with both reported and recorded amounts donated. This in itself is an interesting result: to our knowledge, the relationship of modesty with charitable giving has not been documented earlier. Recent research in personality psychology and positive psychology has conceptualized modesty as a significant motive in prosocial behavior (Gregg et al. 2008). Our results support this view. Modesty has a substantive effect on charitable giving, and is not just a response style.

The fact that we found significant relationships of socio-economic background variables with the difference between reported and recorded donations shows that estimates of the relationships of these variables with giving based on a regression analysis of self-reported donations are biased. In short, respondents with higher education, higher cognitive ability and a tendency to give socially desirable answers overestimate donations in their household, and respondents with a Protestant religious affiliation and respondents from a higher income household underestimate donations in their household. The negative relationship of household income with degree of bias was not expected. It is unclear what caused this relationship. Because the number of charities supported and the amount donated are included in the analysis, we can rule out that the relationship of income with bias is due to a higher frequency or volume of donations, as one would expect from a salience perspective. However, it remains possible that the accuracy of reports on (smaller) consumption decisions in general is lower in higher income households because total consumption budget in these households is higher.

Our results are at odds with several earlier studies on accuracy of self-reports. These studies found that hospital visits were overreported by younger and lower educated people (Ayhan and Isiksal 2004), and age at menopause (Hahn et al. 1997) and mammography screening (Holt et al. 2006) was reported less accurately by lower educated women. We find that age is not related to the accuracy of self-reports on donations and that higher educated people report donations to be higher than recorded. Delinquency (Maxfield et al. 2000) and credit card borrowing (Karlan and Zinman 2008) were more accurately reported by males, but charitable contributions were more accurately reported by females (Cahalan 1968). We find no gender differences in accuracy once we controlled for other characteristics. It is not clear what caused the discrepancies between our results and the results of other studies. These studies examined (1) different behaviors, (2) in different cultures and (3) using different methods. Discrepancies may be the result of either or a combination. Whatever the exact origin may be, it seems safe to conclude that there are no universal correlates of bias in self-reports.

Our finding that education is positively related to overreporting donations is in line with the finding that education is positively related to overreporting voting behavior (Bernstein et al. 2001). A post hoc hypothesis explaining this similarity is that voting and charitable giving are two different expressions of civic-mindedness that are socially valued especially among the higher educated. This hypothesis requires further testing.

This study is the first to ascertain overall accuracy in survey self-reporting, direction of bias, and influence on relationships in the case of giving to charitable organizations. However, it is important to note the limitations of this analysis. The key limitation is the fact that not all donations to KWF are recorded in the KWF database. Because home addresses of smaller donors who give in door-to-door collections are not registered, we can only draw conclusions on a subset of larger, structural donations. This selectivity might have lead us to overestimate the level of overestimation. The result that on average reported donations are 30% higher than recorded donations should be interpreted with these restrictions in mind.

Another limitation is the fact that we compared reported and recorded donations for only one charitable organization in the Netherlands. It is not clear how our findings can be generalized to other specific organizations, or to global reports about donations in a

subsector in the philanthropic market. If reports on donations to individual health organizations are slight overestimates, are global reports to donations in the health sector as a whole overestimates as well? Our results should not be interpreted as implying that estimates of the effects of education and cognitive ability are always biased upward and effects of income and religion are always biased downward in self-report data. We strongly urge our colleagues to conduct more research on the validity of self-reports. Pending this research, we warn researchers that correlates of self-reported donations may not be correlates of actual donations.

A very useful future study would compare reported and recorded donations to organizations that keep records on all donations they receive. In the present study, we were unable to verify the amounts reported as made in door-to-door collections, which constitute a large proportion of all donations reported by GINPS respondents. It would be wise to replicate our study for an organization that only receives donations that are recorded, for example through direct mail campaigns.

Acknowledgements The authors thank Winny ten Damme, Barbara Hellendoorn, Marleen de Goede of KWF Kankerbestrijding and Wouter ter Haar and Fatma Al Harazi of Acxiom for their help on gathering the recorded donations.

References

- Alwin, D.F.: Family of origin and cohort differences in verbal ability. *Am. Sociol. Rev.* **56**, 625–638 (1991)
- Ayhan, H.Ö., Isiksal, S.: Memory recall errors in retrospective surveys: a reverse record check study. *Qual. Quant.* **38**, 475–493 (2004)
- Barger, S.D.: The Marlowe–Crowne affair: short forms, psychometric structure, and social desirability. *J. Pers. Assess.* **79**, 286–305 (2002)
- Bekkers, R., Schuyt, T.: And who is your neighbor? Explaining the effect of religion on charitable giving and volunteering. *Rev. Relig. Res.* **50**, 74–96 (2008)
- Bekkers, R., Wiepking, P.: To give or not to give, that's the question, how methodology is destiny in Dutch data. *Nonprofit Volunt. Sect. Q.* **35**, 533–540 (2006)
- Belli, R.F., Traugott, M.W., Young, M., McGonagle, K.A.: Reducing vote overreporting in surveys: social desirability, memory failure, and source monitoring. *Public Opin. Q.* **63**, 90–108 (1999)
- Bernstein, R., Chadha, A., Montjoy, R.: Overreporting voting: why it happens and why it matters. *Public Opin. Q.* **65**, 22–44 (2001)
- Brooks, A.C.: Does Social capital make you generous?. *Soc. Sci. Q.* **86**, 1–15 (2005)
- Brown, E., Ferris, J.M.: Social capital and philanthropy: an analysis of the impact of social capital on individual giving and volunteering. *Nonprofit Volunt. Sect. Q.* **36**, 85–99 (2007)
- Burt, C.D.B., Popple, J.S.: Memorial distortions in donation data. *J. Soc. Psychol.* **138**, 724–733 (1998)
- Cahalan, D.: Correlates of respondent accuracy in the Denver validity survey. *Public Opin. Q.* **32**, 607–621 (1968)
- Carroll, P., Sweeny, K., Shepperd, J.A.: Forsaking optimism. *Rev. Gen. Psychol.* **10**, 56–73 (2006)
- Crowne, D.P., Marlowe, D.: *The Approval Motive: Studies in Evaluative Dependence*. Wiley, New York (1964)
- Eckel, C.C., Grossman, P.J.: Altruism in anonymous dictator games. *Games Econ. Behav.* **16**, 181–191 (1996)
- Epley, N., Dunning, D.: Feeling “Holier Than Thou”: are self-serving assessments produced by errors in self- or social prediction? *J. Pers. Soc. Psychol.* **79**, 861–875 (2000)
- Gesthuizen, M., Kraaykamp, G.: Verbal ability of low-educated people in the Netherlands: the downside of educational expansion. *Neth. J. Soc. Sci.* **38**, 191–211 (2002)
- GINPS03. Dataset: giving in the Netherlands Panel Study 2003. VU University Amsterdam [computer file], Amsterdam, the Netherlands (2003)
- Gregg, A.P., Hart, C.M., Sedikides, C., Kumashiro, M.: Everyday conceptions of modesty: a prototype analysis. *Pers. Soc. Psychol. Bull.* **34**, 978–992 (2008)
- Hahn, R.A., Eaker, E., Rolka, H.: Reliability of reported age at menopause. *Am. J. Epidemiol.* **146**, 771–775 (1997)

- Hall, M.H.: Measurement issues in surveys of giving and volunteering and strategies applied in the design of Canada's national survey of giving, volunteering and participating. *Nonprofit Volunt. Sect. Q.* **30**, 515–526 (2001)
- Harrell, A.: The validity of self-reported drug use data: the accuracy of responses on confidential self-administered answered sheets. In: Harrison, L., Hughes, A. (eds.) *The Validity of Self-Reported Drug Use: Improving Accuracy of Survey Estimates*, pp. 37–58. U.S. Department of Health and Human Services National Institutes of Health National Institute on Drug Abuse, Rockville, MD (1997)
- Hoffman, E., McCabe, K., Smith, V.L.: Social distance and other-regarding behavior in dictator games. *Am. Econ. Rev.* **86**, 653–660 (1996)
- Holt, K., Franks, P., Meldrum, S., Fiscella, K.: Mammography self-report and mammography claims: racial, ethnic, and socioeconomic discrepancies among elderly women. *Med. Care* **44**, 513–518 (2006)
- Hornstein, S.L., Mulligan, N.W.: Memory for actions: enactment and source memory. *Psychon. Bull. Rev.* **11**, 367–372 (2004)
- Karlan, D., Zinman, J.: Lying about borrowing. *J. Eur. Econ. Assoc.* **6**, 510–521 (2008)
- Koriat, A., Goldsmith, M., Pansky, A.: Toward a psychology of memory accuracy. *Annu. Rev. Psychol.* **51**, 481–537 (2000)
- KWF Kankerbestrijding. Jaarverslag 2005 Kwf Kankerbestrijding. KWF Kankerbestrijding, Amsterdam (2006)
- Lee, N., Halfpenny, P., Jones, A., Elliot, H.: Data sources and estimates of charitable giving in Britain. *Voluntas* **6**, 39–66 (1995)
- Maxfield, M.G., Weiler, B.L., Widom, C.S.: Comparing self-reports and official records of arrests. *J. Quant. Criminol.* **16**, 87–110 (2000)
- McCrae, R.R., Costa, P.T.: Social desirability scales: more substance than style. *J. Consult. Clin. Psychol.* **51**, 882–888 (1983)
- Parry, H.J., Crossley, H.M.: Validity of response to survey questions. *Public Opin. Q.* **14**, 61–80 (1950)
- Rooney, P., Steinberg, K., Schervish, P.: A methodological comparison of giving surveys: Indiana as a test case. *Nonprofit Volunt. Sect. Q.* **30**, 551–568 (2001)
- Rooney, P., Mesch, D.J., Chin, W., Steinberg, K.S.: The effect of race, gender, and survey methodologies on giving in the US. *Econom. Lett.* **86**, 173–180 (2005)
- Schervish, P.G., Havens, J.J.: Social participation and charitable giving: a multivariate analysis. *Voluntas* **8**, 235–260 (1997)
- Schwartz, N., Strack, F., Hippler, H.-J., Bishop, G.: The impact of administration mode on response effects in survey measurement. *Appl. Cogn. Psychol.* **5**, 193–212 (1991)
- Stray, S.: Lies, damned lies and statistics: the accuracy of survey responses. *Qual. Quant.* (2007). doi:[10.1007/s11135-007-9131-1](https://doi.org/10.1007/s11135-007-9131-1)
- Sudman, S., Bradburn, N.M.: *Asking Questions: A Practical Guide to Questionnaire Design*. Aldine, Chicago (1982)
- Sweeny, K., Carroll, P.J., Shepperd, J.A.: Is optimism always best? Future outlooks and preparedness. *Curr. Dir. Psychol. Sci.* **15**, 302–306 (2006)
- Sweeny, K., Shepperd, J.A.: Do people brace sensibly? Risk judgments and event likelihood. *Pers. Soc. Psychol. Bull.* **33**, 1064–1075 (2007)
- Warriner, G.K.: Accuracy of self-reports to the burdensome question: survey response and nonresponse error trade-offs. *Qual. Quant.* **25**, 253–269 (1991)
- Wiepking, P.: *For the love of mankind*. PhD Thesis, Vrije Universiteit Amsterdam (2008)
- Wilhelm, M.O.: The quality and comparability of survey data on charitable giving. *Nonprofit Volunt. Sect. Q.* **36**, 65–84 (2007)