Many', 'Half' or 'One out of Two'?
Assessing the Effectiveness of Counter-Biasing Techniques in Reducing Self-Reported Turnout

Mircea COMSA, Andrei GHEORGHITA

'Many', 'Half' or 'One out of Two'?
Assessing the Effectiveness of Counter-Biasing Techniques in Reducing Self-Reported Turnout

Mircea COMSA[1], Andrei GHEORGHITA[2]

Abstract

When asked in surveys about their voting record, many respondents tend to report having gone to the polls, although they have not. The consequence is a large and significant turnout overreporting, with negative consequences upon the validity of many investigations in the field of voting behaviour. An individual's tendency to exaggerate his/her voting record is usually the combined effect of several factors: social desirability bias, limitations of human memory, and acquiescence response bias. This article assesses the effectiveness of counter-biasing techniques in compensating for the social desirability bias and, consequently, reducing self-reported turnout. It compares the effects of three alternatives of 'loaded' turnout questions, providing variable information on the occurrence of non-voting behaviour (many people/around half of the people/one out of two people did not vote) and invariant face-saving response options. Self-reported turnout is compared against externally validated turnout. We employ a split-ballot survey experiment included in the 2009 Romanian Presidential Election panel study. Our analyses show that the 'half' counter-biasing formula is systematically more effective in reducing the social desirability bias than the alternatives, which perform rather similarly. When compared to validated turnout, it appears that, after applying the counter-biasing techniques, turnout overreporting due to social desirability remains at 5.7 to 8.6 percent.

Keywords: social desirability bias, self-reported turnout, turnout overreporting, counter-biasing, survey experiment, vote validation.

[1] Babeș-Bolyai University of Cluj-Napoca, Department of Sociology and Social Work, Cluj Napoca, ROMANIA. E-mail: mcomsa@socasis.ubbcluj.ro

[2] Lucian Blaga University of Sibiu, Faculty of Social Sciences and Humanities, Department of Sociology and Social Work, Sibiu, ROMANIA. E-mail: andrei.gheorghita@ulbsibiu.ro

Introduction

Much of what we know about the behaviour of fellow human beings comes from self-reported measures (Fisher, 1993). In most of the cases such measures are reliable enough. Still, when sensitive issues are considered, they fail to provide a proper picture of the specific behaviour investigated. For such issues, estimates of behaviour based on self-reported measures largely differ from estimates based upon external sources (registers, recordings, etc.) – thus the spread of the behaviour is either overestimated or underestimated. Voting is high on the list of such sensitive topics: the survey-based estimated turnout rate (no matter the type of survey or the cultural area) is typically higher than the official one (Bernstein, Chadha, & Montjoy, 2001; Clarke, Sanders, Stewart, & Whiteley, 2004; Comsa, 2010; Denver, 2003; Duff, Hanmer, Park, & White, 2007; Mattila, 2003). Turnout overreporting is large and significant, no matter the electoral framework of measurement: an average of 20% for the US national elections between 1972 and 1996 (Bernstein *et al.*, 2001); 13% for the UK national election of 1987 (Swaddle & Heath, 1989); 8% for the European Parliament (EP) elections of 1999 (Mattila, 2003); 20% to 30% for Romanian presidential and EP elections (Comsa, 2010; Gheorghita, 2011).

Systematic overestimation of turnout based on survey data may be due to two categories of factors, namely *survey-related measurement errors* and *official data-related measurement errors* (Holbrook & Krosnick, 2010). The survey-related errors may result from the specific of the *method* (social desirability effects, memory errors, acquiescence) and from the *sampling* (selection bias, panel conditioning). The official data-related errors are mainly the result of incorrect estimations of the population of 'potential' voters (incorrect denominator), but also of the total of votes cast (incorrect numerator).

Amongst the first category of factors, social desirability is most frequently referred to as being at the origin of vote overreporting (Holbrook, Green, & Krosnick, 2003; Holbrook & Krosnick, 2010; Karp & Brockington, 2005). Social desirability refers to the fact that people tend to (or not to) declare certain behaviours or ideas in relation to the presumed manner (positive, negative, and neutral) these would be regarded by the (significant) others (Duff *et al.*, 2007; Rotariu & Ilut, 2006). Thus, a socially valued behaviour tends to be overreported (this is the case of the vote), while a socially reprehended one tends to be underreported. The natural limitations of human memory are also discussed frequently at this level as strong contributors to turnout overreporting (Abelson, Loftus, & Greenwald, 1992; Belli, Traugott, & Beckmann, 2001; Belli, Traugott, Young, & McGonagle, 1999). Especially if a long period of time has passed since the elections in question, some of the respondents tend to forget whether having voted or not and mistake one election for another one. Acquiescence response bias, as respondents' tendency to agree with any assertion, regardless of its content,

is sometimes discussed as a third survey-related source of vote overreporting (Green, Krosnick, & Holbrook, 2001; Holbrook & Krosnick, 2010). Sampling selectivity is particularly relevant for explaining systematic errors in measuring turnout, as the probability of interviewing voters is higher than that of interviewing non-voters (Burden, 2000; Clausen, 1968): voters are more likely to be reached by interviewers, but also to accept taking part in the survey. Finally, as many election studies are panel surveys, turnout overreporting might also be influenced by panel conditioning effects: some of the respondents who otherwise wouldn't have voted decide to go to the polls after participating in the first stage of a survey, as they become more aware of the implications of their participation or just wish to report having voted in the next wave of the panel survey (Clausen, 1968; Kraut & McConahay, 1973; Martinez, 2003; Yalch, 1976).

The errors due to the (lack of) quality of official data are less frequently discussed in the literature, but might also have severe implications for the estimation of turnout. If the electoral body is larger than the population of voters considered in the sampling process (incorrect denominator), the gap between the estimated turnout rate and the official one artificially increases. This is the case of some categories of emigrants, but also of jailed, institutionalized or homeless individuals (Clausen, 1968; Comsa, 2010; Holbrook & Krosnick, 2010). The reverse situation occurs when official data underestimate the population of potential voters. Additionally, errors might occur in the process of counting the ballots or aggregating such results (incorrect numerator), with likely consequences on the precision of estimated turnout rate.

While sources of official data-related measurement errors are usually beyond the range of control of the investigator, there are many attempts in the literature to compensate for the distortions introduced by the survey context. Most of these attempts focused on limiting social desirability pressures through a manipulation of the context in which turnout questions are asked: placing the vote of interest in a list of electoral contexts (Abelson *et al.*, 1992; Presser, 1990), offering counter-biasing information to transmit how common non-voting behaviour is or providing socially acceptable excuses for abstention (Belli, Traugott, & Rosenstone, 1994). Their success was rather limited. Other strategies addressed simultaneously the issue of social desirability and memory-related errors (Belli, Moore, & VanHoewyk, 2006; Belli *et al.*, 1999) and managed to obtain a significantly lower rate of vote overreporting.

This article aims to extend the discussion on the effectiveness of counter-biasing information for reducing turnout overreporting. Thus, it compares the effects of three alternative formulations of the turnout question, providing invariant face-saving response options and variable information for the extension of non-voting behaviour (many people/around half of the people/one out of two people did not vote). For this purpose, it employs a split-ballot survey experiment included in the 2009 Romanian Presidential Election panel study (three waves,

N=1,504). The article starts with an extensive discussion on the effects of social desirability for measuring turnout and how counter-biasing techniques can contribute to the limitation of such effects. Next, it introduces the research design of the split-ballot survey experiment and the hypothesis to be tested. The following section is dedicated to assessing the impact of the counter-biasing formulae on turnout overreporting, both in terms of overall effectiveness to bypass the constraints of social desirability, and in terms of variations from one formula to another. Self-reported turnout is compared against both the official turnout rate and the validated turnout. The main findings and conclusions are then discussed extensively.

Reducing social desirability effects in turnout research: counter-biasing techniques

Social desirability bias is a significant problem in survey research. It involves reporting inaccurate answers to questions about sensitive topics, for which socially shared norms or values generate strong conformity expectations. Such inaccurate or erroneous answers may emerge either as a result of a conscious strategy of impression management (the presentation of self in a positive light, in accordance with the expectations of others) or as a result of an unconscious distorted self-perception (self-deception) (Paulhus, 2002; Sackheim & Gur, 1978). The social desirability bias usually occurs in relation to reporting behaviours, intentions, attitudes, values or opinions. It works in a *proscriptive* manner in areas related to the private life of the individual (sexual practices, alcohol consumption, health issues, income and savings) or to illicit or illegal behaviours (drug use, tax evasion, crime). But social desirability may also generate *prescriptive* effects when it comes to civic participation (voting, community engagement), life styles and related behaviours (reading books, attending cultural events, environmental issues) or commitment to social and moral responsibilities (charity, helping friends, being employed) (Bradburn, Sudman, & Wansink, 2004). The first scenario generates a tendency to underreport non-desirable behaviours, while the second one induces the overreporting of desirable behaviours.

Over time, there have been many attempts from the scientific community to reduce the effects of social desirability in survey research. Most of these methodological approaches are focused on providing more anonymity and a better confidentiality of responses to the survey respondents. Sometimes these strategies are obtrusive: collecting responses to sensitive questions in sealed envelopes (the sealed ballot approach) or indicating the response by choosing a card (the numbered card approach) (Bradburn *et al.*, 2004), indirect questioning (Fisher, 1993), loading techniques (Bradburn *et al.*, 2004), evaluating the response style to sensitive questions using social desirability scale scores (Paulhus, 1991) and

warning the respondents of the possibility of detecting the accuracy of their responses in various settings of the bogus pipeline technique (Roese & Jamieson, 1993). There are also non-obtrusive techniques developed for the purpose of reducing the social desirability bias in survey research: the randomized response technique (Warner, 1965) or the item count technique (Droitcour *et al.*, 2004).

The counter-biasing (loading) techniques are among the most successful a-ttempts to reduce the overreporting of desirable behaviours. This set of techniques involves formulating the questions in a less threatening way by deliberately 'loading' them in a way that welcomes the socially undesirable answer (Groves *et al.*, 2009). The aim is to reduce the ego-defensive tendency to offer socially desirable answers by devaluing the social norm associated with that behaviour. Thus the undesirable behaviour is presented as 'normal', which 'reduces the embarrassment associated with admitting to not performing a socially desirable behaviour, and thereby, makes responding to the sensitive question less threa-tening' (Raghubir & Menon, 1996: 635).

The literature discusses four main strategies for counter-biasing sensitive questions: 'everybody does it', 'assume the behaviour', 'authorities recommend', and 'provide reasons why not' (Bradburn *et al.*, 2004: 110-111). The first strategy, 'everybody does it', attempts to reduce the social desirability bias associated with a specific behaviour by presenting it as common and widespread. The question wording employed in the second strategy genuinely assumes the presence of the investigated behaviour and targets various details in relation to it. The third approach attempts to frame an underreported behaviour as recommended by someone highly respected or admired by the respondent. Finally, the 'provide reasons why not' technique[3] offers the respondent socially acceptable reasons as justification for non-performing a socially desirable behaviour. The first three strategies are usually employed for compensating the tendency to underreport socially undesirable behaviours, while the last one usually addresses overreported behaviours.

Only two of these techniques are appropriate in the attempt to compensate for turnout overreporting in survey research, namely 'everybody does it' and 'provide reasons why not'. In this particular case, the 'loading' should either transmit that non-voting is a largely widespread behaviour (first strategy) or provide face-saving socially acceptable reasons for not going to the polls (second strategy). In fact, the two strategies can easily be combined. Thus, a turnout question that posits the 'normality' of non-voting and includes acceptable justifications for voter's abstention is expected to reduce fake turnout reporting. This was actually the aim of the ANES (American National Election Studies) investigators for rephrasing the standard turnout question as:

---

[3] Or 'forgiving' wording, as referred to by other authors (Groves et al., 2009, p. 241).

In talking to people about elections, we often find that a lot of people were not able to vote because they were not registered, they were sick, or they just didn't have the time. How about you, did you vote in the elections this November?

    1. Yes, voted;
    2. No, didn't vote.

By combining the two counter-biasing techniques into a single question, it was expected to obtain a more precise self-reported measure of turnout. This was the case to a certain degree: the ANES wording was successful in reducing vote overreporting, but not to eliminate the phenomenon (Belli *et al.*, 2001; Groves *et al.*, 2009). From that point on, many alternative wordings of the turnout question were issued in the literature, with various degrees of success in reducing turnout overreporting. Among the less successful attempts, those developed by Presser (1990), Abelson, Loftus, and Greenwald (1992), Belli, Traugott, and Rosenstone (1994), and Holbrook, Green, and Krosnick (2003) deserve a particular mention for their contribution to the development of the field.

Still, several approaches proved to be remarkably effective in counterbalancing social desirability effects in self-reported turnout. Thus, the ANES 2000 turnout question employed the previous wording (as discussed above), but replaced the response alternatives with a more elaborate list, capable to induce additional face-saving loading:

In talking to people about elections, we often find that a lot of people were not able to vote because they were not registered, they were sick, or they just didn't have the time. How about you, did you vote in the elections this November?

    1. I did not vote in the election this November;
    2. I thought about voting this time but didn't;
    3. I usually vote, but didn't this time;
    4. I am sure I voted.

The new format apparently reduced the overreporting rate by 8% (Duff *et al.*, 2007). Another successful attempt to reduce social desirability effects in measuring turnout belongs to Belli, Traugott, Young, and McGonagle (1999). Building on previous experience with a failed methodological experiment (Belli *et al.*, 1994), they conceptualize turnout overreporting as a combined effect of social desirability bias and source confusion. The question format issued targets both phenomena and was able to reduce overreporting by 9%. Unfortunately, it is a quite lengthy format, raising serious practical issues for ordinary use:

The next question is about the elections in November. In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. We also sometimes find that people who thought that they had voted actually did not vote. Also, people who usually vote may have trouble saying for sure whether they voted in a particular election. In a moment, I'm going to ask you whether you voted on Tuesday, November 5th, which was [time fill] ago. Before you answer, think of a number of different things that will likely come to mind if you actually did vote this past election day; things like whether you walked, drove, or were driven by another person to your polling place [pause], what the weather was like on the way [pause], the time of day that was [pause], and people you went with, saw, or met while there [pause]. After thinking about it, you may realize that you did not vote in this particular election. [pause]. Now that you've thought about it, which of these statements best describes you?

1. I did not vote in the November 5th election;
2. I thought about voting this time but didn't;
3. I usually vote but didn't this time;
4. I am sure I voted in the November 5th election.

Besides the face-saving 'provide reasons why not' loading, all the turnout question formats have in common the fact that the respondents are provided the information that a large share of the electorate was not able to go to the polls. This is basically done by employing the expression 'a lot of people were not able to vote', which presents turnout as a phenomenon less frequent than it is supposed to be by the respondent. However, the meaning of 'a lot of people' may significantly differ from one respondent to another.

Moreover, it is rather likely that the particular way this counter-biasing information is formulated influences the degree to which turnout overreporting is reduced. In other words, for example, formulations such as 'many people', 'a lot of people' or 'a significant number of people' might not be perfect synonyms in their capacity to reduce exaggerations in turnout self-reporting. Such counter-biasing information ('everybody does it') can be provided to the respondent in many ways: aggregate (about an overall population average) or individuating (about an average member of the population), as a number or as percentage, etc. (see for details Raghubir & Menon, 1996), with likely consequences on the capacity to reduce the overreporting of a desirable behaviour. Rather few studies have explored this aspect. Among these, the experimental research of Raghubir and Menon (1996) deserves a particular mention, due to its two main conclusions: first, the aggregate way of presenting counter-biasing information appears to be more effective than the individuating one; and, second, using actual numbers appears to be more effective in reducing overreporting than using percentages.

Research objectives and design

This study aims to extend the discussion on the effectiveness of the counter-biasing technique in reducing turnout overreporting. It compares the effects of three alternative formulations of the turnout question, each one including two components of counter-biasing information. The first component is *invariant* across the three question formats and takes the form of face-saving response options. It is expected to work in a similar way across the three question formats. The second component includes information about the official turnout rate. Its wording *varies* across the three question formats: apart from the classic ANES 'many people were not able to vote', two alternative formulations are tested, namely 'around half of the people' and 'around one out of two people'. The two alternative formulations are designed to reflect rather closely the official turnout (54.4% in the first round of the 2009 Presidential Elections and 58.0% in the runoff). Variations between the three alternative questions in their capacity to reduce the overreporting bias are expected to be due to differences in the wording of this component.

Thus, the three alternative survey questions tested differ exclusively in how the 'everybody does it' counter-biasing information is formulated. They are built on a core question combining elements from the ANES turnout question and the one proposed by Belli *et al.* (1999):

For different reasons, many people/around half of the people/around one out of two people were not able to vote at [elections], while others did. Which of the following statements describes you best?

1. I did not vote in [election];
2. I thought about voting this time but didn't;
3. I usually vote but didn't this time;
4. I am sure I voted in the [election].

The face-saving response options used are identical to those proposed by Belli *et al.* (1999), also employed in ANES 2000. There are two major differences between the original version of the question (Belli *et al.*, 1999) and our core question. First, our question does not address the issue of memory failures. As our data are collected in an interval of only two weeks after the election day, there were no solid reasons to consider memory issues as potential sources of overreporting (for an extensive argument, see Belli *et al.*, 2006). Second, our

question does not mention explicitly socially accepted reasons for the abstention. Our preference went towards a generic wording ('for different reasons'), thus allowing respondents to self-project their own reasons for abstention. This decision was based on at least two categories of reasons: on the one side, a restrictive enumeration of reasons for non-voting might have become a source of bias by itself; on the other side, the lack of such an enumeration allowed the length of the turnout question to remain within reasonable limits.

To summarize, this article investigates the effects of three alternative formulations of the turnout question on the magnitude of overreporting. Two major research objectives guide our analysis. First, we are interested in variations in effects from one formulation to another, likely to be generated by the alternative wordings of the turnout rate counter-biasing information (many people/around half of the people/around one out of two people). At this level two core research questions emerge. Are clearly-specified formulations (half/one out of two) more effective in reducing turnout overreporting than the vaguely specified one (many)? And are there variations in the way people interpret the two equivalent clearly-specified formulations and, consequently, in their success in reducing overreporting?

The second research objective is to assess the effectiveness of turnout overreporting reduction strategies across various categories of individuals, based on socio-demographic and political characteristics. Thus, we analyse if the use of counter-biasing techniques induces an alternative bias in measuring turnout[4], as a result of interactions with the characteristics of respondents. In other words, we try to provide an answer to the question: are there significant variations among categories of respondents in the effectiveness of the three counter-biasing formulae to reduce turnout overreporting?

Our data come from a split-ballot survey experiment carried out within the framework of the 2009 Romanian Presidential Election Study (RPES), a three-wave national panel survey on a representative sample of 1,504 respondents. Data collection was scheduled as following: for wave 1 during the two weeks (November 9-21, 2009) prior to the first round of elections; for wave 2 between the two rounds (November 23-December 5, 2009); for wave 3 during the two weeks after the runoff (December 7-20, 2009). The split-ballot survey experiment was included in the third wave of RPES (n=1,403). Thus, the respondents were randomly assigned to three groups and each group received a different version of the turnout question. In order to increase the number of evidences available, each version of the turnout question was asked in relation to three previous electoral

---

[4] If a formula performed differently in reducing overreporting across categories of respondents, then analyses using that specific turnout question would be biased.

competitions in Romania, namely: the first and the second round of the presidential election (November 22 and December 6, 2009) and the last legislative election (November 30, 2008, the only case in which self-reported vote might be affected of by memory failures).

As mentioned previous, the survey respondents are randomly assigned to three groups, each group receiving a different format of the question. Consequently, the success of such an experiment is essentially dependent on the comparability of the groups, meaning their similarity in terms of socio-demographic characteristics. The multivariate analyses we ran (on age: 18-34, 35-49, 50-64, 65+; gender: male versus female; residency: urban versus rural) showed the absence of any statistically significant ($p \leq 0.05$)[5] differences between the three sub-samples randomly generated. In the light of these findings, the random assignment was considered successful. However, for multivariate comparability, data were weighted in order to insure a perfect overlap for the socio-demographic categories mentioned above.

The findings are tested against a validated turnout measure. Vote validation is one of the most reliable methods for the identification of genuine voters (see Duff *et al.*, 2007). In a traditional approach, it refers to verifying survey self-reported vote against the official records. In other words, each respondent is checked against the official list of voters in his/her constituency to be registered and to have cast his/her vote. This 'traditional' approach to vote validation studies poses several important limitations that are hard to ignore: the issue of overreporting is solved outside the survey context, the solution is rather expensive both in terms of money and time, there are frequent systematic errors in the official records (Presser, Traugott, & Traugott, 1990), and there are many countries forbidding public access to such official records.

Figure 1. 'Voted' stickers on the backside of an identity card in Romania

---

[5] Significance levels for the Chi-square tests varied between 0.5 and 0.8. The highest standardized residual value associated to a joint frequency was 1.1.

Our approach to vote validation follows an alternative path, originating in a particular aspect of the Romanian electoral legislation. According to this, a 'Voted' sticker (stamped and dated) is attached on the backside of the identity card of each citizen casting his/her vote by the personnel of the polling station (see *Figure 1*). This is intended to prevent situations of multiple voting. In the particular context of our study, respondents were rewarded (gift tickets) for answering the questions, at the end of wave 1 and in the beginning of wave 3. At each of the two moments of time, in order to receive the incentive, respondents were asked to present their identity card to the interview operators. The latter were requested to record the presence or absence of the voting sticker on the respondent's identity card, thus offering us a quite accurate external validation of the vote. Such a strategy of vote validation still holds several important limitations that cannot be ignored: first, a small share of respondents (around 5%) did not show their identity cards to the interview operators; second, there is a non-zero chance of error in the process of sticker identification, as the interviewer had to deal with many such stickers, coming from various electoral contexts, attached to the backside of the same identity card; third, there may have been cases of respondents having detached the proof of voting from their identity card, although such a behaviour can be interpreted as against the law. Regardless these limitations, it allowed us to gain a significant amount of knowledge on the sample's voting behaviour, in a rather inexpensive way (easy to use, controllable efforts).

Analyses and discussions

The first part of our analysis is dedicated to the *variations between the three question formats* in terms of self-reported turnout (*Table 1*). As groups are presumed to be highly similar, it is rather likely that variations in measuring self-reported turnout are due to variations in the wording of the counter-biasing information. Data show that, for every single election, the 'half' format generates the lowest self-reported turnout. Nevertheless, the differences are statistically significant in only four out of six cases. The 'many' and 'one out of two' formulae produce nearly the same results, with no difference statistically significant.

If questions related to all three electoral competitions are considered simultaneously (aggregated), the differences between the 'half' formula and the alternatives are significant both statistically and practically. On the average, the 'half' formula induces an additional decrease of 5 to 6 percent in the self-reported turnout rate compared to the two alternatives. The difference between the 'many' and 'one out of two' formulae remains statistically insignificant.

Table 1. Self-reported turnout: results of the split-ballot experiment survey

The version of turnout question \ Answer category	I did not vote	I thought about voting this time but didn't	I usually vote but didn't this time	I am sure I voted	Overreporting compared to version ... A	B	C
					2009 Romanian Presidential Elections – first round		
A ('many')	11.6	0.8	2.5	84.4	-	5.7†	-0.2
B ('half')	15.2	1.9	2.5	78.7	-5.7†	-	-5.9*
C ('one out of two')	10.9	1.6	1.6	84.6	0.2	5.9*	-
					2009 Romanian Presidential Elections – runoff		
A ('many')	8.9	1.7	2.3	86.7	-	4.5	-1.3
B ('half')	12.3	2.1	2.5	82.3	-4.5	-	-5.7*
C ('one out of two')	8.6	1.4	1.1	88.0	1.3	5.7*	-
					2008 Romanian Parliamentary Elections		
A ('many')	15.1	1.0	2.3	75.3	-	7.0*	3.1
B ('half')	18.0	0.8	4.0	68.3	-7.0*	-	-3.9
C ('one out of two')	14.5	2.5	3.6	72.2	-3.1	3.9	-
					Averages		
A ('many')	11.9	1.2	2.3	82.2	-	5.7**	0.6
B ('half')	15.2	1.6	3.0	76.4	-5.7**	-	-5.2**
C ('one out of two')	11.3	1.8	2.1	81.6	-0.6	5.2**	-

Note: Data are row percentages; the difference to 100% represents "Don't know" and "Refused" answers.

$\dagger$ *p = .10; * p = .05; ** p = .01; *** p = .001 (Mann-Whitney test).*

A comment is needed at this point. Rather counter-intuitively, one clearly-specified formulation ('half') appears to be significantly more effective in reducing turnout overreporting than its perfect equivalent in content ('one out of two'). The reasons behind this finding are hard to assess and can only be subject to speculations. However, we expect this difference to rely on the higher cognitive accessibility of 'half' (a common phrase, frequently used) compared to 'one out of two' (involving a ratio, longer, and more difficult to manage cognitively in a rather long question format). On the other side, the 'half' format performs better than the vaguely-specified 'many', which is consistent with our expectations. 'Half' is better specified, so easier to process cognitively, and more comforting in its meaning, as it transmits a higher (and unambiguous) magnitude of the non-desirable behaviour.

In order to receive a more solid empirical support for these findings, self-reported turnout measured on the three subsamples should be compared against the validated turnout. The results are presented in *Table 2*. In this second scenario, differences in success rates between the three question formats are significantly

lower compared to the initial scenario, when validated turnout was not considered (*Table 1*). No matter the counter-biasing format employed, turnout overreporting is reduced by a comparable margin. Yet, even under these specific circumstances, the 'half' formula generates self-reported turnout rates closer to the validated turnout. On the average, the 'half' formula reduces self-reported turnout by an additional 2% compared to the alternatives, but differences are not statistically significant.

Table 2. Self-reported turnout compared against validated turnout (2009 Presidential Election)

Turnout question format	Self-reported turnout	Validated turnout	Turnout overreport	Overreporting compared to turnout question version …		
				A	B	C
First round						
A ('many')	84.4	76.7	7.7***	-	2.0	-0.7
B ('half')	78.7	73.0	5.7***	-2.0	-	-2.7
C ('one out of two')	84.6	76.2	8.4***	0.7	2.7	-
Full sample	82.6	75.3	7.3***	-0.4	1.6	-1.1
Runoff						
A ('many')	86.7	78.1	8.6***	-	1.5	0.6
B ('half')	82.3	75.2	7.1***	-1.5	-	-0.9
C ('one out of two')	88.0	80.0	8.0***	-0.6	0.9	-
Full sample	85.7	77.8	7.9***	-0.7	0.8	-0.1
Averages						
A ('many')	85.6	77.4	8.2***	-	1.8	0.0
B ('half')	80.5	74.1	6.4***	-1.8	-	-1.8
C ('one out of two')	86.3	78.1	8.2***	0.0	1.8	-
Full sample	84.1	76.5	7.6***	-0.6	1.2	-0.6

Note: Data are percentages.

*† p = .10; * p = .05; ** p = .01; *** p = .001 (paired-samples test)*

The next part of the analysis is dedicated to *assessing the effectiveness of turnout overreporting reduction strategies across various categories of individuals*, the second research objective. Although the experimental groups are comparable at univariate level based on a set of variables, the observed difference between the 'half' formula and the alternatives might still rely on some differences in the structure of the groups at multivariate level. In order to reject this scenario, we ran a multinomial logistic regression analysis with vote self-report as dependent variable and question format as independent variable, while controlling for a set of socio-demographic and political characteristics[6]. Even under these circumstances, turnout overreporting remains significantly lower in the case of the 'half' formula. Thus, all other things being equal, the 'half' formula brings a supplementary decrease of 4% in overreporting (p=0.05) compared to the two alternatives.

[6] The list of socio-demographic variables controlled for includes: gender, age, education, ethnicity, and residency. The political variables controlled for are the following: general interest in politics, interest in elections, political knowledge (subjective and objective), and partisanship.

At this point, a further question arises naturally: to what degree the 'half' formula produces similar results in reducing overreporting for the sub-populations analysed? If some sub-populations differed significantly to the sample as a whole in turnout (higher or lower), the use of the 'half' formula would introduce an additional bias[7] in estimating the self-reported turnout. In order to assess to what degree is plausible such a scenario, we ran a new logistic regression analysis, using the same predictors employed in the previous model. Additionally, interaction effects were added between each independent variable and the type of turnout question (namely 'half' versus the alternatives). In this model, finding a statistically significant interaction between one variable and the turnout question format would mean that the 'half' formula produces self-reported turnout estimations that vary in relation to one or more categories of that specific variable. In order to extend the size of the three sub-samples, we pooled the data (valid n=4,079) for all the three electoral competitions questioned about (the two rounds of the 2009 presidential elections and the 2008 legislative election) under each question format. In this second model, none of the twenty six interaction effects tested proved to be statistically significant at p 0.05 and only five[8] did at p 0.10. This means that the 'half' formula does not induce significant biases on vote self-reporting in relation to the respondents' characteristics. Consequently, the findings issued for the whole population remain valid for the case of sub-populations.

A last comment is dedicated to the *magnitude of vote overreporting* in the context of counter-biasing techniques. By comparing the self-reported turnout against the validated turnout (*Table 2*), we can have a proper picture of how large remains the social desirability bias after 'loading' the questions with the counter-biasing information. A question arises naturally at this point: why not comparing self-reported turnout against the official turnout rate? There are at least two reasons for avoiding such a choice: first, it is possible that our sample (part of a panel study, third wave) differs from the population in terms of turnout, due to panel conditioning effects and non-response selection bias; second, official data are likely to underestimate massively the turnout rate, mainly due to the inclusion of temporary emigrants in the overall electoral body.

Thus, based on the validation technique discussed in the previous section, we estimate a turnout rate of 75.3% in the first round of the 2009 Presidential Elections and of 77.8% in the runoff for our sample[9]. In a similar way, validated

---

[7] If the induced effect is higher for the categories of respondents with levels of turnout overreporting above average, such a bias is useful for accurately identifying genuine voters.

[8] Central unit of a commune, interest in politics, knowledge about elections, vote satisfaction, and perceived share of friends who voted.

[9] These results are still far away from the official turnout rates issued by the Central Electoral Bureau (54.4% for the first round and 58.0% for the runoff). Beyond possible random validation errors, the reasons for this difference are likely to be found in three sources: the quality of official data, sample issues (non-response selection bias), and panel conditioning.

turnout rates are calculated for each of the three subsamples. Having validated turnout as a reference, it appears that turnout overreporting due to social desirability bias remains somewhere between 5.7% and 8.6% after applying the counter-biasing techniques. These results are statistically significant in every alternative scenario tested. However, as stated previously, the uncompensated turnout overreporting appears to be lower in the case of the 'half' format compared to the alternatives.

Conclusions and limitations

Turnout overreporting is a constant presence in all political surveys, with serious negative implications for the relevance of many analyses in the field of political behaviour. The main survey-related reasons being at the origin of this phenomenon are related to distortions in sample representativeness and to social desirability. This article focused on the latter: due to a perceived social pressure in favour of voting, some respondents tend to report having voted, although they did not. Employing data from a split-ballot survey experiment included in the 2009 Romanian Presidential Election Study, we compared the effectiveness of three different question formats in reducing turnout overreporting. In the attempt to compensate for the social desirability bias, the question formats included variable formulae of counter-biasing, suggesting that non-voting is a common behaviour ('many people/around half of the people/one out of two people did not vote'), and offered invariant face-saving response options. Findings are tested against an original strategy of vote validation.

After applying the counter-biasing techniques, the uncompensated social desirability bias is of 5.7 to 8.6 percent in self-reported turnout, varying across question formats and electoral competitions targeted. Overall turnout overreporting is likely to be higher, since additional sources of error that cannot be reduced through counter-biasing intervene (either survey-related, or official data-related). The 'around half of the people did not vote' counter-biasing formula proved systematically more effective in reducing the social desirability bias than the alternatives. The 'many' and 'around one out of two people' formulae performed rather similarly, in spite of their different logics (vaguely versus clearly) of transmitting the normalcy of the non-desirable behaviour. Therefore, it cannot be stated that clearly-specified counter-biasing formulations are more successful than the vaguely-specified one, but only that 'half' formula outperforms the alternatives. We expect the success of the 'half' formula to rely on a combination of clear content specification and higher cognitive accessibility.

The conclusions issued for the whole population remain valid across sub-populations defined in terms of socio-demographic characteristics and patterns of political involvement. None of the three counter-biasing formulae induced any

significant bias ($p \leq .05$) on self-reported turnout in relation to the respondent characteristics considered. This provides a solid support in favour of considering the use of the 'half' formula for having a more precise measure of turnout in political surveys.

However, in spite of the relative success of the 'half' counter-biasing formula, there are serious limitations in recommending it as a 'must do' solution. First, it is *dependent of the real turnout*: it is a great solution to measure electoral participation if official turnout is close to 50 percent, but it cannot be used in other circumstances. Furthermore, there are no arguments to state that a similar clearly-specified formulation for other turnout circumstances (for example 'one third' or 'two thirds') may (or may not) work in the same way in reducing overreporting. Second, it definitely *requires further testing in various electoral contexts*, in order to reduce the chance of having context-dependent findings.

Acknowledgements

References

Abelson, R.P., Loftus, E.F., & Greenwald, A.G. (1992). Attempts to improve the accuracy of self-reports of voting. In Tanur, J. M. (Ed.), *Questions about Questions*. New York: Russell Sage Foundation, pp. 138-153.

Belli, R.F., Moore, S.E., & VanHoewyk, J. (2006). An experimental comparison of question forms used to reduce vote overreporting. *Electoral Studies*, *25*(4), 751-759.

Belli, R.F., Traugott, M.W., & Beckmann, M.N. (2001). What Leads to Voting Overreports? Contrasts of Overreporters to Validated Voters and Admitted Nonvoters in the American National Election Studies. *Journal of Official Statistics*, *17*(4), 479-498.

Belli, R.F., Traugott, M.W., Young, M., & McGonagle, K.A. (1999). Reducing Vote Overreporting in Surveys: Social Desirablity, Memory Failure, and Source Monitoring. *Public Opinion Quarterly*, *63*(1), 90-108.

Belli, R.F., Traugott, S.A., & Rosenstone, S.J. (1994). *Reducing Over-Reporting of Voter Turnout: An Experiment Using a "Source Monitoring" Framework* (NES Technical Report No. nes010153). Retrieved from http://www.electionstudies.org/resources/papers/documents/nes010153a.pdf

Bernstein, R., Chadha, A., & Montjoy, R. (2001). Overreporting Voting: Why It Happens and Why It Matters. *Public Opinion Quarterly*, *65*(1), 22-44.

Bradburn, N.N., Sudman, S., & Wansink, B. (2004). Asking Questions: The Definitive Guide to Questionnaire Design - For Market Research, Political Polls, and Social and Health Questionnaires. San Francisco: Jossey-Bass Publishers.

Burden, B.C. (2000). Voter Turnout and the National Election Studies. *Political Analysis*, *8*(4), 389-398.

Clarke, H.D., Sanders, D., Stewart, M.C., & Whiteley, P. (2004). *Political Choice in Britain*. Oxford: Oxford University Press.

Clausen, A.R. (1968). Response Validity: Vote Report. *The Public Opinion Quarterly*, *32*(4), 588–606.

Comsa, M. (2010). Participarea la vot. De ce (nu) voteaza oamenii? In Comsa, M., Gheorghita, A., Tufis, C. D. (Eds.), *Alegerile pentru Parlamentul European. România, 2009*. Iasi: Polirom, pp. 55-108.

Denver, D.T. (2003). *Elections and voters in Britain*. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan.

Droitcour, J., Caspar, R.A., Hubbard, M. L., Parsley, T. L., Visscher, W., & Ezzati, T. M. (2004). The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application. In Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., Sudman, S. (Eds.), *Measurement Errors in Surveys*. New Jersey: John Wiley & Sons, pp. 185-208.

Duff, B., Hanmer, M.J., Park, W.H., & White, I.K. (2007). Good Excuses: Understanding Who Votes With An Improved Turnout Question. *Public Opinion Quarterly*, *71*(1), 67-90.

Fisher, R.J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, *20*(2), 303-315.

Gheorghita, A. (2011). Vote Transfers, Thwarted Voters and Newcomers in the 2009 Presidential Runoff in Romania. *Social Change Review*, *9*(2), 135-155.

Green, M.C., Krosnick, J.A., & Holbrook, A.L. (2001). The Survey Response Process in Telephone and Face-to-Face Surveys: Differences in Respondent Satisficing and Social Desirability Response Bias. University of Florida. Retrieved from http://www.clas.ufl.edu/users/kenwald/pos6757/spring02/tch62.pdf

Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. New Jersey: John Wiley & Sons.

Holbrook, A. L., & Krosnick, J. A. (2010). Social desirability bias in voter turnout reports Tests using the item count technique. *Public Opinion Quarterly*, *74*(1), 37-67.

Holbrook, A.L., Green, M.C., & Krosnick, J.A. (2003). Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias. *Public Opinion Quarterly*, *67*(1), 79-125.

Karp, J.A., & Brockington, D. (2005). Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries. *The Journal of Politics*, *67*(03), 825-840.

Kraut, R.E., & McConahay, J.B. (1973). How Being Interviewed Affects Voting: An Experiment. *The Public Opinion Quarterly*, *37*(3), 398-406.

Martinez, M.D. (2003). Comment on "Voter Turnout and the National Election Studies." *Political Analysis*, *11*(2), 187-192.

Mattila, M. (2003). Why bother? Determinants of turnout in the European elections. *Electoral Studies*, *22*(3), 449-468.

Paulhus, D.L. (1991). Measurement and Control of Response Bias. In Robinson, J. P., Shaver, P. R., Wrightsman, L. S. (Eds.), *Measures of Personality and Social Psychological Attitudes*. San Diego: Academic Press, pp. 17-60.

Paulhus, D.L. (2002). Socially Desirable Responding: The Evolution of a Construct. In Braun, H. I., Jackson, D.N., & Wiley, D.E. (Eds.), *The Role of Constructs in Psychological and Educational Measurement*. Mahwah: Lawrence Erlbaum Associates, Publishers, pp. 49-69.

Presser, S. (1990). Can Changes in Context Reduce Vote Overreporting in Surveys? *The Public Opinion Quarterly*, *54*(4), 586-593.

Presser, S., Traugott, M.W., & Traugott, S.A. (1990). Vote "Over" Reporting in Surveys: The Records or the Respondents? Paper presented at the International Conference on Measurement Errors, Tucson, Arizona.

Raghubir, P., & Menon, G. (1996). Asking Sensitive Questions: The Effects of Type of Referent and Frequency Wording in Counterbiasing Methods. *Psychology & Marketing (1986-1998)*, *13*(7), 633-652.

Roese, N.J., & Jamieson, D.W. (1993). Twenty Years of Bogus Pipeline Research: A Critical Review and Meta-Analysis. *Psychological Bulletin*, *114*(2), 363-375.

Rotariu, T., & Ilut, P. (2006). Ancheta sociologica si sondajul de opinie. Teorie si practica. Iasi: Polirom.

Sackheim, H.A., & Gur, R.C. (1978). Self-deception, other-deception and consciousness. In Schwartz, G.E., Shapiro, D. (Eds.), *Consciousness and self-regulation: Advances in research*. New York: Plenum Press, pp. 139-197.

Swaddle, K., & Heath, A. (1989). Official and Reported Turnout in the British General Election of 1987. *British Journal of Political Science*, *19*(4), 537-551.

Warner, S.L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, *60*(309), 63-69.

Yalch, R.F. (1976). Pre-Election Interview Effects on Voter Turnout. *The Public Opinion Quarterly*, *40*(3), 331-336.