The invitation methods of web surveys have been studied from various perspectives, but less is still known about how demographic factors affect the selection of entry options in mailed web survey invitation letters. In a postal invitation letter, we provided the following three options to enter our web survey: using the URL address, emailing the researcher to get a link to the web survey, and texting one’s email address to the researcher to get a link to the web survey. The results of the multinomial logistic regression model show that the odds of selecting the option “Response link by email” is 4.1 times higher for those who have a primary education than for those who have an upper secondary education. In addition, an increase of one year in the respondent’s age increased the odds of selecting the “Response link by email” option by approximately 5%. In conclusion, older and less educated people tend to select less cognitively burdening entry options.

Introduction

Web-based surveys are a cost-effective method to collect responses in comparison to other survey modes. They are also the primary method in mixed-mode designs (web-push), typically offering alternative modes of subsequent data collection (Dillman 2019). The importance of mixed-mode designs is increasing due to the precipitous drop in response rates of telephone surveys for both landline and cell phone frames (Olson et al. 2020). It is important to note that mixed-mode design refers both to multiple contact modes and to multiple modes of data collection (De Leeuw 2018). Different mode combinations produce varying results in terms of coverage, the nonresponse rate, and the quality of measurements (De Leeuw 2018). The best results can be achieved by combining the primary features of different modes. This is the case, for instance, when combining the self-administered and interview modes by applying the self-administered mode to sensitive questions and the interview mode to other, often more complex questions (De Leeuw 2018).

Even if statistically representative data is pursued at the general population level, the invited individuals must be identified regardless of the mode of the survey. In some countries, the identification can be based on the population registry representing a sampling frame of the general population. As the contact information offered by the population registry is typically a postal address, the invitations to a web survey must be delivered by mail using a postcard or a letter. Postal delivery can further be combined with the Internet, email, the mobile phone, and the landline phone (Dillman 2019; Dillman, Reips, and Matzat 2010).
Despite the extensive research on mixed-mode design, less is known about the layout design of mailed invitation letters and the demographic factors affecting the selection of entry options in the invitation letter. Our study extends the research of mixed-mode design in these directions. In this respect, it is essential to understand that the self-selection of entry options has the potential to affect the nonresponse bias, which is a function of both the nonresponse rate and the difference between respondents and nonrespondents.

**Previous research and research hypotheses**

Web survey invitation procedures have been studied at least in terms of prenotification (De Leeuw et al. 2007; Dillman, Smyth, and Christian 2009), invitation to a PC survey (Bandilla, Couper, and Kaczmirek 2012; Kaplowitz et al. 2012), and invitation to a mobile survey (Mavletova and Couper 2014). The previous research has also focused on the design of mailed invitations and advance letters, showing that they affect the response rates in both mail surveys and interviewer-administered surveys (Dillman 2007; Hembroff et al. 2005). Attention has been focused mostly on how the invitation mode (email / postal letter / postcard) affects participation in web surveys. This is due to a growing number of problems with email as an invitation mode, including problems with spam, the rapid turnover (churn) of email addresses, privacy in providing an email address when solicited, and the ability to deliver a prepaid incentive (Bandilla, Couper, and Kaczmirek 2012).

Bandilla, Couper, and Kaczmirek (2012) experimentally varied both the invitation mode and the prenotification mode in a general population sample. They found that a mailed letter, whether a prenotification or invitation, is more effective than an email alone. If the recipients were not prenotified, the response rate in a web survey with a mailed invitation was 51%, being 40% with an email invitation. These findings are supported by studies showing that a postcard and an email invitation together work better than an email invitation alone (Kaplowitz et al. 2012; Porter and Whitcomb 2007). A proposed reason why recipients are more likely to overlook a prenotification sent via email than one sent via traditional communication channels is that in traditional survey modes, the researcher’s tangible investment in contacts is perceived to indicate the importance and legitimacy of the survey (Daikeler, Bošnjak, and Lozar Manfreda 2020).

While progress has been made in the research of invitation procedures, there is not much research focusing on how entry options in invitation letters are addressed and selected by invitees in terms of their demographic backgrounds and other characteristics. Nevertheless, preference studies can be applied in the formulation of hypotheses with regard to the selection of entry options. Diment and Garrett-Jones (2007) found that when the respondents were given the choice to complete either a web or a paper questionnaire, the majority chose the paper version. The web respondents tended to be young, male, middle-ranking, and working in the information technology sector. Similarly,
Bech and Kristensen (2009) found that a significantly high proportion of 65 to 75-year-old individuals responded to a postal version instead of a web survey. These results are supported by the findings of Parast et al. (2019) showing that from telephone, mail, and web respondents, the telephone respondents were most similar to the sampled population of the emergency department patients in terms of the age. Compared to mail respondents, web respondents tended to be younger. According to Smyth et al. (2010), these types of demographic differences between web and mail respondents can be explained by the computer use factor that eventually affects the choice of response mode.

Regarding the order of multiple modes offered to the recipients, two designs are commonly applied, the sequential design and the concurrent design. In the concurrent design, the different modes are offered simultaneously; in the sequential design, the web option is typically offered first. Mauz and colleagues (2018) found the sequential design and the concurrent design equivalent in terms of response rates and sample composition while they also found support for the Millar and Dillman (2011) observations, whereby web response rates may increase if the web option is offered first in the sequential design. These results are supported by the findings of Suzer-Gurtekin et al. (2020) who divided random samples of recipients into groups that received either a concurrent request to complete the survey by mail or web or a web-intensive request to complete the survey by web before offering a mail alternative, an approach equivalent to the sequential design. They found the web-intensive approach shifted mail respondents to the web mode, yet did not attract different subgroups of people to participate in the web survey who would not have participated in a mail survey.

To summarize the previous research, it seems that those who are relatively young, educated, and experienced in information technology prefer a web survey over a paper one. Overall, they are individuals who are somewhat familiar with cognitively demanding tasks. Based on this, older and less educated individuals are likely to select a cognitively less burdening entry option in the invitation letter.

**Method**

In the beginning of August 2014, a three-stage stratified sample of 1,329 persons received a postal advance letter regarding participation in a web survey on the Finnish competence-based education system. The sample frame consisted of individuals who had completed a Finnish competence-based education program in 2007. The ethics committee of the Finnish National Agency for Education approved the study.

The web survey questions covered various topics ranging from the benefits received from the education program to one’s living conditions after the program. The questionnaire had 27 questions including single rating questions, matrix questions, and open-ended questions.
sampled persons to the web survey, the following three entry options were presented in the invitation letter, from which the recipients were asked to select one to participate:

1. Write the following address in your web browser’s address bar and press “enter”: http://www.webropolsurveys.com/oph.net. Then write the password “näyttö” in the input field and press the “Entry” button.

2. Send an email with the heading “Näyttötutkinto” to the following address: arto.selkala@ulapland.fi. No content is needed in the message; the heading will suffice. Having sent the message, you will receive an email containing the response link to the survey. Click on the link to participate in the survey.

3. On your phone, text your personal email address to the following number: XXX-XXXXXXX. You will then receive an email with a link to the survey.

The recipients received the first postal invitation letter on August 1, 2014. A reminder letter with corresponding information was sent on August 15, 2014. The researchers’ email addresses and phone numbers were listed in the invitation letter (see Appendix) for requests of information about the study. No incentives were offered for completing the survey. The web survey was available only in Finnish and it was closed on August 24, 2014.

Data

Of the persons (1,329) who received the advance letter, 296 (22%) selected the first entry option (typing the URL in the web browser’s address bar), 54 (4%) the second option (email), and 26 (2%) the third option (SMS) (Table 1). Altogether 950 persons (72%) did not participate. Three persons selected both the email and the SMS option; their data were removed from the final data set because of the ambiguous response. The difference between the groups was statistically significant at the p<0.001 level when the expected proportions between the categories were set equal (X² = 1671.611, d.f. = 3). Of the 1,329 invited individuals, 359 eventually participated in the web survey, of whom 296 persons selected the URL option, 42 the email option, and 21 the SMS option. The response rate of the survey was 27.0% (American Association for Public Opinion Research (AAPOR) 2008: RR2 definition). Table 1 shows the distribution of the cases and the response rate in the web survey.

According to the research regarding the association of age with cognitive functioning, some cognitive abilities such as spatial orientation and perceptual speed remain relatively constant until a person’s mid-40s, after which their decline starts to gain speed (Hedden and Gabrieli 2004; Schaie, Willis, and
Table 1. Characteristics of the sample and web survey participants

<table>
<thead>
<tr>
<th>Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross sample size</td>
<td>1329</td>
</tr>
<tr>
<td>Selection of the type of entry into the web survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typing the URL</td>
<td>296</td>
<td>22.3</td>
</tr>
<tr>
<td>Response link by email</td>
<td>54</td>
<td>4.1</td>
</tr>
<tr>
<td>Response link by SMS</td>
<td>26</td>
<td>2.0</td>
</tr>
<tr>
<td>No selection</td>
<td>950</td>
<td>71.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Web survey participants</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondents and response rate</td>
<td>359</td>
<td>27.0</td>
</tr>
</tbody>
</table>

Selection of the type of entry of web survey participants

| Typing the URL | 296 | 82.5 |
| Response link by email | 42 | 11.7 |
| Response link by SMS | 21 | 5.8 |

<table>
<thead>
<tr>
<th>Age</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20–44 years old</td>
<td>154</td>
<td>43.4</td>
</tr>
<tr>
<td>45–54 years old</td>
<td>116</td>
<td>32.7</td>
</tr>
<tr>
<td>55–70 years old</td>
<td>85</td>
<td>23.9</td>
</tr>
<tr>
<td>(Missing)</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>226</td>
<td>63.3</td>
</tr>
<tr>
<td>Male</td>
<td>131</td>
<td>36.7</td>
</tr>
<tr>
<td>(Missing)</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary education</td>
<td>54</td>
<td>15.3</td>
</tr>
<tr>
<td>Lower secondary education</td>
<td>171</td>
<td>48.4</td>
</tr>
<tr>
<td>Upper secondary education</td>
<td>128</td>
<td>36.3</td>
</tr>
<tr>
<td>(Missing)</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Note. The response rate was calculated by AAPOR RR2 definition (2008). “Education” refers to nonvocational education in accordance with the International Standard Classification of Education 2011.

Caskie 2004; Singh-Manoux et al. 2012). Given that cognitive abilities have the potential to affect the selection of entry options, we categorized the age variable as follows: 20–44, 45–54, and 55–70 years.

Results

Using the final data on the web survey participants, we compared the conditional distribution of the selection of entry options by demographic factors (Table 2). The results show that the difference between the selections made by male and female respondents was not statistically significant ($X^2 = 1.444$, d.f. = 2, $p = 0.486$). A comparison between age categories revealed that the URL option was selected most often by the youngest respondents (20–44 years, 89%, n = 154), second most often by the middle category (45–54 years, 81%, n = 116), and least often by the oldest respondents (55–70 years, 73%, n = 85). The email option was selected most often by the oldest respondents (55–70 years, 21%), then by the middle category (45–54 years, 12%), and least often (7%) by the youngest respondents (20–44 years). The SMS option was
Table 2. Web survey entry option selections by demographic factors

<table>
<thead>
<tr>
<th>Web survey entry options</th>
<th>Typing the URL</th>
<th>Typing the URL</th>
<th>Response link by email</th>
<th>Response link by email</th>
<th>Response link by SMS</th>
<th>Response link by SMS</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Sex ($X^2 = 1.444, d.f. = 2, p = 0.486$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>186</td>
<td>82.3</td>
<td>25</td>
<td>11.1</td>
<td>15</td>
<td>6.6</td>
<td>226</td>
<td>100</td>
</tr>
<tr>
<td>Male</td>
<td>109</td>
<td>83.2</td>
<td>17</td>
<td>13.0</td>
<td>5</td>
<td>3.8</td>
<td>131</td>
<td>100</td>
</tr>
<tr>
<td>Age ($X^2 = 12.4, d.f. = 4, p &lt; 0.05$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–44 years old</td>
<td>137</td>
<td>89.0</td>
<td>10</td>
<td>6.5</td>
<td>7</td>
<td>4.5</td>
<td>154</td>
<td>100</td>
</tr>
<tr>
<td>45–54 years old</td>
<td>94</td>
<td>81.0</td>
<td>14</td>
<td>12.1</td>
<td>8</td>
<td>6.9</td>
<td>116</td>
<td>100</td>
</tr>
<tr>
<td>55–70 years old</td>
<td>62</td>
<td>72.9</td>
<td>18</td>
<td>21.2</td>
<td>5</td>
<td>5.9</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>Education ($X^2 = 20.653, d.f. = 4, p &lt; 0.001$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary education</td>
<td>38</td>
<td>70.4</td>
<td>14</td>
<td>25.9</td>
<td>2</td>
<td>3.7</td>
<td>54</td>
<td>100</td>
</tr>
<tr>
<td>Lower secondary education</td>
<td>138</td>
<td>80.7</td>
<td>20</td>
<td>11.7</td>
<td>13</td>
<td>7.6</td>
<td>171</td>
<td>100</td>
</tr>
<tr>
<td>Upper secondary education</td>
<td>118</td>
<td>92.2</td>
<td>6</td>
<td>4.7</td>
<td>4</td>
<td>3.1</td>
<td>128</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. In a contingency table of the web survey entry options by sex, 0 cells have an expected count of less than 5. The minimum expected count is 7.34. In a contingency table of the web survey entry options by age, 1 cell (11.1%) has an expected count of less than 5. The minimum expected count is 4.79. In a contingency table of the web survey entry options by basic education, 1 cell (11.1%) has an expected count of less than 5. The minimum expected count is 2.91.

selected most often by the middle category (7%), second most often by the oldest respondents (6%), and least often by the youngest respondents (5%). The difference between the groups was statistically significant at the p < 0.05 level ($X^2 = 12.4, d.f. = 4$).

The results also show (Table 2) that the difference between the entry option selections made by those who had completed primary education, lower secondary education, and upper secondary education was statistically significant at the p < 0.001 level ($X^2 = 20.653, d.f. = 4$). The URL option was selected most often by respondents with an upper secondary education (92%, n = 128), second most often by those with a lower secondary education (81%, n = 171), and least often by those with a primary education (70%, n = 54). The email option was selected most often by people with a primary education (26%), then by those with a lower secondary education (12%), and least often by those with an upper secondary education (5%). The SMS option was selected most often by respondents with a lower secondary education (8%), second most often by those with a primary education (4%), and least often by respondents with an upper secondary education (3%).

We employed the multinomial logistic regression model to examine the self-selection mechanism behind the web survey entries by age, sex, and education. Given, however, that the education of the participants was measured twofold,
by nonvocational and vocational education, we had to decide which one of the two variables should be preferred. With regard to this issue, when the cross-tabulation of a response variable with a given categorical predictor results in one or more empty cells, it is not possible to estimate the effects associated with those cells in the logistic regression model (DeMaris 2004, 268). This is due to the fact that the maximum likelihood estimation, which is applied in logistic and multinomial regression analyses, does not work properly in the presence of empty cells (Cook, Niehaus, and Zuhlke 2018). Having found empty cells in the cross-tabulation of “Entry option” (response variable) and “Vocational education” (categorical predictor), we decided to predict the selection of web survey entries by nonvocational instead of vocational education. The education complies with the International Standard Classification of Education 2011.

Table 3 provides the results of the multinomial logistic regression model predicting the selection of the “Response link by email” option and the “Response link by SMS” option as compared with typing the URL in the web browser’s address bar (reference category). Education predicted to a significant extent the selection of “Response link by email” as well as the selection of “Response link by SMS,” while sex was a non-significant predictor in both cases. The odds of selecting the “Response link by email” option was 4.2 times higher for those with a primary education and 3.1 times higher for those with a lower secondary education as compared with participants with an upper secondary education (ref. category). The participant’s age predicted to a significant extent the selection of “Response link by email” but not the selection of “Response link by SMS.” An increase of one year in the respondent’s age raised the odds of selecting the “Response link by email” option by approximately 5%. Further, an increase of ten years in age raised the odds by 61% ((exp(10 x β)-1) x 100). As regards the comparison of “Response link by SMS” and “Typing the URL” options, the odds of selecting the “Response link by SMS” option was 3.3 times higher for those with a lower secondary education. In the final stage of the analysis, we applied the multinomial regression model using the same factors excluding sex, given that it was not a significant predictor in the first model. We found the odds ratios as well as their significance levels were almost identical with the previous model presented in Table 3. These results are not included in the article.

Conclusion

In general, the most frequently chosen entry option in the invitation letter was the URL address option, requiring one to insert the URL address in the web browser’s address bar. The second most popular option required one to send an email to the researcher in order to receive the response link to the web survey. The least frequently chosen option was the third alternative, requiring the respondent to send a text message to the researcher to receive the survey link by email. In addition, the selection of the second most popular option, “Response link by email,” could be predicted by age and education as compared with the selection of the URL address option. The results of
the study support our hypothesis. Older and less educated individuals prefer cognitively less burdening entry options in the invitation letter to cognitively more demanding ones, such as inserting the URL address in the web browser’s address bar.

Given that the results attest to a nonresponse bias, it should be noted that the bias has two components: the nonresponse rate and the difference between respondents and nonrespondents (De Leeuw et al. 2007). Therefore, instead of being merely a function of the nonresponse rate, the bias increases as the difference between the respondents and nonrespondents becomes more pronounced (Groves and Peytcheva 2008). Based on our results, we should expect the nonresponse bias to increase in cases where the URL option is offered as the only entry option in the invitation letter. Thus, we recommend researchers to consider offering alternative entry options along with the URL option in order to decrease the nonresponse bias. This becomes especially important in studies on the general population involving a large variety of individuals.
Limitations of the present study and further research

An obvious limitation of the present study is the lack of experimental control of the essential confounding variables. Previous research has shown that the text length of an invitation letter or URL affects the participation rate of surveys and web surveys (Reips and Franek 2004). Therefore, the various sources of cognitive load generated by the entry options become impossible to identify unless they are experimentally controlled (Artino 2008). Another confounding variable that should be controlled is the order of the entry options on the invitation letter. This can be done, for instance, by counterbalancing the entry options in the experimental design (Reips and Krantz 2010; Zeelenberg and Pecher 2014). Age is, as often, confounded with cohort (e.g., Schaie, Willis, and Caskie 2004), so the crucial difference between younger and older people in this study could really have been that the older ones were born at an earlier time in history and thus may have responded differently because they grew up in a different techno-social environment. One should also be cautious of the relatively small sample size of the study, particularly in terms of the subgroup analyses. It should also be noted that the studied invitation method is fully usable only in countries where population registers are available. In countries where only lists of buildings or addresses are available, a within-household selection is required as an additional stage in order to create a representative sample. There is, however, no reason why the invitation method introduced in the present article could not be adjusted to perform in these occasions as well. Future research will certainly delve into further clarifying most practical and usable options in entry selection, for example in comparing the presently researched options to the recently introduced option of scanning quick response (QR) codes in devices with cameras. Invitation methods of web surveys do have a strong impact on participation and quality of data and will thus remain a focus of investigation.

Acknowledgements

The authors are grateful for the funding provided by the Finnish National Agency for Education.

Corresponding Author:
Arto Selkälä, PhD
University of Lapland, P.O. Box 122, FI-96101 Rovaniemi, Finland
E-mail: arto.selkala@ulapland.fi
Phone: +358 45 6351256

Submitted: October 28, 2020 EDT, Accepted: January 26, 2021 EDT
REFERENCES


APPENDIX

1.8.2014

Arvoisa vuonna 2011 näyttötutkinnon suorittanut!

Opetushallitus rahoittaa tutkimushanketta, jonka avulla selvitetään näyttötutkintojärjestelmän vaikuttavuutta ja pyritään kehittämään näyttötutkintojärjestelmää entistä paremmaksi. Teidät on satunnaisesti valittu näyttötutkinnon suorittaneiden joukosta vastaajaksi tutkimuksen kyselysionon. Osallistumiseenne tutkimukseen on erittäin tärkeää tutkimuksen onnistumisen kannalta. Koska kyselyyn osallistujat on poimittu satunnaisesti, mahdollista poisjäämistä tei ei voida korvata muilla näyttötutkinnon suorittaneilla.

Hankkeen vastuullisena johtajana toimii professori Antti Syväjärvi Lapin yliopistosta sekä tutkijoina emeritusprofessori Asko Suikkanen ja tilastotieteen lehtori Arto Seikälä Lapin yliopistosta sekä yliopettaja Leena Viinanmäki Lapin ammattikorkeakoulusta.


Volitte valita seuraavista vastausvaihoista itseellesen sopivimman:


b) Lähetätte sähköposti otsikolla ”Näyttötutkinto” seuraavaan sähköpostiosoitteeseen: arto.seikal@ulapland.fi Viestissä ei tarvitse olla mitään sisältöä, vain otsikko riittää. Tämän jälkeen saatte omaan sähköpostinne viestin, joka sisältää vastauslinkin kyseisyn. Vastauslinkkiä klikkaaamalla pääsette vastaanaan kyselyyn.

c) Lähetätte oma sähköpostiosoitteen tekstiviestillä seuraavaan puhelinnumeroon: 045-6351256, jonka jälkeen saatte sähköpostinne viestin, joka sisältää vastauslinkin kyseisyn.


Kiitämme kyselyyn vastaamisesta jo ennakoon ja toivotamme Teille mukavaa alkuysyyn aikaa.

Lisätietoja allekirjoittaneilta tai Opetushallituksesta: Opetusneuvos Seppo Hyppönen (seppo.hypponen@oph.fi) tai tekn. lis. Matti Ropponen (Matti.Ropponen@oph.fi).

Arto Seikälä
Arto Seikälä
tilastotieteen lahtori, YTT

Leena Viinanmäki
Leena Viinanmäki
Lapin ammattikorkeakoulu

Asko Suikkanen
Asko Suikkanen
emeritusprofessori, YTT

Antti Syväjärvi
Antti Syväjärvi
professori, PhD, HTT

Lapin yliopisto
Lapin yliopisto
asiko.suikkanen@ulapland.fi

leena.viinanmaki@lapinamk.fi

leena.viinanmaki@lapinamk.fi

aso.suikkanen@ulapland.fi

aritt.syvaarvi@ulapland.fi

Puh. 045 6351 256
Puh. 0400 142 445
Puh. 0405 5442 744
Puh. 040 606 244

---

1 Verkkoselain tarkoittaa ohjelmaa, jolla mennään nettiin, esim. Internet Explorer tai Google Chrome.