What is the impact of using personalized salutations on survey participation and data quality?

The case of a web survey

Chiara Respi and Emanuela Sala
chiara.respi@unimib.it
emanuela.sala@unimib.it

Work in progress. Do not quote without authors’ permission

1. Introduction

Response rates are key indicators of data quality and may have an impact on the research potential of survey data. There are two main reasons why obtaining high levels of survey participation is important: high response rates are likely to reduce the occurrence of response bias and can enhance the power of the analysis, both in a statistical sense, as the results of some tests also depend on the sample size available, and in relation to the type of analysis that can be carried out (i.e., larger sample sizes allow to break down research findings by different socio-economic and demographic characteristics).

However, in a context in which response rates have been constantly declining over time (see, for example, Dillman, Reips, & Matzat, 2010), obtaining high level of response has become a challenge for many researchers. The situation is particularly serious for surveys conducted over the Internet, that are characterised by lower response rates than those usually obtained with other modes of data collection (Manfreda, Bosnjak, Berzelak, Haas, & Vehovar, 2008). Indeed, in case of online surveys, the absence of interviewers who encourage participation constitutes an additional obstacle to survey participation.

To boost response rates, survey designers can adopt different strategies; they can increase the number of contact attempts, use a combination of different contact modes, introduce a set of monetary and non-monetary incentives. In addition, they can personalize the survey materials, e.g., advance letters or contact e-mails, by “tailoring” survey instruments to respondent characteristics. Compared to other strategies, personalization of survey materials offers the advantage of being less costly. However, by violating the assumptions of standardisation (Fowler 1984, 1991; Fowler and Mangione 1990), it may also have an impact on data quality. Indeed, the personalization of survey materials may also affect the degree of perceived anonymity amongst respondents and, ultimately, the accurateness of the
survey data, for example, by leading to a higher number of missing data or social desirable answers (Rogelberg et al. 2006; Tourangeau and Yan 2007).

In this paper we focus on personalization of SMS salutations in the context of a web survey of university graduates in Italy. The aim of this work is to evaluate the impact of personalized vs. generic (standardized or “neutral”) SMS invitations on two aspects of non-sampling error, i.e., nonresponse and measurement error. Previous research on this topic has focused on personalization of salutations in e-mail. To the best of our knowledge, this work is the first that deals with the personalization of SMS salutations. Results from this work are meant to contribute to expand the knowledge in this research field.

2. The impact of personalized salutation in web surveys. Literature review

In postal surveys, there is consistent evidence that shows that personalization of survey materials and, in particular, advance letters have a positive impact on response (for a review, Dillman 2000). In web surveys, however, the effects of personalization on different sources of error, especially measurement error, are still unclear. To contextualise our work, we grouped the studies in this field into three main strands of research.

2.1 Impact on response rates

A number of papers looked at the impact of personalization on response rate only. One of the first studies on this topic is by Pearson and Levine (2003). Experimenting with a group of Stanford alumni, the authors looked at the impact of four types of e-mail personalization on response rates. Although personalized salutations were associated to higher response rates, the differences between the experimental groups were not statistically significant. However, when exploring differences in the sample composition, they did find evidence for significant interactions (e.g., older and younger respondents as well as engineer reacted positively to a formal personalized salutation). Recently, two other important studies belonging to this first research strand have been carried out by Australian researchers. In the community study that Sinclair et al. (2012) conducted in Melbourne they found that 4.7% of respondents who received a personalized e-mail responded to the survey compared with 2.2% of those who were addressed using a general salutation. The authors also found little evidence for selectivity in the sample composition; respondents in the “generic salutation group” were less likely to be homeowners and more likely to be greywater users than the other respondents. Similar
results were found in Short et al. (2015)’s work. Experimenting with a population of Australian breast cancer survivors, the authors found that 69% of respondents who were addressed using their name and surname completed the eligibility questionnaire compared to 50% of those who were addressed with the more generic salutation “dear member”.

2.2 Impact on response rate and measurement error

Although nonresponse error is key for survey quality, measurement error is paramount important. The second research strand on the impact of personalization in Internet surveys was initiated by Heerwegh et co-authors and includes studies that deal with these two sources of survey error.

Heerwegh et co-authors, carried out three experiments on Belgian university students to assess the link between personalization, response and measurement error. In the first study they found that personalized e-mails have a positive effect on login and response rates and no effect on drop-out rate (Heerwegh et al. 2005). For example, 57.7% of respondents who received a personalized e-mail clicked on the “Start” survey button compared with 49.1% of those who got a generic e-mail. Although the authors used different indicators to measure response, this finding was consistent amongst the other studies that they carried out later (Heerwegh 2005; Heerwegh and Loosveldt 2006). In their first paper, they also found little evidence that personalization impacts data quality. Indeed, only one of the six indicators of measurement error seems to be related to the type of e-mail salutations respondents received. As Heerwegh et al. (2005, p. 97) conclude: “there are indications that personalization may increase the tendency of respondents to answer questions in a socially desirable way”. In their later studies they further investigated the relationship between personalization and data quality, using a larger number of questions on sensitive topics. In both studies, the authors found very little evidence for social desirability bias. Heerwegh (2005) found no relationship between type of e-mail salutations and the tendency to answer in a social desirable way whereas Heerwegh and Loosveldt (2006) found no effects for five of the six measures considered in the analysis.

More recently, a number of papers have been published on these topics that reported results from studies carried out on differ types of populations. Experimenting with a population of German employees of a financial service company, Mueller et al. (2014) found no differences in response rates (measured as response to any survey question) between respondents who received e-mails with a personalized form of address and access procedure and those who received e-mails with a generic form of address and no personalized login codes. In addition, with the exception of the indicator of
perceived anonymity, they found no differences in the quality of the survey answers provided by the two groups. As the authors clarify: “the findings of the current study do not support the notion that implementation strategies that involve the use of passwords and personalized forms of address substantially reduce the perceived anonymity of the survey” (Mueller et al., 2014, p. 177). Sauermann and Roach (2013) carried out an experiment with US graduate students and post-doctoral researcher. They found that 48% of the young scientists addressed using their name completed the survey compared with 24% of those addressed using their first and last name. However, when looking at the impact of personalization on measurement error, they found no evidence for social desirability bias, measured with the indicator “self-reported preferences for making a contribution to society”.

2.3 Interactions with other design features

This last strand of research focuses on the evaluation of interaction effects between personalization of salutations and other survey features. Porter and Whitcomb (2003) carried out an experiment with a complex design on high school students aimed to evaluate the effect of personalization of e-mail salutations, e-mail source address, authority of the e-mail signature, and the profile of the requesting office on click-through and response rates. They found no effects of the above mentioned design features on the indicators of response and no evidence for interactions between personalized salutations and the other survey features considered in the analysis. They conclude that “in contrast to research on paper surveys, personalization of the e-mail contact, whether through personalized greetings or a personal e-mail address, appears to have little impact on response rates” (p. 586).

Drawing on Porter and Whitcomb (2003)’s work, Joinson, Reips and co-authors further assessed the impact of personalized e-mail salutations and power of the sender on response and measurement error experimenting with a population of UK students. Contrary to Porter and Whitcomb (2003)’s work, Joinson and Reips (2007, Study 1) found that response increased from 12.4% to 16.6% when respondents were addressed using their name (“dear John”) rather than an impersonal salutation (“dear Student”). In addition, in Study 3 they also found evidence for interactions between type of salutation and nature of the power of the sender. Indeed, the relationship between these two variables is significant only for respondents in the high power condition: 53.4% of respondents in this experimental condition addressed in a personalized manner submitted the first page of the questionnaire compared to 42.1% of those addressed with “Dear Student”. A few years later, Joinson, Woodley and Reips (2007) replicated and extended Joinson and Reips (2007)’s study. Consistently with their previous research, Joinson, Woodley and Reips (2007) found evidence for interactions between type of salutations and the nature of the sender. In addition, when assessing the impact of
personalization on the propensity to disclose sensitive information, they found only a marginally significant association between personalization and disclosure to sensitive questions \((p = .10)\). The authors conclude that “there is some slight evidence that a personalized salutation leads to reduced disclosure to a sensitive personal question, in the form of selection of an “I prefer not to say” option” (p. 279).

Other important studies on these topics are the ones carried out by Sánchez-Fernández, Muñoz-Leiva and colleagues (Muñoz-Leiva et al. 2010; Sánchez-Fernández, Muñoz-Leiva, Montoro-Ríos 2012). The focus of their research is on the evaluation of the impact of interaction effects between personalization, frequency of reminders and incentives on response and measurement error. Experimenting with Spanish university students, Muñoz-Leiva et al. (2010) found a positive effect of personalization on login and retention rates, no evidence for interaction effects between personalization and frequency of reminders and no association between personalization and missing data and number of blocks answered. Regarding retention rates, for example, they found that 61.0 % of respondents in the “personalized” condition completed the survey compared to 52.0 % of those who were assigned to the “anonymous” condition. These findings were partially confirmed in a study that Sánchez-Fernández, Muñoz-Leiva, Montoro-Ríos (2012) published a few years later. Experimenting with a population of members of a Spanish Internet Panel representative of the general population, they found a small but significant effect of personalization on response. When personalized messages were used to invite panel members to take part in the survey, the number of respondents reaching the end of the questionnaire was higher (9.4%) than when anonymous mailings were used (8.6%). More importantly, they found a significant interaction effect between personalization and frequency of reminder, i. e., personalized e-mails sent on a weekly basis reached higher response rates than the other experimental conditions. When assessing the impact of personalization on measurement error, they found no evidence that personalization impacts missing data. Although not central for this literature review, Trespalacios and Perkins (2016) carried out an experiment with a population of US university students to assess the joint effect of the amount of e-mail personalization and information provided on the survey topic on response. Contrary to the studies described so far, the authors did not modify the salutations that remain unchanged amongst the four experimental groups. Their analysis showed that “sending a personal or impersonal e-mail with low or high level of survey information did not differ response rates significantly” (p. 332). Importantly, they also found no selection bias regarding main socio-demographic variables.

In conclusion, the review of the main studies in this field has shown that:
• Personalized salutations in e-mails lead to higher response rates or, in the worst case scenario, have no effects on survey participation. Indeed, we are not aware of any studies that found that personalized salutations lead to lower response rates;
• Although the research evidence with this regard is not always consistent, personalized salutations do not seem to impact data quality and lead to social desirability bias;
• Although research findings are quite mixed, the effect of personalizing salutations in e-mails on response rates may be dependent on other survey design features such as the status/power of the sender or the frequency of reminder. In addition, there is little evidence for interaction effects between personalization and other survey features when looking at data quality indicators.
• There is some evidence that personalization may have a different impact on respondents with different socio-demographic characteristics.

Interestingly, this review has also shown that some of these findings are independent of cultural differences and the nature of the study population. Indeed, the studies are carried out in different nations and on different types of populations (students, researchers, employees, etc.)

3. Research questions

This paper sits within this empirical context. The aim of this work is to assess the impact of personalized salutations in SMS on response and measurement error. Consistently with previous works, we assume that personalizing salutations in SMS is positively related to survey response. In addition, we posit that it does not have an impact on survey quality. This work also aims to test whether personalized salutations are more effective on respondents with specific socio-demographic characteristics.

4. Data

In this paper we consider three data sources: a national web survey on Italian graduates, experimental data from a follow-up study and administrative data.
4.1 The Study on Labour Market Outcomes of Italian Graduates in Social Work

We use data from a national web survey on labor market outcomes of Italian students who graduated in Social Work between 2006 and 2012, conducted in 2013-2014. Twenty-one of the forty-three Italian courses on Social Work took part in the study (that is, 59% of all Italian students who graduated in Social Work between 2006 and 2012). Eligible study members (N=6,294) were randomly allocated to the three experimental groups, the E-mail, SMS and No Reminder groups. The questionnaire collected information on three main areas: respondents’ socio-demographic characteristics, characteristics of the first and current employment, and evaluation of the quality of the university courses.

4.2 The follow-up study

The follow-up study aimed to collect updated information on sample members’ characteristics of the current employment. It was carried out in 2015 on the SMS experimental group only. We checked for differences in the sample composition between the main study and the follow-up and we found no evidence for bias.

All graduates were randomly allocated to two experimental groups: the Personalized salutation and the Generic salutation groups. The first group received an invitation with a personalized salutation. The text of the SMS was: “Dear [First name], we’re conducting a follow-up study on graduates in Social Work. Please join you too! Here is the link to the 6 questions: [URL]”. The second group received an invitation with a generic salutation. In this case the text of the SMS was: “Dear graduate, we’re conducting a follow-up study on graduates in Social Work. Please join you too! Here is the link to the 6 questions: [URL]”. The sender of both types of SMS was Uni-Bicocca (short for Università di Milano-Bicocca).

To send the SMS we used an online service (https://www.smshosting.it/it) that allows to personalize the text of the messages and provides a report on the status of the SMS messages (i.e. delivered or not delivered). Non-respondents received three reminders.

After sending the SMS invitation 133 out of the 708 study members were excluded from the analysis as they did not receive the SMS. After checking for differences in main socio-demographic characteristics\(^1\) between the two groups, we found very little evidence for selection bias.

\(^1\) We considered sex, type of secondary school, year of birth, high school score, graduation year, graduation score, distinction, geographic area of residence and of the location of the university. We found statistically significant differences between the two groups on the two geographic variables: the Southern graduates tend to have more invalid mobile telephone numbers and the North-Eastern graduates seem to have more valid mobile telephone numbers.
4.3 Administrative data

University administrative offices provided graduates’ contact details including their telephone numbers and e-mail addresses (when available), their socio-demographic characteristics and key information on their university performance.

5. Methods

5.1 Analysis
To pursue our aims we mainly perform bivariate analysis and compute appropriate statistical tests to check for statistically significant differences between groups (i.e., Pearson’s Chi-Square Test for Independence). For the analysis of response and measurement error, outcome variables are, respectively, response rates computed as AAPOR RR2 (AAPOR 2016), and two indicators of data quality. For the analysis of the impact of personalization on different respondents’ characteristics we use both socio-demographic variables and data collected by the main study on Labour Market Outcomes of Italian Graduates in Social Work. The key independent variable is the variable that indicates the two treatment groups (Personalized and Generic groups) and all the analyses are carried out on respondents’ sample.

5.2 Indicators of measurement error
We consider response speed and the number of answers provided to open-ended questions as indicators of data quality.
Response speed is computed as a dummy variable that identifies respondents who answer before/after the first reminder. The indicator “answers to open-ended questions” is also a dummy variable, computed on the basis of the following question: “Could you please specify the year in which you started to work as […]?” This indicator identifies respondents who provided an answer to this question.

5.3 Indicators of respondents’ characteristics
We focus on the following indicators: response rate to the main survey, graduation year and score, and area of residence. The first variable is from the dataset of the main survey, the others, from the administrative dataset.
6. Results

6.1 The effect of personalization on response

We compare response rates between Personalized and Generic salutation groups. Results are shown in Table 1. We find that sample members who received a personalized invitation are more likely to take part in the survey (17.5%) than people receiving a generic invitation (12.1%). The difference of 5.4 percentage points between the two groups is statistically significant at the 10% level.

Table 1
Response rate by experimental group

<table>
<thead>
<tr>
<th>Experimental group</th>
<th>Response rate</th>
<th>Total (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalized salutation</td>
<td>17.5</td>
<td>285</td>
</tr>
<tr>
<td>Generic salutation</td>
<td>12.1</td>
<td>290</td>
</tr>
<tr>
<td>Total</td>
<td>14.8</td>
<td>575</td>
</tr>
</tbody>
</table>

Note. $x^2 = 3.420$, df 1, $p = 0.064$

6.2 The effect of personalization on measurement error

To assess the impact of personalization on data quality, we consider two variables: response speed and answers provided to open-ended questions. The analysis on time of response is carried out comparing response rates before and after the first reminder. As shown in Table 2, sample members who received a personalized invitation tend to answer more quickly (54.0%) than those receiving a generic invitation (42.9%). Even if the personalization seems to have a positive effect on response speed, the differences between the two experimental groups are not statistically significant. When we focus on the second indicator of measurement error, we find that 72% of respondents who received the personalized SMS provided an answer to the open-ended question compared to 60% of those belonging to the Generic group. However, note that also in this case the relationship between the two variables is not statistically significant. This part of the analysis is hampered by the small sample size.

---

2 We checked for differences in the sample composition between respondents and non-respondents and we found no evidence for bias.
Table 2

*Response speed and answers provided to an open-ended question by experimental group (%)*

<table>
<thead>
<tr>
<th>Experimental group</th>
<th>Personalized salutation</th>
<th>Generic salutation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response speed</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR before 1&lt;sup&gt;st&lt;/sup&gt; reminder</td>
<td>54.0</td>
<td>42.9</td>
<td>49.4</td>
</tr>
<tr>
<td>RR after 1&lt;sup&gt;st&lt;/sup&gt; reminder</td>
<td>46.0</td>
<td>57.1</td>
<td>50.6</td>
</tr>
<tr>
<td>Total (N)</td>
<td>50</td>
<td>35</td>
<td>85</td>
</tr>
<tr>
<td><strong>Specification of the year in which the respondent started to work (“yes”)</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>72.0</td>
<td>60.0</td>
<td>66.7</td>
</tr>
<tr>
<td>Total (N)</td>
<td>25</td>
<td>20</td>
<td>45</td>
</tr>
</tbody>
</table>

<sup>a</sup> $\chi^2 = 1.023, df 1, p = 0.312$

<sup>b</sup> $\chi^2 = 0.720, df 1, p = 0.396$

6.3 The effect of personalization on respondents’ characteristics

To assess whether personalization has a differential impact on respondents, we focus on four characteristics, participation to the main study, graduation year and score, and area of residence. With the exception of the variable graduation score, we find some indication that personalization may “attract” different types of respondents, i. e., those who participated to the main survey, who graduated recently, and are from the Centre or the South of Italy. For example, 64% of those who received a personalized salutation graduated between 2009 and 2012 compared to 57.1% of respondents belonging to the Generic group. The Chi-Square Tests, however, are not significant.

Table 3

*Respondents’ characteristics by experimental group (%)*

<table>
<thead>
<tr>
<th>Experimental group</th>
<th>Personalized salutation</th>
<th>Generic salutation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Response rate to the main survey</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46.0</td>
<td>40.0</td>
<td>43.5</td>
</tr>
<tr>
<td><strong>Graduation year</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006-2008</td>
<td>36.0</td>
<td>42.9</td>
<td>38.8</td>
</tr>
<tr>
<td>2009-2012</td>
<td>64.0</td>
<td>57.1</td>
<td>61.2</td>
</tr>
<tr>
<td><strong>Graduation score</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>up to 100</td>
<td>38.8</td>
<td>40.0</td>
<td>39.3</td>
</tr>
<tr>
<td>more than 100</td>
<td>61.2</td>
<td>60.0</td>
<td>60.7</td>
</tr>
</tbody>
</table>
7. Conclusions

Nonresponse and measurement errors are very serious issues in survey research. The personalization of survey materials may contribute to reduce the impact of nonresponse, as shown in different papers. However, personalization can also have an impact on measurement error. With personalization some of the assumptions of standardization are violated and this may have a negative impact on data quality. Our paper focuses on an unexplored research field: the impact of personalized salutation in SMS invitation on response and measurement error and on respondents’ characteristics.

Consistently with previous research in this field, we found that personalization increases response rates and this relationship is significant at 10% level. Contrary to our hypothesis, the analysis on indicators of data quality (response speed and answers provided to an open-ended question) offers some preliminary indication that personalization may have an impact on measurement error. Unfortunately, the low occurrence of inaccurate response behaviors and the low response rates reduced the number of cases available for the analysis and therefore the Chi-Square Tests are not statistically significant. Finally, the analysis of respondents’ sample composition shows some evidence that personalization may have a different impact on respondents with different characteristics. We found that personalization tends to have a major influence on those who participated to the main survey, who graduated recently, and are from the Centre or the South of Italy. The sense of belonging to university context could be a reason for the positive influence of personalized salutation on participation to the main survey and on recently graduated students, while cultural differences could explain why personalization is more effective on respondents from the Centre or the South of Italy. The small sample size for this analysis is the major limitation of this part of the paper. Overall, we can speculate that personalization, that is a violation of the principle of standardization, may have a detrimental effect on data quality.

Further studies based on larger sample and higher response rates may contribute to strengthen the findings obtained with our study on Italian graduates.
We conclude this paper with a consideration on costs. To the best of our knowledge, previous research focused on the e-mail personalization. Sending personalized e-mails may be costly, due to the implementation time of personalizing e-mail salutations. Sending personalized SMS, on the other hand, may be cheaper as the implementation time is reduced. However, researchers need to factor in the cost of sending SMS, that, in our case, amounted to 0.081€ per SMS.
References


