Political Involvement and Memory Failure as Interdependent Determinants of Vote Overreporting

Volker Stocké* 
and Tobias Stark**

January 2006

We gratefully acknowledge stimulating discussions with Hartmut Esser and Bob Belli. Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim, is gratefully acknowledged.

*Sonderforschungsbereich 504, email: vstocke@rumms.uni-mannheim.de

**Sonderforschungsbereich 504, email: tstark@rumms.uni-mannheim.de
Abstract
Survey respondents have been found to systematically overreport their participation in political elections. Although the sociodemographic correlates of this response bias are well known, only a few studies have analyzed the determinants predicted by two prominent theoretical explanations for vote overreporting: memory failure and social desirability bias. Both explanations have received empirical support in studies in which the probability of vote overreporting was found to increase (a) with the time between the election and the survey interview and (b) when respondents were more politically involved. In the present paper, we argue that the effect of each of these determinants is not simply additive, but depends on the value of the respective other factor. This interaction effect has been found with data from the American National Election Studies: The probability of vote overreporting increases significantly stronger with the respondents’ political involvement when more time has elapsed since the election day.
Introduction

Research about the determinants of electoral participation relies heavily on survey respondents’ retrospective self-reports about whether they voted in the election under consideration. However, these self-reports have been found to be highly susceptible to response bias in the direction of the subjects’ falsely reporting to have voted (cf. for example: Belli et al., 1999; Presser, 1984; Presser and Traugott, 1992; Sigelman, 1982). Based on data from the American National Election Study (ANES), 7.8 and 14.2 percent of all respondents reported that they voted, although they did not (Belli et al., 2001). In Great Britain, this percentage was found to be 3 percent in 1987 and in Sweden between 3.2 and 5.9 percent in national elections from 1979 to 1988 (Granberg and Holmberg, 1991; Swaddle and Heath, 1989). Since researchers are mainly interested in the antecedence conditions of electoral participation, one could argue that biased univariate response distributions do not pose a serious threat to the research results. This however is likely to be the case when certain groups, with shared characteristics, are differently prone to overreporting. The observed differences in the self-reported voting mistakenly suggest an association between the characteristics of these groups and their disposition to vote in political elections. Empirical research has provided evidence that vote overreporting indeed biases the results of the determinants of political participation. Accordingly, the respondents’ education, the strength of their party identification and political interest on the one hand, and their propensity to vote on the other, are substantially more strongly related when self-reported rather than when validated voting behavior was utilized in the analysis (Abramson and Claggett, 1984, 1986, 1989; Cassel, 2003; Bernstein et al., 2001; Presser and Traugott, 1992; Tittle and Hill, 1967).

The subjects’ political involvement is frequently utilized when explaining their participation in political elections. However, the susceptibility to overreport voting has been found to increase substantially with the respondents’ political involvement, as measured by their political interest (Granberg and Holmberg, 1991; Presser 1984) and the strength of their party identification (Bernstein et al., 2001). These differences may result when respondents who are more politically involved and thus hold stronger participation norms, perceive voting to be more socially desirable and consequently perceive a stronger incentive to overreport voting. Aside from differences in the motivation from social desirability, cognitive factors have been found to be relevant in the probability of vote overreporting as well. Accordingly, subjects are increasingly more likely to report that they voted, although they have not, when more time has elapsed between the election and the survey interview (Abelson et al., 1992;
Belli et al., 1999). It has been assumed that this effect results from the fact that behavioral episodes, which are more deeply in the past are less available in the respondents’ memory and thus are more likely to be misreported. Consistent with this explanation, other kinds of response effects have been found to be stronger when, according to self-reported response certainties or response latencies, the requested behavior or attitudes are less cognitively available (Lavine et al., 1998, Stocké, 2004a; for inconsistent results cf. Bassili and Krosnick, 2000).

From the theoretical perspective of Rational-Choice Theory (RCT), it is expected that stronger incentives for social desirability (SD) bias, associated with a higher political involvement, and insufficient cognitive availability of behavioral episodes in memory due to elapsed time since the behavior took place, may not be independent determinants for vote overreporting. Instead, it is predicted that increasing incentives for socially desirable responding and thus a stronger political involvement has a stronger positive effect on the probability of vote overreporting when the behavioral episode of electoral participation is less cognitively available in memory (Esser, 1991; Stocké, 2004c; Tourangeau et al., 2000: 281). From this perspective, it is expected that the political involvement and the elapsed time between the election and the survey interview in interaction explain the respondents’ susceptibility to vote overreporting. Previous research has not tested whether this effect exists.

**Theoretical Framework**

Rational-Choice Theory (RCT) assumes that answering a survey question is a goal-directed, instrumentally rational selection between response options (Esser, 1991; Stocké, 2004c; Tourangeau et al., 2000: 281). Respondents in survey interviews are assumed to be motivated to realize two different goals. Firstly, respondents strive to answer questions in a way which represents as closely as possible their true inner beliefs, feelings and evaluations. This goal originates on the one hand from the respondents’ need for expressive authenticity, which is their desire to authentically express their inner states and thus reaffirm their personal identity. On the other hand, subjects strive to express what they believe to be the true answer in order to comply with norms of honesty and to avoid psychic costs resulting from deliberately telling a lie. The respondents’ accuracy motive can be expected to have the strongest effect on the answers when they have a clear conviction about what represents the true response, and no effect when all responses are regarded to be equally (in-)valid because no relevant information is cognitively available.
Secondly, respondents strive to answer survey questions in a way that creates a positive impression in others and thus realizes the goal of reaching social approval (Crowne and Marlowe, 1960). Consistent with this assumption, individual differences in the subjects’ need for social approval have been found in many studies to explain socially desirable responding (cf. for example: Bernardi et al., 2003; Larson, 2000). In order to gain the approval of others, subjects select the response option they regard as most socially desirable and thus expect to provoke positive evaluative reactions from others. This is however only expected, when others are both present and able to observe the answers. Furthermore, the approval motive leads to SD-bias to the degree that respondents perceive differences in the desirability of the available response options (Esser, 1991). These desirability beliefs are based on social norms, which the respondents perceive to be in force in society (Stocké, 2004c). Accordingly, social desirability should be a more prevalent determinant of response behavior for groups of respondents who hold stronger norms and thus perceive clearer desirability differences between the response options.

Depending on whether the respondents’ subjectively held true scores are consistent with those they believe to be most socially desirable or whether this is not the case, their accuracy motive and their need for social approval will have identical or conflicting implications for response behavior. In the case of consistency, the joint effect of both motives will ensure that subjects will reliably select the response option, which most closely represents their true inner state. The accuracy motive however contributes decreasingly to the response validity when the requested information is less available in their memory. When the respondents’ true characteristics are not consistent with those regarded as socially desirable, it is a matter of the relative strength of the two then conflicting motives which determines which answers will be selected. Respondents will increasingly select the socially desirable answer, when either the differences in the perceived desirability between the response options increases or when the

---

1 Survey data may be biased by social desirability either because of the respondents’ need for self-deception or because they want to create a positive impression on others (Paulhus, 2002). Self-deceptive SD-bias is a defensive, unconscious and self-directed distortion of survey responses, where the underlying motivation is to preserve a positive self-concept. In contrast, the aim of deliberate impression-management strategies is to manipulate an external audiences’ evaluation of the own person and to gain social approval from others. In the first case, survey reports are expected to be biased in the direction of social norms, irrespective of whether an evaluation from outside is expected. In RCT, only impression management-based SD-bias is predicted. The anonymity of answers, defined as the probability that responses cannot be identified after the interview and thus not be sanctioned at a later point in time, can be a determinant of SD-bias as well. However, we assume that such expectations are primarily relevant for questions, for instance about illegal drug use, where legal sanctions are possible. In the case of topics, as for example in the participation in political elections in political systems without mandatory voting, ‘only’ informal social sanctions are expected. Here, the possibility of immediate evaluative reactions of others and therefore the response privacy is the important factor.
accuracy motive becomes weaker due to there being less available information in memory. Furthermore, the respondents’ response behavior will be more strongly determined by either of the two response motives, when the respective other is becoming less dominant. This implies that differences in the perceived desirability of the response options should have a stronger effect on the prevalence of SD-bias when this tendency is less counterbalanced by clear beliefs about which response option represents the true answer.

Norms of civil engagement prescribe voting in political elections as a citizens’ duty, and thus, respondents are likely to assume that reporting that they voted is more socially desirable in comparison to reporting that they have not done so. How strongly these desirability differences are may differ according to the respondents’ political involvement. Consistent with this assumption, subjects who highly value different indicators of their political involvement were found to perceive a stronger obligation to participate in political elections (Knack, 1992). Thus it can be firstly assumed, as consistent with empirical studies, that political involvement leads to stronger desirability beliefs in favor of reporting voting, to more incentives for SD-bias and thus to a higher probability of overreporting. Secondly, behavioral episodes, which took place more deeply in the past are less available in the memory (Tourangeau, 2000). When respondents are asked about their participation in an election later rather than earlier in the field period of a post-election study, their reports can be assumed to be less determined by their desire to accuracy. Thus, empirical evidence suggests, the probability of vote overreporting increases with the time elapsed between election day and the survey interview.

The less respondents remember whether they actually voted at a certain election the less their accuracy motive counterbalances incentives to answer in a way they regard as socially desirable. Thus, for subjects interviewed more distantly in time from the election under consideration, differences in their desirability beliefs, associated with varying degrees of political involvement, are expected to have a stronger positive effect on their probability to overreport voting. Therefore, it can be expected that both the political involvement and the elapsed time in interaction explain the respondents’ susceptibility to overreport voting. The main aim of our empirical study is to test this hypothesis.

**Previous Research**

Most research on the determinants of vote overreporting has concentrated on its socio-demographic correlates. Accordingly, it was found that subjects who overstate their participation in political elections are on average more educated (Belli et al., 2001; Granberg and Holmberg, 1991; Hill and Hurley, 1984), less wealthy (Traugott and Katosh, 1979) and at-
tend religious services more often (Bernstein et al., 2001) than respondents who correctly reported that they failed to vote in a political election. A very consistent result from U.S.-American studies is that African Americans are more prone to vote overreporting compared with other groups of respondents (Abramson and Claggett, 1984, 1986, 1991; Anderson et al., 1988; Belli et al., 2001; Bernstein et al., 2001; Hill and Hurley, 1984; Sigelman, 1982; Traugott and Katosh, 1979). However, the results in terms of gender and age are inconsistent. Hill and Hurley (1984) as well as Belli and colleagues (2001) reported results in which men are more likely to overreport their electoral participation than women, but Traugott and Katosh (1979) as well as Cahalan (1968) did not find any gender differences in this respect. In some studies, the probability of overreporting was found to increase with the respondents’ age (Belli et al., 2001), whereas in others, it did not differ according to this dimension (Hill and Hurley, 1984) or was found to be highest for subjects in the middle of the age continuum (Granberg and Holmberg, 1991).

Abelson and colleagues (1992) analyzed the effect of increasing time distance between the election and the survey interview with post-election data from the U.S.-presidential elections in 1986 and 1988, as well as the 1988 primary election. With individual-level validation data from the official voter register, they found that 5 months after the election in 1986, the percentage of non-voters who reported to have voted was 16.3 and this proportion increased to 40.0 percent when respondents were asked about their electoral participation 6.5 months after this election. The results for the primary election in 1988 were similar: 3 months after the election, 31.6 percent, but 8 months later, 57.1 percent of the non-voters answered that they voted. However, in the case of the presidential election in 1988, the difference in vote overreporting when the surveys were conducted either 5 or 8 months after the election day was small and not statistically significant: overreporting just increased from 54.3 percent at the earlier point in time to 57.3 percent 3 months later. More evidence for the role of the elapsed time between the election day and the survey interview was found in a more recent study with two different data sources (Belli et al., 1999). The first was a telephone survey, conducted with a nationwide sample after the U.S.-presidential election in 1996. The fieldwork for this survey started the day after the election and took 85 days to be completed. It was found that shortly after the election in November, 59.8 percent of the respondents reported they voted and this figure increased significantly to 74.9 percent by the end of the survey in January, 1997. The second data source was a survey, which was conducted in Oregon after the senate election in 1996, and validation data was available for each respondent. Field work started
directly after the election day and lasted 42 days. It was found that in the first week of the survey, 15.5 percent of the respondents reported that they voted although they did not, whereas this figure grew to 29.2 percent later during the field period. This effect of elapsed time was statistically significant as well. More evidence of the relevance of memory problems leading to vote overreporting was provided in a study in which respondents were asked about whether they participated in elections held on average 44, 92 or 148 months before the survey interviews, and in two surveys where the second one was conducted 11 months after the first (Stocké, 2005). It has been found that the survey-estimated turnout rate was for all elections and both surveys higher than the official election outcome. In particular, this deviation increased when subjects answered questions about elections further in the past and in the later rather than the earlier field period. A statistically significant overestimation of electoral participation was found when subjects reported whether they voted in the survey further back in time about elections held longer ago than the most recent one.

A few studies tested the hypothesis that SD-bias is the causal mechanism, which underlies vote overreporting. In an experimental study, subjects were instructed to answer questions about their participation in past elections either in a way which will provoke a positive or negative evaluations from others (Holbrook et al., 2003). It was found that subjects under the ‘fake good’-instruction claimed to have voted significantly more often than those under the ‘fake bad’-condition. Thus, subjects regarded reporting that they voted as more instrumental for creating a positive impression than not having done so. Impression management-based SD-bias depends on whether others are able to perceive, evaluate and possibly sanction the response behavior. Thus, more vote overreporting can be expected under the condition of low response privacy. Such privacy effects have been found for a great number of other sensitive survey topics (cf. for example: Currivan et al., 2004). Whether more response privacy leads to reduced vote overreporting is an undecided question. Visser and colleagues (1996) compared the predictive power of the Columbus Dispatch Newspaper pre-election mail survey to the outcome of the state election with that of two statewide telephone surveys. In the period between 1980 and 1994, the mail survey, despite lower cooperation rates, was consistently more successful in predicting the election results than the telephone surveys. The good performance of the mail survey was attributed to the more private response situation, which discourages respondents who in fact did not intend to vote to report a candidate preference, and in this way to bias the survey results. A more direct test of privacy effects was undertaken in an early study with individual-level validation data. Here, the probability of overreporting participation in the 1972 primary election between three different modes of administration
was compared (Locander et al., 1976). These modes were self-administered drop-off pick-up questionnaires on the one hand, and non-private interviewer-administered telephone and face-to-face interviews on the other. According to the results, self-administration led to 3 percentage points less overreporting compared to the face-to-face mode, but to 5 percentage points more than in the telephone interviews. However, none of the differences were statistically significant. Another study with a locally defined random-probability sample compared the percentage of respondents who claimed to have voted in three federal elections in Germany when they answered the questions either self- or interviewer-administered (Stocké, 2005). Subjects reported that they were significantly more likely to have participated in the elections when the privacy of the response situation was low in interviewer-administered interviews. Furthermore, it was found that the survey-based measure for the electoral turnout did not differ significantly from the official figures under the condition of self-administration. In contrast, when the vote reports were recorded interviewer-administered, there was a significant overestimation of the electoral participation rate.

Several studies have analyzed the effect of different indicators for the respondents’ political involvement on the probability of vote overreporting. Presser (1984) found with validated self-reports from the Denver Community Study that the probability of falsely reporting to have voted in five different elections significantly increases with the respondents’ political interest. This effect was replicated with data in a post-election study in Sweden (Granberg and Holmberg, 1991) and with data from the ANES 1964-1990 (Belli et al., 2001). In these studies, and in one with ANES data from 1980-1988, it was found that subjects with a stronger party identification and thus a higher commitment to one of the political candidates were more likely to overreport their political participation (Bernstein et al., 2001). Furthermore, it has been shown with validated vote reports from Great Britain, New Zealand and the U.S., that respondents who reported a stronger sense of civil duty and thus a stronger obligation to vote showed a higher probability of vote overreporting (Karp and Brockington, 2005).

The study from Belli and colleagues (1999) tested with two datasets whether modifications in the question wording reduce vote overreporting. In the experimental question wording, respondents were first asked to think about different details from the election day and than to consider carefully whether they really voted in the respective election. The aim of this modification was to improve the cognitive availability of whether respondents voted or not. Furthermore, beside the response options ‘yes’ and ‘no’, the additional alternatives ‘I thought about voting this time but didn’t’ and ‘Usually I vote but didn’t this time’ were added. Al-
though the authors did not assume such an effect, this modification might have led to less negative desirability beliefs for answering that they did not participate in the respective election. Compared with the standard ANES-question wording, the experimental wording reduced vote overreporting in a survey conducted after the 1996 senate election in Oregon. The modified question reduced the probability of vote reports in a nationwide telephone survey after the 1996 U.S.-presidential election as well. Furthermore, it was found that the question wording had an increasingly stronger effect on reducing vote overreporting when more time had elapsed since election day. From our theoretical perspective, this interaction effect may result from the fact that the modified wording reduces the perceived undesirability of not having participated in the election, and this difference in the incentives from social desirability has stronger effects when memory of true behavior is weaker, as in those survey interviews conducted in the later part of the field research.

**Empirical Study**

The aim of this study is to test the hypothesis that the respondents’ political involvement and the amount of time elapsed between the election and the survey interview are interdependent determinants for the probability of vote overreporting. The following data and operationalizations are utilized to realize this aim.

**Data and Measures**

The following analyses were conducted with the 1948-2002 American National Election Studies (ANES) Cumulative Data File (The American National Election Studies, 2005). We utilized all studies from those years with national elections when respondents were asked about their participation and when these reports were validated with data from official voter registers. As the only exception, following the suggestion of Belli and colleagues (2001), we did not include the ANES panel study, which was conducted between 1972 and 1976. Thus, our data set consisted of the ANES post-presidential election studies of 1964, 1980, 1984, 1988, and the survey conducted after the Congress elections in 1978, 1986 and 1990. The outcome variable and explanatory factors were operationalized as follows:

---

2 The reason for not including this study is the strong sample-selection bias caused by panel attrition (see Belli et al., 2001 for more details).
- **Vote overreporting:** In the upper part of table 1, the results from the validation of the respondents’ reports about their electoral participation are presented for each election included in our study. *First,* as found in other studies, the prevalence of non-voters falsely reporting to have voted is much higher than that of validated voters failing to report their electoral participation: only a proportion of between 0.2 and 1.5 percent of all respondents did not report their participation in the elections, but between 7.8 and 13.2 percent of the whole sample answered that they voted but actually did not. On average across all 7 elections included in our study, 0.7 percent of the respondents were classified as underreporters, but 10.4 percent over-reported their participation. Thus, the respondents’ errors are by no means random, but are systematically in the direction of overreporting. *Second,* we found a trend in the direction of less overreporting in more recent elections: in 1965, there were 12.6 percent overreporters, and this proportion decreased nearly monotonically to 7.8 percent in the ultimate election in 1990.

An important question is in which way to construct the dependent variable for our analysis. Two different approaches have been proposed in the literature. In the first one, only validated non-voters, and thus the population at risk for overreporting, is included in the analysis (Anderson and Silver, 1986; Bernstein *et al.*, 2001). It is then tested what explains whether subjects admit their failure to have voted or falsely report an electoral participation. The second approach is to include those respondents who reported that they did vote and analyze what determines whether respondents really voted or overreported their participation in elections (Belli *et al.*, 1999). In our view, vote overreporting is the joint result of a two-stage decision process. First, all eligible citizens have to choose whether or not to participate in a particular election. In the second step, during the interview, they have to decide whether they will report a possible non-participation or not. The first decision is about participation behavior, the second about response behavior. In the case in which only self-reported voters are included into the analysis, the focus is on the explanation of why they voted or not, and thus on their participation decision. However, in our study, we restrict our analysis to response behavior and thus to the question of what determines correct or incorrect answers, given the decision not to vote in the first step of the decision process. Thus, we only included non-voters, and our dependent variable indicated whether the respondent was an admitted non-voter (coded 0) or an overreporter (coded 1). As presented in the lower part of table 1, the percentage of non-voters who inaccurately reported to have voted differed considerably between the elections included in our analysis: whereas this percentage varied between 13.0 and
27.9 percent in the elections between 1978 and 1990, the size of this group rose to 36.7 percent in the 1964 election. Significance tests have shown that in 1964, the respondents’ susceptibility to overreporting was significantly higher, compared to all remaining elections (p < 0.05). Moreover, validated non-voters showed a lower probability to report that they voted in the non-presidential elections of 1978, and especially in those of 1986 and 1990. With the exception of the difference between the election in 1978 and that in 1988, all contrasts between presidential and non-presidential elections proved to be significant at the 5-percent level.

-- Table 1 here --

- Political involvement: Consistent with other studies of the determinants of vote overreporting, we utilized three indicator variables in order to determine the degree of the respondents’ political involvement. These were (a) the respondents’ reports about the strength of their party identification, (b) those about their interest in the election campaign and (c) their general interest in governmental and public affairs. The first two variables were measured on a four-point and the latter on a three-point response scale. We assume that only respondents with a high political involvement, compared with those with medium or low involvement, hold stronger participation norms, regard voting to be substantially more socially desirable and perceive stronger incentives to overreport voting. Thus, we first created a set of dummy variables, indicating whether a respondent chose the most extreme response category, expressing the strongest party identification, the highest political interest and strongest interest in public affairs. A value of one indicated a maximum of political involvement, and a value of zero a lower level. Due to the nominal nature of these measures, we utilized principle-component analyses with tetrachoric correlations in order to test the dimensionality of the measures.

---

3 The question wording was as follows (response coding in parenthesis): Strength of party identification: ‘Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?’; if Republican or Democrat: ‘Would you call yourself a strong (Republican/Democrat) (4) or a not very strong (Republican/Democrat) (3)?’; if Independent, other or no preference: ‘Do you think of yourself as closer to the Republican or Democratic party?’ (close (2) vs. absolutely independent (1)). Interest in political campaigns: ‘Some people don’t pay much attention to political campaigns. How about you, would you say that you were very much interested (3), somewhat interested (2), or not much interested (1) in following the political campaigns this year?’ Interest in public affairs: ‘Some people seem to follow what’s going on in government and public affairs most of the time, whether there’s an election going on or not. Others aren’t that interested. Would you say you follow what’s going on in government and public affairs most of the time (4), some of the time (3), only now and then (2), or hardly at all (1)?’

4 As the original variables were measured on an ordinal scale, it would not be appropriate to use them in a factor analysis with a Pearson-correlation matrix.
formed a single latent dimension (see table 5 in the appendix). We therefore constructed an index of political involvement by combining the three dummy variables into a simple additive index and divided the resulting measure by the number of questions answered by each respondent. The index finally ranged from 0 (no answer indicated a high political involvement) to 1 (all answers indicated a high political involvement). For the elections between 1978 and 1990, the index values were found to be similar and ranged between the values of 0.15 and 0.18, indicating that the respondents had on average a low political involvement (cf. table 2, first column). Only the elections in 1980 and 1990 differed significantly from the one in 1978. However, in the case of the election in 1964, the political involvement was 0.28 and thus 40 percent higher, compared with the average of all other elections. As in the case of the probability of vote overreporting, the differences between the election in 1964 and all other election years included in our study were statistically significant (p < 0.05).

-- Table 2 here –

- Memory failure: We expected that the respondents experienced increasingly more problems in remembering whether they participated in the election under consideration when increasingly more time elapsed between the election day and the survey interview. Thus, we computed for each respondent this time distance in days. As a result, we found that there was a tendency for less time to be necessary to conduct post-election interviews after presidential elections (cf. table 2, second column). Whereas after the 1964 election, on average 25.0 days were needed to realize the post-election interviews, in 1980 these were 20.3 days, 15.1 days in 1984, and 16.9 days in 1988. Each of these decreases were statistically significant (p < 0.05). However, in the years of non-presidential elections, the length of the field period remained relatively stable at around 20 days, and none of the differences between the election years have proven to be statistically significant (p > 0.05). Again, the election in 1964 must be regarded as an extreme case: with the exception of the elections in 1978 and 1990, it took in all other years significantly less time to realize the survey interviews.

Results
In the following, we tested the hypothesis that the incentives for vote overreporting associated with a higher political involvement have a stronger effect on the probability of overreporting when more time has elapsed between the election and the survey interview. On the
level of statistical hypothesis testing, this implied that an interaction effect between the involvement index and the number of days elapsed since the election is expected to be a significant predictor for the probability of non-voters to falsely report to have voted. In order to hold all other variables constant that are already known to influence vote overreporting and to test whether their effect could be replicated in our study, we included first the respondents’ age, their sex, race and their subjectively perceived social class in the analysis. Second, we constructed an index representing the subjects’ political efficacy and controlled for it in the analyses. This index was computed by adding the answers on three items with dichotomous response options (0=low efficacy, 1=high efficacy) and then dividing the result by the number of answers which were available for each respondent. The results ranged from 0 (no feeling of efficacy) to 1 (strong feeling of efficacy). Third, we controlled for how familiar the respondents were with the candidates in the particular elections. Respondents were asked to report the names as well as the party affiliation of up to three candidates for congress and the number of correct answers was counted (0=no name and affiliation correct; 6=all correct). After normalizing this measure of candidate familiarity, it ranged from 0 (no familiarity) and 1 (high familiarity). This measure is available for the elections in the period from 1978 to 1990, but the underlying items were not asked in 1964. Fourth, subjects were asked how much they cared which candidate or party would win the election (Response options: ‘Don’t care very much’, ‘Care a good deal’ and ‘Care very much’). Responses were dichotomized and the resulting variable varied between 0 (‘Don’t care very much’ or ‘Don’t know’) and 1 (‘Care a good deal’ or ‘Care very much’). Fifth, we controlled for differences in vote overreporting between the election years by including dummy variables for these elections in the regression equation. In order to prevent a sample-selection bias due to the listwise deletion of cases with missing values on the control variables, we included missing dummies for these

5 For the question regarding respondents’ self-categorization of social class, we used the following categories: ‘average working class’ (1), ‘upper working class’ (2), ‘average middle class’ (3), ‘upper middle class’ (4). Since the answers were recorded in 1964 in a less differentiated way, it was only possible to code the respondents into either ‘working class’ (1) or ‘middle class’ (3).

6 This index was constructed from the answers to the following questions: ‘Please tell me whether you agree or disagree with these statements: (1) I don’t think public officials care much what people like me think. (2) People like me don’t have any say about what the government does. (3) Sometimes politics and government seem so complicated that a person like me can’t really understand what’s going on.’

7 Question wording: ‘Do you happen to remember the names of the candidates for Congress, that is, for the House of Representatives in Washington?’; ‘Which party does this candidate belong to?’

8 Question wording: ‘Generally speaking, would you say that you personally care a good deal which party wins the presidential election this fall, or that you don’t care very much which party wins?’ (after presidential elections); ‘How much would you say that you personally cared about the way the elections to the U.S. House of Representatives came out?’ (after non-presidential elections).
variables into the analyses. Due to space limitations, the regression coefficients for these variables are not reported.

The outcome variable of our analysis is binary, and thus, we estimated logistic regression models. Since the election in 1964 has proven to be an exceptional case with respect to the prevalence of vote overreporting and the value of all explanatory variables, we run separate analyses for this election and the cumulative data from the other election years. For each of the two samples, we first estimated a model with only the control variables and the main effects of the political involvement as well as the time elapsed since the election day. With these models, we tested whether the results from other studies about the relevance of these determinants of vote overreporting can be replicated. In a second step, we then included the interaction effect between the respondents’ political involvement and the time distance into the regression equations.

The results from regression model 1 with the 1978-1990 cumulative ANES data were in most respects consistent with those found in previous research. First, respondents who classified themselves in a higher social class proved to be more susceptible to vote overreporting: compared with subjects who felt they belonged to the working class, those who expressed an affiliation with the middle or upper middle class were significantly more likely to falsely report that they voted. Second, the susceptibility to vote overreporting increased with the respondents’ age. Third, white subjects who reported electoral participation were significantly less likely to be found to have in fact not voted, compared with non-white respondents. Fourth, when respondents had a stronger feeling of political self-efficacy, a higher knowledge about the candidates of the election and cared more about the election outcome, the probability of vote overreporting increased significantly. Fifth, the respondents’ sex did not prove to be a significant predictor for vote overreporting. Sixth, compared with the most recent election in 1990, in all other election years, subjects who reported to have voted were significantly more likely to be non-voters. Furthermore and in line with the results from other studies, we found significant main effects of our two key variables: subjects with higher values on the index of political involvement and who were interviewed later during the post-election studies had a significantly higher disposition to overreport their electoral participation. It should be noted that these factors are all determinants of vote overreporting net of the effects of all other variables included in the analysis. In model 2, we added the interaction variable between the index of political involvement and the time elapsed since election day into the regression equation (cf. table 3). The positive and statistically significant regression
The parameter indicated that an increasingly longer time between the election and the survey interview intensified the positive effect of the subjects’ political involvement on their probability of vote overreporting.

The regression models 3 and 4, which present the results from the same analysis for the election in 1964, show a surprisingly different picture (cf. table 3). First, only the effects of how strongly the respondents cared about the election outcome and that of their political involvement was completely consistent with those found for the other election years: both factors had a significantly positive effect on the susceptibility to vote overreporting. Second, for the respondents’ subjective class affiliation, their age and their self-reported political efficiency, we found for the post-election study in 1964 regression parameters of the same sign as for the other election years, but these effects were not statistically significant. This difference may be attributed to the smaller sample size and the consequently reduced statistical power. Third, in 1964, the effects of the subjects’ race as well as the elapsed time since the election day and the interaction parameter were found to be completely different from those found for all other elections. The respondents’ race had in 1964 a significant regression parameter, but the sign of the effect is in the opposite direction: here, white respondents showed a substantially higher probability of vote overreporting, compared with all other groups. The main effect of the elapsed time was found to be non-significant and had, compared with the other post-election surveys, a sign in the opposite direction. The same is true for the interaction between political involvement and elapsed time: the parameter proved to be insignificant and the sign is inconsistent with the one found for all other election years. Thus, the data from the 1964 election did not only substantially differ with respect to the marginal distributions of the factors included in our analysis from the other elections, but showed substantially different associations between overreporting and its antecedence conditions as well.

In order to more easily interpret the significant interaction parameters for all post-election surveys between 1979 and 1990, we computed predicted probabilities for vote overreporting using regression model 2 (cf. table 4). This was done for all combinations of the respondents’ political involvement (high vs. low) on the one hand and the length of time the interview was conducted after the election day (long vs. short). The results first showed that the differences in the incentives from social desirability, associated with varying levels of political involvement, had an effect of 0.119 points on the probability of vote overreporting, when the survey
interview was conducted immediately at the day after the election: the probability to overreport their electoral participation was 0.060 for subjects with minimum political involvement and increased to 0.179 for subjects with a maximum involvement. Second, the same differences in political involvement had an approximate double effect when the interview was conducted on the 95th day after the election, which represents the longest field period found in the surveys included in our study. Here, the probability of vote overreporting was 0.115 under the condition of weak political involvement and increased to 0.240 points to a value of 0.355 for subjects with strong political involvement. From the opposite perspective, looking at the differences in the effect of elapsed time for weak and strong political involvement, we found that the difference in overreporting between the minimum and maximum time period was 0.055 probability points for weakly involved persons, but 0.176 points under the condition of strong involvement. Thus, the effect of a reduced cognitive availability of the participation behavior in the respective election was more than three times as strong when the respondents perceived strong rather than weak incentives for socially desirable response behavior. These results are completely in line with our theoretical expectations.

-- Table 4 here --

Summary and Discussion

Previous research has shown two factors explaining respondents’ susceptibility in overreporting their participation in political elections. First, politically more involved subjects and second, respondents who are interviewed longer after the election day were found to be more likely to falsely report their voting behavior. In the present article, we tested the hypothesis that the respondents political involvement as a proxy variable for how socially desirable voting is perceived and the time elapsed since the election day as a determinant for the cognitive availability of whether subjects really voted are interdependent determinants for the probability of vote overreporting. Our empirical analysis first replicated different effects of the respondents’ sociodemographic characteristics and their political attitudes on their susceptibility to vote overreporting. In particular, it has been shown that the respondents’ political involvement and the time elapsed since the election day both have independent net effects on the susceptibility to vote overreporting. Second, and consistent with our theoretical hypothesis, we found that the positive effect of the respondents’ political involvement on their overreporting is significantly stronger when more time has elapsed between the election and the
survey interview. Accordingly, the accuracy motivation associated with a strong conviction about what represents the correct answer is increasingly less capable in counterbalancing incentives from social desirability when a longer time since the encoding of the behavioral episode increasingly extinguishes the information from memory.

The aforementioned results were found for all 6 ANES post-election studies conducted after the presidential and non-presidential elections in the period between 1978 and 1990. However, those found for the election survey in 1964 were substantially different. Here, the elapsed time since the election did not explain vote overreporting and consequently this factor did not moderate the effects of political involvement. This election was exceptional with respect to both the disposition of non-voters to overreport and the antecedent conditions of this bias as well: compared with all other elections, we found the highest susceptibility to overreport, the strongest political involvement and the longest elapsed time since the election day. A possible explanation for the failure of the elapsed time to explain whether subjects misreported their electoral participation is that the exceptional high level of political involvement may have made the decision to participate very salient, which in turn led to a more intense encoding and thus less forgetting of voting behavior. If this explanation holds, our theoretical model would be incomplete, since no causal effect of the respondents’ level of political involvement on the cognitive availability of the true score in memory and thus on the accuracy motivation is assumed. In order to test this possibility we would need a direct measure for how cognitively available the voting behavior was during the interviews. However, the ANES studies do not provide such a measure.

From a practical point of view, our results suggest that a longer field period of post-election studies does not only, as already known, increase vote overreporting, but that the strength of this effect differs for certain groups of respondents. Accordingly, subjects with a higher political involvement are much more strongly affected, compared to those who are less involved. Thus, the later the interviews are conducted in the field period, the stronger the difference in the self-reported electoral participation according to the dimension of political involvement, which is simply an artifact of differences in vote overreporting. Hence, the observation that political involvement is more strongly correlated with self-reported than validated voting behavior may in particular result from interviews which were conducted late in the field period (Bernstein et al., 2001; Cassel, 2003). Not only does late interviewing have a direct negative effect on the data quality of election surveys, but it intensifies the one of political involvement as well. Our results thus offer an additional argument for conducting post-election surveys as soon as possible after the election day.
Although our study shows for the first time the interdependence between two major determinants of vote overreporting, it leaves questions unanswered. Firstly, we are not able to present direct empirical evidence for the assumption that more political involvement leads respondents to perceive voting to be more socially desirable, and by this mechanism to increase incentives for socially desirable responding. For other questionnaire topics, for instance for racial attitude answers, this has been done and differences between groups of respondents have been analyzed (Stocké, 2004b). Another open question is whether the social desirability effect associated with the subjects political involvement is the consequence of impression management-driven ‘other-deception’ or results from the respondents’ tendency for self-deceptive enhancement (cf. Paulhus, 2002). In the first case, the presence of others, their ability to perceive and evaluate response behavior is the precondition for SD-bias. Under private response conditions, SD-bias can only be expected in the form of ‘self-deception’. The interviews in the ANES studies were, however, all conducted interviewer-administered, and thus both forms of social desirability effects are possible sources of vote overreporting. In order to decide about the nature of SD-bias, it would be necessary to conduct a split-ballot study, in which the privacy of the response situation is experimentally varied, and thus, SD-effects due to impression management and ‘self-deception’ can be separated.

ACKNOWLEDGEMENTS
We gratefully acknowledge stimulating discussions with Hartmut Esser and Bob Belli. This research was supported by a grant from the German Science Foundation (DFG) to the Sonderforschungsbereich 504 at the University of Mannheim.

REFERENCES


### Appendix

**Table 5:** Principal-Component Factor Analysis with Tetrachoric Correlations of the Involvement Indicators

<table>
<thead>
<tr>
<th>Election Year</th>
<th>Strength of Party Identification</th>
<th>Interest in Election Campaign</th>
<th>Interest in Public Affairs</th>
<th>Eigenvalue 1st Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>0.70</td>
<td>0.86</td>
<td>0.79</td>
<td>1.86</td>
</tr>
<tr>
<td>1978</td>
<td>0.69</td>
<td>0.91</td>
<td>0.80</td>
<td>1.94</td>
</tr>
<tr>
<td>1980</td>
<td>0.74</td>
<td>0.84</td>
<td>0.85</td>
<td>1.99</td>
</tr>
<tr>
<td>1984</td>
<td>0.69</td>
<td>0.92</td>
<td>0.74</td>
<td>1.87</td>
</tr>
<tr>
<td>1986</td>
<td>0.70</td>
<td>0.84</td>
<td>0.88</td>
<td>1.96</td>
</tr>
<tr>
<td>1988</td>
<td>0.67</td>
<td>0.85</td>
<td>0.76</td>
<td>1.73</td>
</tr>
<tr>
<td>1990</td>
<td>0.64</td>
<td>0.90</td>
<td>0.86</td>
<td>1.96</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>0.69</strong></td>
<td><strong>0.87</strong></td>
<td><strong>0.82</strong></td>
<td><strong>1.92</strong></td>
</tr>
</tbody>
</table>
Tables for

‘Political Involvement and Memory Failure as Interdependent Determinants of Vote Overreporting’

**Table 1**: Respondents’ Self-Reported and Validated Electoral Participation in American National Election Studies (ANES)

<table>
<thead>
<tr>
<th>Year of Election</th>
<th>Validated Voters %</th>
<th>Admitted Non-voters %</th>
<th>Over-reporters %</th>
<th>Under-reporters %</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>65.4</td>
<td>21.7</td>
<td>12.6</td>
<td>0.4</td>
<td>100.0 (1450)</td>
</tr>
<tr>
<td>1978</td>
<td>41.1</td>
<td>44.5</td>
<td>13.2</td>
<td>1.3</td>
<td>100.0 (2299)</td>
</tr>
<tr>
<td>1980</td>
<td>60.4</td>
<td>28.3</td>
<td>10.9</td>
<td>0.4</td>
<td>100.0 (1408)</td>
</tr>
<tr>
<td>1984</td>
<td>63.7</td>
<td>26.2</td>
<td>9.9</td>
<td>0.2</td>
<td>100.0 (1989)</td>
</tr>
<tr>
<td>1986</td>
<td>43.5</td>
<td>47.2</td>
<td>9.0</td>
<td>0.3</td>
<td>100.0 (2174)</td>
</tr>
<tr>
<td>1988</td>
<td>59.7</td>
<td>29.7</td>
<td>9.9</td>
<td>0.7</td>
<td>100.0 (1773)</td>
</tr>
<tr>
<td>1990</td>
<td>38.8</td>
<td>51.9</td>
<td>7.8</td>
<td>1.5</td>
<td>100.0 (1980)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>51.9</strong></td>
<td><strong>37.0</strong></td>
<td><strong>10.4</strong></td>
<td><strong>0.7</strong></td>
<td><strong>100.0 (13072)</strong></td>
</tr>
</tbody>
</table>

Subsample included in the Study: Population at Risk

<table>
<thead>
<tr>
<th>Year of Election</th>
<th>Validated Voters %</th>
<th>Admitted Non-voters %</th>
<th>Total % (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>63.3</td>
<td>36.7</td>
<td>100.0 (496)</td>
</tr>
<tr>
<td>1978</td>
<td>77.1</td>
<td>22.9</td>
<td>100.0 (1325)</td>
</tr>
<tr>
<td>1980</td>
<td>72.2</td>
<td>27.9</td>
<td>100.0 (553)</td>
</tr>
<tr>
<td>1984</td>
<td>72.6</td>
<td>27.4</td>
<td>100.0 (719)</td>
</tr>
<tr>
<td>1986</td>
<td>84.0</td>
<td>16.0</td>
<td>100.0 (1222)</td>
</tr>
<tr>
<td>1988</td>
<td>75.1</td>
<td>24.9</td>
<td>100.0 (702)</td>
</tr>
<tr>
<td>1990</td>
<td>87.0</td>
<td>13.0</td>
<td>100.0 (1182)</td>
</tr>
<tr>
<td><strong>Sub-Total</strong></td>
<td><strong>78.1</strong></td>
<td><strong>21.9</strong></td>
<td><strong>100.0 (6199)</strong></td>
</tr>
</tbody>
</table>
Table 2: Respondents’ Political Involvement and the Time Elapsed between the Election Day and the Survey Interview

<table>
<thead>
<tr>
<th>Election Year</th>
<th>Index of Political Involvement Mean (Std.)</th>
<th>Elapsed Time between Election and Interview (Days) Mean (Std.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>0.28 (0.32)</td>
<td>25.0 (16.5)</td>
</tr>
<tr>
<td>1978</td>
<td>0.15 (0.25)</td>
<td>21.2 (16.4)</td>
</tr>
<tr>
<td>1980</td>
<td>0.18 (0.27)</td>
<td>20.3 (13.9)</td>
</tr>
<tr>
<td>1984</td>
<td>0.17 (0.26)</td>
<td>15.1 (11.7)</td>
</tr>
<tr>
<td>1986</td>
<td>0.17 (0.26)</td>
<td>20.8 (18.4)</td>
</tr>
<tr>
<td>1988</td>
<td>0.17 (0.25)</td>
<td>16.9 (13.1)</td>
</tr>
<tr>
<td>1990</td>
<td>0.18 (0.26)</td>
<td>20.2 (16.2)</td>
</tr>
<tr>
<td>Total</td>
<td>0.18 (0.26)</td>
<td>20.0 (16.0)</td>
</tr>
</tbody>
</table>
Table 3: Determinants of Vote Overreporting in the American National Election Studies (ANES), 1978-1990 and 1964 (Logistic Regression Results)

<table>
<thead>
<tr>
<th>Elections</th>
<th>1978-1990</th>
<th>Election</th>
<th>1964</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
</tr>
<tr>
<td>B (Std.)</td>
<td>B (Std.)</td>
<td>B (Std.)</td>
<td>B (Std.)</td>
</tr>
</tbody>
</table>

**Control variables**

Subjective Social Class a)
- *upper working*\( \beta = .15 (.12) \) \( \beta = .15 (.12) \) -- --
- *average middle*\( \beta = .23 (.09)** \) \( \beta = .23 (.09)** \) \( \beta = .29 (.22) \) \( \beta = .29 (.22) \)
- *upper middle*\( \beta = .66 (.15)** \) \( \beta = .66 (.15)** \) -- --

Age (years) \( \beta = .01 (.00)** \) \( \beta = .01 (.00)** \) \( \beta = .01 (.01) \) \( \beta = .01 (.01) \)

Sex (female) b)
- *female*\( \beta = -.04 (.07) \) \( \beta = -.04 (.07) \) \( \beta = -.29 (.20) \) \( \beta = -.29 (.20) \)

Race (white) c)
- *white*\( \beta = -.26 (.08)** \) \( \beta = -.26 (.08)** \) \( \beta = .64 (.30)** \) \( \beta = .64 (.30)** \)

Index Political Efficacy \( \beta = .47 (.12)** \) \( \beta = .47 (.12)** \) \( \beta = .51 (.32) \) \( \beta = .51 (.32) \)

Index Candidate Knowledge \( \beta = 1.41 (.21)** \) \( \beta = 1.42 (.21)** \) -- --

Outcome Important? (yes) \( \beta = .47 (.08)** \) \( \beta = .82 (.08)** \) \( \beta = .47 (.21)** \) \( \beta = .47 (.21)** \)

**Election Year d)**
- *1978*\( \beta = .76 (.12)** \) \( \beta = .77 (.12)** \) -- --
- *1980*\( \beta = .81 (.14)** \) \( \beta = .80 (.14)** \) -- --
- *1984*\( \beta = .90 (.13)** \) \( \beta = .90 (.13)** \) -- --
- *1986*\( \beta = .33 (.14)** \) \( \beta = .34 (.14)** \) -- --
- *1988*\( \beta = .82 (.13)** \) \( \beta = .82 (.13)** \) -- --

**Model variables**

Elapsed Time (days) \( \beta = .01 (.00)** \) \( \beta = .01 (.00)** \) \( \beta = -.00 (.01) \) \( \beta = -.00 (.01) \)

Index Political Involvement \( \beta = 1.51 (.13)** \) \( \beta = 1.21 (.20)** \) \( \beta = .99 (.35)** \) \( \beta = 1.07 (.59)* \)

Time • Involvement -- \( \beta = .02 (.01)** \) -- \( \beta = -.00 (.02) \)

Constant \( -3.34 (.17)** \) \( -3.27 (.17)** \) \( -2.04 (.53)** \) \( -2.06 (.54)** \)

N 5703 5703 494 494

Pseudo-R² 0.13 0.13 0.08 0.08

Significance: * p ≤ 0.10; ** p ≤ 0.05; omitted categories: a) average working, b) male, c) non-white, d) 1990.
Table 4: Predicted Probabilities of Vote Overreporting for the Interaction between Political Involvement and Time Elapsed since the Election

<table>
<thead>
<tr>
<th>Political Involvement</th>
<th>Elapsed Time</th>
<th>Difference: Long - Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.060</td>
<td>0.115</td>
</tr>
<tr>
<td>High</td>
<td>0.179</td>
<td>0.355</td>
</tr>
<tr>
<td>Difference: High - Low</td>
<td>0.119</td>
<td>0.240</td>
</tr>
</tbody>
</table>

The predicted values are computed for the theoretically possible maximum and minimum of political involvement, and the maximum and minimum of involvement which at the same time empirically exists in the sample. For the elapsed time, the empirically existing range is used, where ‘short’ represents subjects who were interviewed one day after the election and ‘long’ represents subjects who were interviewed 95 days after the election (longest observed field period). The continuous control variables ‘political efficacy’, ‘candidate knowledge’ and ‘age’ were fixed at the sample means (efficacy=0.3; knowledge=0.1; age=40.4), whereas for ‘sex’, ‘subjective social class’, ‘race’, ‘importance of election outcome’ and ‘election year’, the reference categories were inserted in the regression equation.
<table>
<thead>
<tr>
<th>Nr.</th>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>06-01</td>
<td>Volker Stocké &amp; Tobias Stark</td>
<td>Political Involvement and Memory Failure as Interdependent Determinants of Vote Overreporting</td>
</tr>
<tr>
<td>05-43</td>
<td>Volker Stocké &amp; Tobias Stark</td>
<td>Stichprobenverzerrung durch Item-Nonresponse in der international vergleichenden Politikwissenschaft</td>
</tr>
<tr>
<td>05-42</td>
<td>Volker Stocké</td>
<td>Response Privacy and Elapsed Time Since</td>
</tr>
<tr>
<td>05-41</td>
<td>Josef Hofbauer, Jörg Oechssler &amp; Frank Riedel</td>
<td>Brown-von Neumann-Nash Dynamics: The Continuous Strategy Case</td>
</tr>
<tr>
<td>05-40</td>
<td>Markus Glaser, Thomas Langer, Jens Reynders &amp; Martin Weber</td>
<td>Framing Effects in Stock Market Forecasts: The Difference Between Asking for Prices and Asking for Returns</td>
</tr>
<tr>
<td>05-39</td>
<td>Tri Vi Dang</td>
<td>Alternating Offer Bargaining with Endogenous Information: Timing and Surplus Division</td>
</tr>
<tr>
<td>05-38</td>
<td>Tri Vi Dang</td>
<td>On Bargaining with Endogenous Information</td>
</tr>
<tr>
<td>05-37</td>
<td>Patric Andersson</td>
<td>Overconfident but yet well-calibrated and underconfident: A research note on judgmental miscalibration and flawed self-assessment*</td>
</tr>
<tr>
<td>05-36</td>
<td>Peter Dürsch, Albert Kolb, Jörg Oechssler &amp; Burkhard Schipper</td>
<td>Rage Against the Machines:</td>
</tr>
<tr>
<td>05-35</td>
<td>Siegfried K. Berninghaus, Hans Haller &amp; Alexander Outkin</td>
<td>Neural Networks and Contagion</td>
</tr>
<tr>
<td>05-34</td>
<td>Jacques Durieu, Hans Haller, Philippe Solal</td>
<td>Interaction on Hypergraphs</td>
</tr>
<tr>
<td>05-33</td>
<td>Markus Glaser, Martin Weber</td>
<td>Which Past Returns Affect Trading Volume?</td>
</tr>
</tbody>
</table>