

Embedding the first question in the e-mail invitation: THE EFFECT ON WEB SURVEY RESPONSE

Introduction:

Low response rates in web surveys are a challenging issue. Researchers explore several response inducements, e.g. when inviting participants through e-mail. However, we are aware of only two studies that focus on embedded questions in the e-mail invitation (Varghese, Moore and Earnhart, 2018; Liu and Inchausti, 2017). The main aim of this poster is to assess the impact of tailoring the e-mail invitation text on response. In particular, we evaluate the impact of an e-mail invitation that includes the first question of the web questionnaire vs a standard e-mail invitation on questionnaire completion, completion time, break-offs, and respondents' composition.

Methods & Data:

We use experimental data from a web survey conducted in 2018 on delegates of the trade union Italian General Confederation of Labour (CGIL). Sample members (N=5,494) were stratified by geographic area and type of trade-union category, and then they were randomly (within the strata) assigned to two groups: the "link" and the "first question" e-mail invitation group. The text of the e-mail sent to the two groups was different only in the final statement. At the end of the e-mail text (the same for both groups), in the "first question" group, the first question of the questionnaire was reported, while, in the "link" group, the survey link was included. The first question of the questionnaire is "How many years have you joined the trade union Italian General Confederation of Labour?". The delegates' e-mails were not all available at the same time. For this reason the delegates received the e-mail invitation (also called "first stimulus") at four different times, between 22 May and 15 June. To analyze our data we adopt both bivariate and multivariate analysis.

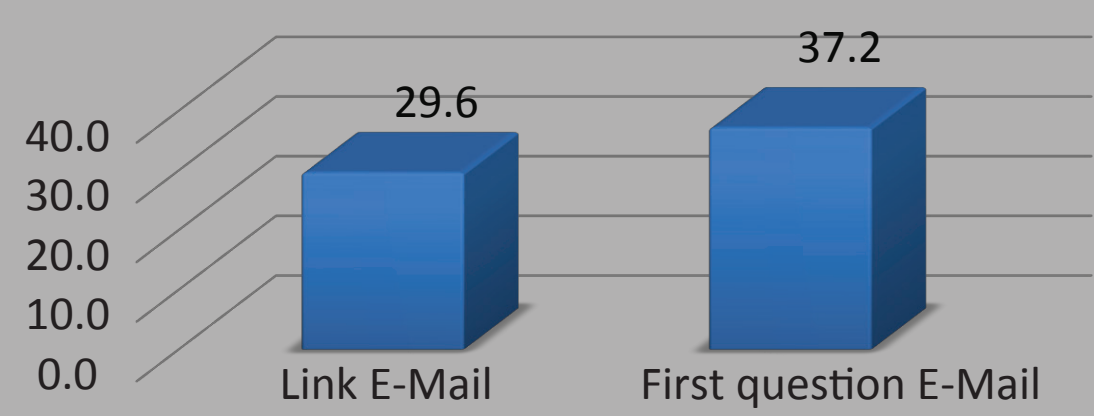
Results:

Four main findings stand out from our analyses:
1. The embedded e-mail invitation is more effective in increasing survey response than the link invitation
2. There are no differences in completion time
3. The number of break-offs is higher in the "first question" group than in the "link" group
4. There are no differences by age group, gender and enrollment to CGIL

Questionnaire completion:

From the bivariate analysis on questionnaire completion we find that the overall response rate (calculated as the number of completed questionnaires divided by the number of the e-mails sent to the initial sample of delegates; we refer to RR1 defined by AAPOR – American Association for Public Opinion Research) is 33%. When looking at the differences between the two experimental groups (Graph 1), the response rate of the "first question" group is higher (37%) than that of the "link" group (30%).

Graph 1. Response rates by type of e-mail invitation.



To investigate the net effect of the experimental group on the response propensity, we perform a binomial regression analysis with a stepwise forward selection of variables. As control variables we use the date of the first stimulus, the presence of the CGIL delegates in Italy (i.e. high, medium and small presence), the geographic area, and the productive sector to which the delegates belong (not statistically significant). Results from the logistic regression model (Table 1) show that, ceteris paribus, the "first question" group has a higher probability to respond to the web survey than the control group (i.e. the "link" group). We speculate that there are two main reasons for these results: i) people who receive the e-mail invitation with the first question embedded in the text message, can immediately identify the e-mail as a web survey invitation and not as a spam e-mail, and ii) the "first question" e-mail invitation may arouse the respondents' curiosity.

Table 1. Results from the logistic regression that estimates the probability of response.

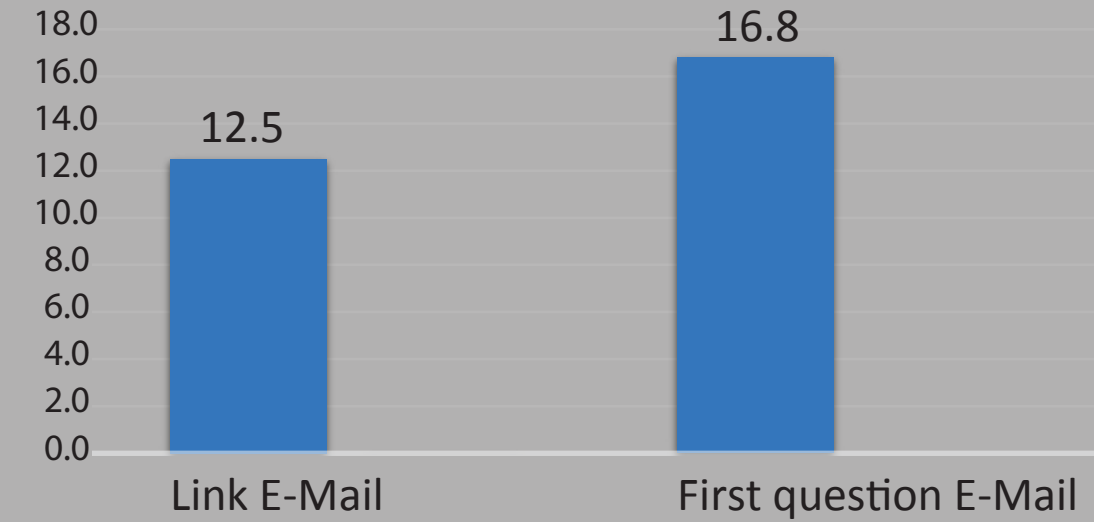
	Estimate	Std. Error	Z-value	Significance
(Intercept)	-0.125	0.105	-1.193	0.233
Invitation time (first stimulus until 23/5)*				
First stimulus from 24/5 to 30/5	0.010	0.086	0.116	0.908
First stimulus from 31/5 to 6/6	-0.574	0.096	-5.989	0.000
First stimulus from 7/6 to 15/6	-0.309	0.094	-3.290	0.001
Type of e-mail invitation (link e-mail)*				
First question e-mail	0.362	0.058	6.216	0.000
Delegates' presence in Italy (small)*				
Medium	-0.513	0.092	-5.547	0.000
High	-0.514	0.101	-5.093	0.000
Geographic area (North)*				
Center	-0.314	0.086	-3.669	0.000
South	-0.319	0.101	-3.154	0.002
Islands	-0.360	0.131	-2.748	0.006

*Category of reference

Break-offs:

As an additional response metric, besides the response rate, we computed the break-offs, that we defined as individuals who start filling out the questionnaire, but do not submit it. Comparing the two experimental groups (Graph 3), we found that break-offs are higher for the "first question" sample (16.8%) than for the "link" sample (12.5%). This difference is statistically significant (p-value < 1%).

Graph 3. Break-offs by type of e-mail invitation.



To assess the net effect of the type of e-mail invitation on break-offs, we perform a binomial regression analysis with stepwise forward selection of variables (Table 3). After adjusting for the effect of invitation time (statistically significant), geographic area (not statistically significant) and delegates' presence in Italy (not statistically significant), we found a statistically significant difference between the "first question" group and the "link" group. In particular, respondents who received the embedded e-mail invitation are more likely not to complete the survey.

Table 3. Results from the logistic regression that estimates the probability of break-off.

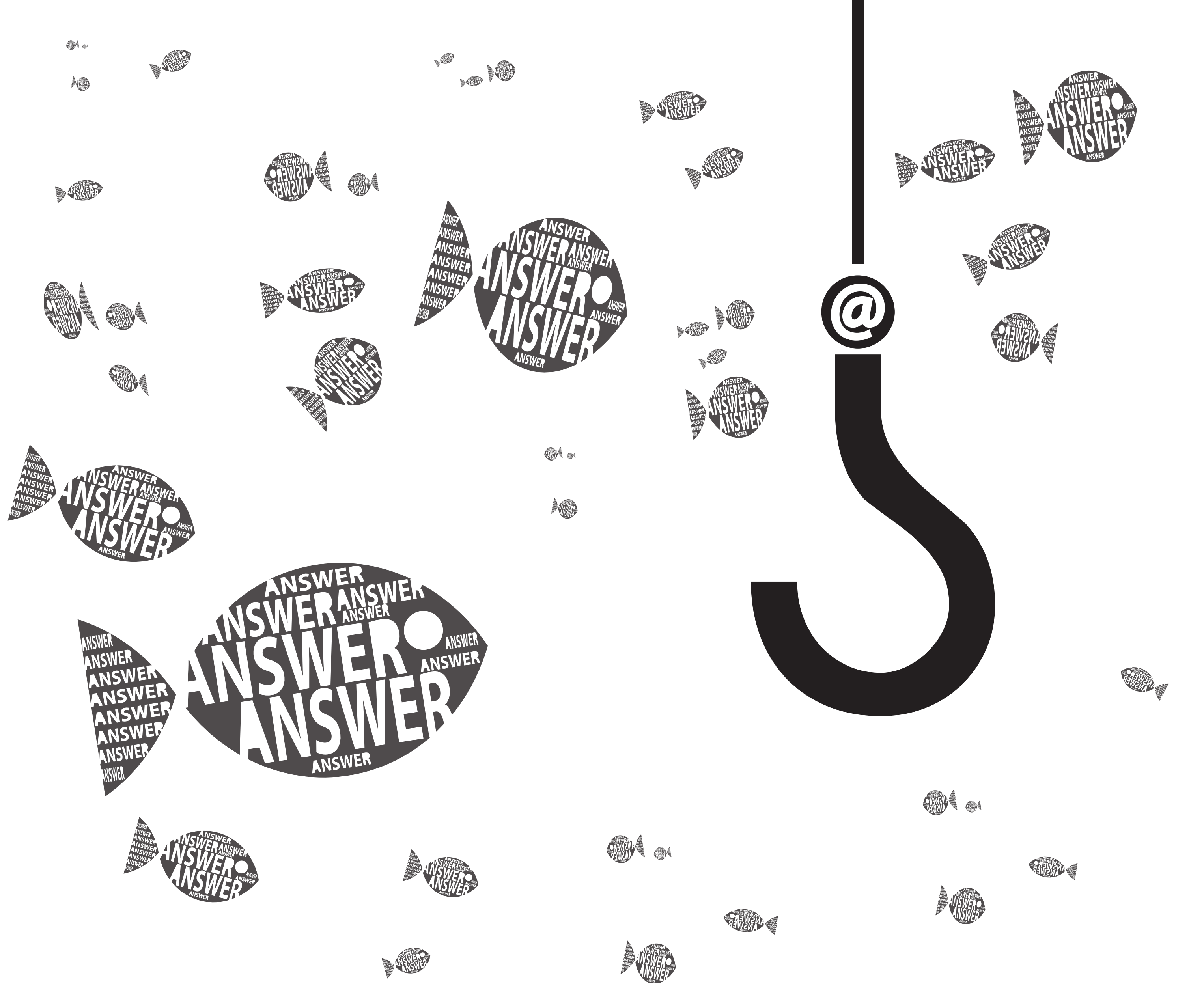
	Estimate	Std. Error	z-value	Significance
(Intercept)	-2.030	0.128	-15.887	0.000
Invitation time (first stimulus until 23/5)*				
First stimulus from 24/5 to 30/5	-0.245	0.176	-1.396	0.163
First stimulus from 31/5 to 6/6	0.401	0.154	2.595	0.009
First stimulus from 7/6 to 15/6	0.212	0.174	1.219	0.223
Type of e-mail invitation (Link e-mail)*				
First question e-mail	0.335	0.126	2.666	0.008

*Category of reference

We speculate that the embedded e-mail invitation is also effective in encouraging "reluctant" respondents to start the questionnaire. Even if some of these respondents do not submit the questionnaire (higher number of break-offs in the "first question" group than in the "link" one), some others do complete the survey, thus increasing the response rate for the "first question" sample.

Conclusions:

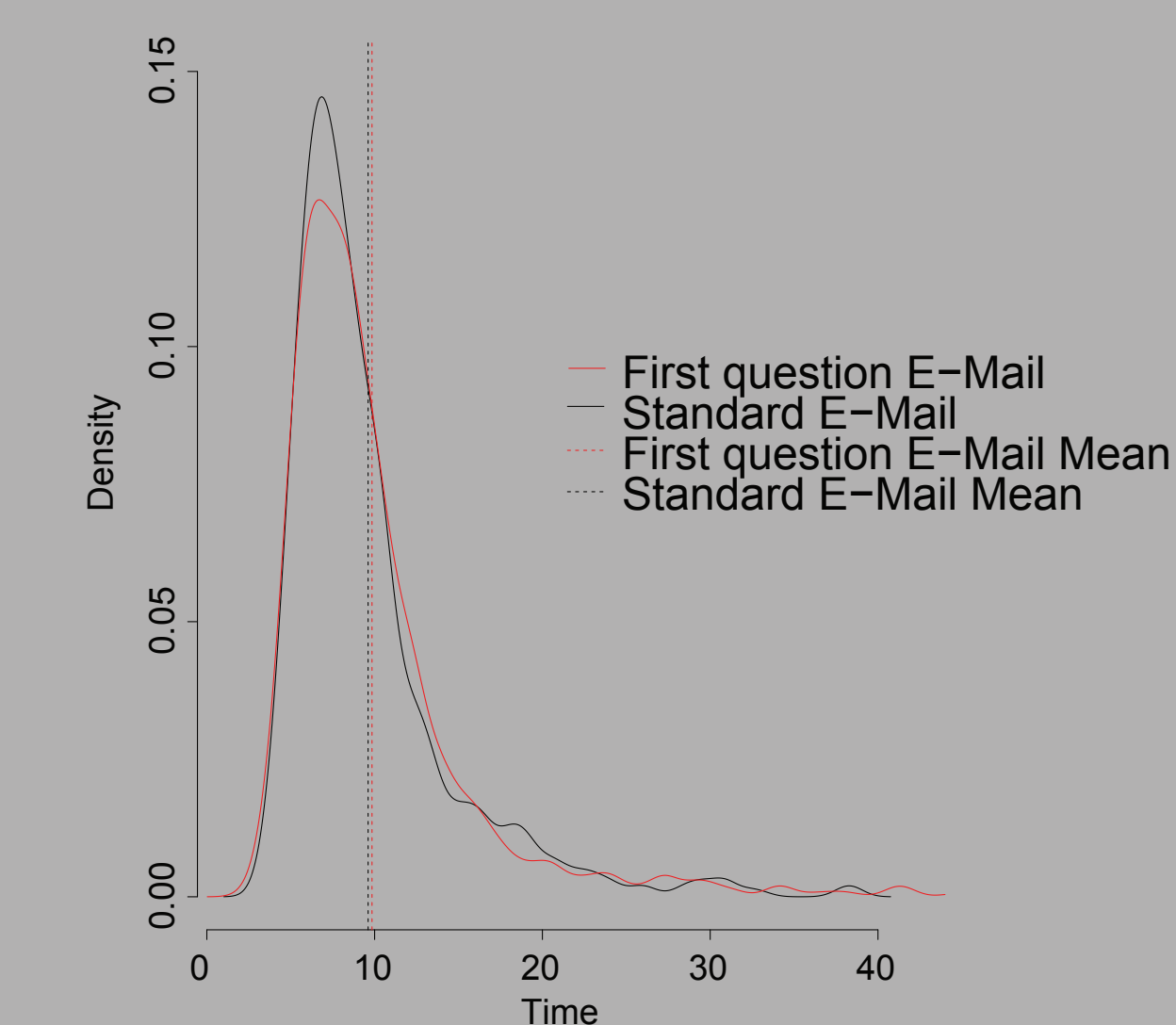
We believe that our work may contribute to expand the knowledge on the effectiveness of embedding a question in the e-mail invitation on response. Indeed, to the best of our knowledge, this is the first study that looks at the impact of the embedded invitation on completion time and respondents' composition.



Completion time:

Another indicator of response quality that we use is the time spent to complete the questionnaire. Graph 2 shows the distribution of completion time (for the web surveys completed within 45 minutes, 91% of the sample) for each e-mail invitation group.

Graph 2. Completion time curve by type of e-mail invitation.



The mean completion time is 9.84 minutes for respondents who received the "first question" invitation, whereas it is 9.61 minutes for those who received the "link" invitation. We computed a statistical test (i.e. t-test) to check for the differences between the two groups and we found no significant differences (p-value=38%). Looking at gender and age group of the respondents (Table 2), we found no differences in completion time.

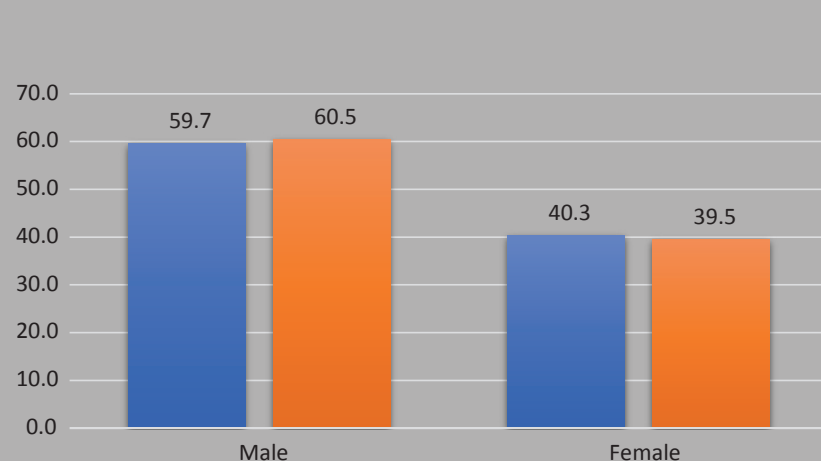
Table 2. Completion time by gender and age group.

	Mean	Standard Deviation
Gender		
Male	9.822	5.402
Female	9.616	5.417
Age group		
18-44	9.286	5.754
45-55	9.665	5.094
56+	10.265	5.546

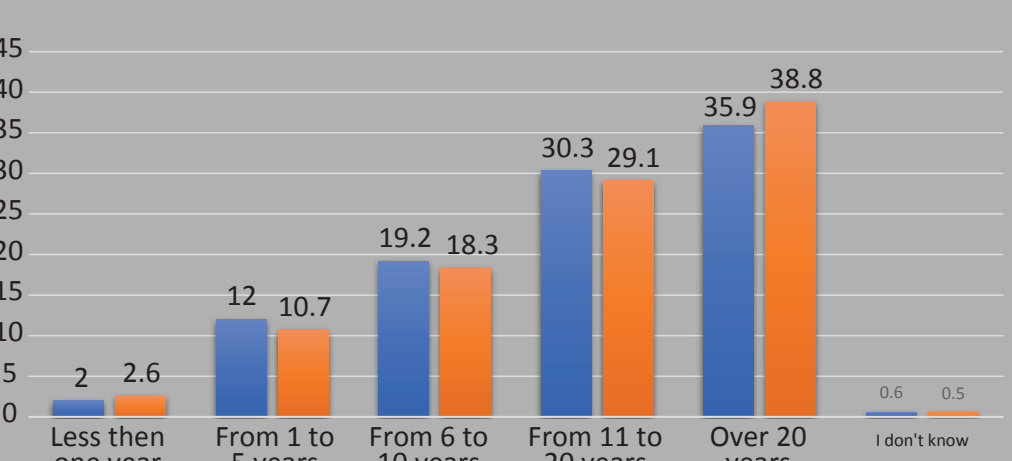
Respondents' composition:

In this section we report results on the respondents' composition of each sample (i.e. the "first question" and the "link" groups). We focused on three variables: gender, years of enrollment in CGIL, and age group. When looking at gender, age group, and enrollment in CGIL, the sample composition is not influenced by the type of e-mail received.

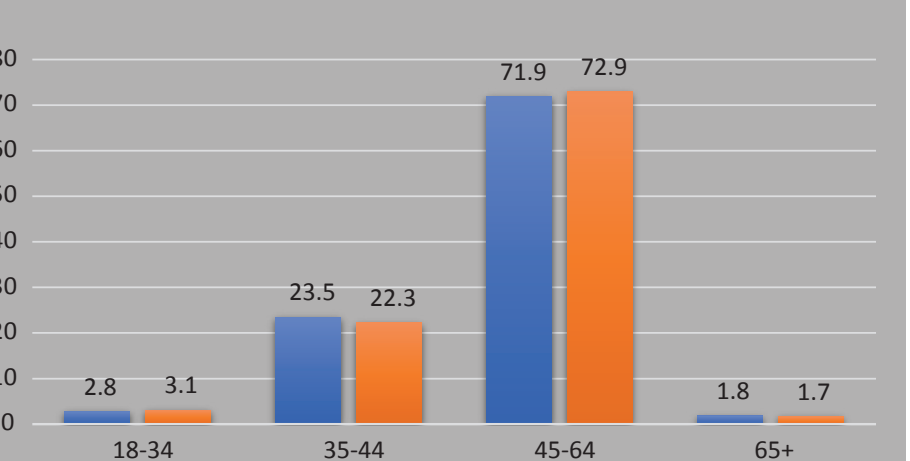
Graph 4: Sample composition by gender



Graph 5: Sample composition by Enrollment to CGIL



Graph 6: Sample composition by age group



References:

Liu, M. and N. Inchausti. 2017. Improving survey response rates: the effect of embedded questions in web survey email invitations. Survey Practice 10(1)

Varghese, Moore and Earnhart, 2018; Are Embedded Survey Items the Solution to Low Web Survey Response Rates? An Investigation of the Interaction Between Embedded Survey Items and Time of Survey Administration. Survey Practice 11(2)