# Experiments with web survey length to inform Split Questionnaire Design – with a sidestep to QR codes

Annemieke Luiten[[1]](#footnote-1)

Statistics Netherlands

## Introduction

For some time now, Statistics Netherlands has embraced the idea of Split Questionnaire Design for longer surveys. Several simulations have been performed, and the first field test is foreseen in 2019. One of the most important reasons to embrace Split Questionnaire Design, is to increase (web) response rates by substantially shortening the questionnaire. There is some indication that shorter questionnaires in web surveys may lead to a higher response rates (Yan, Conrad, Tourangeau and Couper, 2010; Galesic and Bosnjak, 2009; Crawford, Couper and Lamias, 2001), at least if the stated questionnaire length corresponds to the actual length. If the survey is actually longer than the stated length, this may lead to survey break-off. But how short is ‘short’? This question is important in informing us how long the splits should be in our split questionnaires. The literature above experimented with short questionnaires of 5, 8, and 10 minutes, versus long questionnaires of 20, 25, 30, and 40 minutes. Revilla and Ochoa (2017) simply asked respondents how long web questionnaires should be. According to them, the optimal length was about 10 minutes, and the maximal length 20 minutes.

Most Statistics Netherlands’ social surveys have a mixed mode design, starting with web. Web questionnaires are typically around 20 minutes long (range 8 to 33 minutes). Web response rates range from 19 to 46%, depending on topic and the inclusion and kind of incentives. Web response rates of younger persons (16 to 30 years old) lag about 10 percentage points behind those of the group with the highest response rates: the middle aged persons (45 to 65 years old). Although Revilla and Ochoa’s respondents pinpointed 10 minutes as the optimal length, qualitative research with groups of young persons at Statistics Netherlands (CBS) indicated that 5 minutes would be more attractive for this group.

Whether we should strive for 5 minute splits, or if 10 minute splits would also increase response over the standard approach (which does not mention the duration of surveys longer than 10 minutes), was the topic of this experiment. Four different time conditions were used, as is explained in the next section.

Beside attracting especially young persons with a short questionnaire, we suspected that facilitating smartphone response would be beneficial to young persons’ response rates as well. To that end, CBS includes QR codes in advance letters for surveys shorter than 10 minutes. Whether including a QR code actually draws in young people to do the survey on smartphone and what the general effect is on response rates, was not tested formally however. Previous research suggested that including a QR code could actually be detrimental for response (Smith 2017). We took the opportunity of the times experiment to also test the effect of including QR codes.

## Method

The experiment was performed in the survey of consumer sentiments (SCS). This survey is eight minutes long. It has a web – CATI sequential mixed mode design. An advance letter is sent with a login code and password for the web questionnaire. The SCS is incentivized with a lottery incentive of iPads. Two reminder letters are sent. Nonrespondents of the web phase are subsequently approached for a CATI interview, provided that a telephone number can be found. The SCS web questionnaire is suited for smartphone. The sample consists of persons of 16 year and older.

For the experiment, a parallel sample was drawn. There were four experimental Time conditions: ‘5 minutes’ (a few minutes shorter than the actual length), ’10 minutes’ (a few minutes longer than the actual length), ‘a few minutes’, or no time indication. The latter condition is CBS’s operational procedure for longer surveys. ‘A few minutes’ is the phrase that is used in the standard SCS advance letter. For the experiment, only the web part of the SCS was used. The web results of the regular SCS served as control group.

The QR code experiment was crossed with the Time conditions. Table 1 shows the design and sample size of the experiment.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 5 min | 10 min | No time | ‘few minutes’  | Total n |
| QR +  | 321 | 321 | 321 | 2100 | 3000 |
| QR -  | 321 | 321 | 321 | 321 | 1284 |
| Total n | 642 | 642 | 642 | 2421 | 4284 |

It was determined beforehand that the analytic power would be too low to look at the interactions between time and QR code, so the analysis will focus on Time and QR code separately.

Three measures are analyzed: the percentage of sample units starting the questionnaire, the percentage break off, and the response rate. In addition, the device that people used to login was studied in relation to Time and QR conditions. Although analytic power is too low to draw firm conclusions concerning the effect of age, I did look into that, and report preliminary findings.

## Results

Tables 2 to 4 show the percentages of sample persons starting the questionnaire, the percentage of people who started filling in the questionnaire, but broke off prematurely, and the percentage response[[2]](#footnote-2) by Time and QR condition.

***Table 2 % sample persons starting the web questionnaire by Time and QR condition***



Although there is a difference of about 4 percentage points between the condition with the highest percentage of starters (5 minutes) and the lowest (a few minutes and 10 minutes), this difference was not significant. The amount of break-off in the 5 minutes condition was also the highest, reminiscent of earlier findings that break-off is higher when the actual questionnaire duration is longer than expected. The differences in break-off rate by Time are not significant, however. Neither are the differences in response rates.

Including the QR code in the letter has no significant effect on either starting the questionnaire and subsequent response, but there is a marginally significant effect (p=.10) on the percentage of break-off. Break-off is higher when a QR code is offered.

***Table 3. % break- off of people starting the questionnaire by Time and QR condition***



***Table 4. % response by Time and QR condition***



I wondered whether the higher break-off in the QR condition had something to do with device use. It is known from previous research that break-off is higher when the questionnaire is filled in on smartphone than on other devices. In this experiment we see the same pattern, with 9.2% break-off when people use the smartphone, 4.5% for tablet users, and 3.8% for pc/laptop users (χ2(4) = 36.9, p < .001). We studied the effect of Time and QR code on device use. Time did not have influence on device choice, but the presence of a QR code did, see table 5.

***Table 5. % of people choosing smartphone, tablet of pc/laptop by QR code***



When a QR code is offered, people do indeed choose to respond via smartphone more often than in the no-QR condition (χ2(4) = 25.8, p < .001). These data suggest that the higher break-off in the QR condition may partially be an effect of pushing people to smartphone. The QR by device interaction on the percentage of break-off was however not significant, see table 6.

***Table 6. % of people breaking off by device and QR code***

 

A marginally significant (p<.10) interaction of device and Time indicated that not mentioning time leads to a higher break-off in smartphones than the ‘few minutes’ condition, while the opposite pattern was found for tablet users. The difference between either time and the other times were not significant for smartphones nor tablets. For pc/laptop users, the differences between Time conditions were not significant. Table 7 shows these results.

***Table 7. % of people breaking off by device and Time condition***



An important reason to facilitate smartphone questionnaires and to strive for short questionnaires, is the hope to increase response of young persons. We studied whether the inclusion of a QR code or the promise of a short questionnaire did increase response rates in this group. The results in table 8 show that this may indeed be the case. Although the number of 16 to 30 year old is quite limited (about 70 in each experimental condition), the 5 minute condition leads to marginally significantly more response (p = .10) in this age group. The differences in the other ages groups were not significant. There were no differences between age groups in the percentage of people logging in or breaking off.

***Table 8. % response by age and Time condition***



The inclusion of a QR code did not lead to a higher number of young people logging in or responding. The only effect of the QR code was that elderly people (65 and over) broke of significantly more if they used a QR code to log on (5.5% with QR code, 0,7% without, p <. 01).

## Discussion

The four Time conditions studied here did not have large effects on either login, break-off and response rates. The five minute condition looks promising however, especially in view of the high login rate of young persons. The results for the ‘no time mentioned’ condition are interesting too. It looks like the combination of providing a QR code, but not mentioning time gives respondents a clue and expectation of short duration.

Providing a QR code in the letter did not influence login or response, but did lead to a higher level of break-off. This may be the result of the fact that including the QR code led people to use their smartphone more. And smartphone respondents tend to brake-off to a much larger extent than respondents on other devices, a finding that was replicated in these data as well.

CBS invests in questionnaire development for smartphones. Embracing SQD is one of the measures taken in this regard. Ideally, however, these measures should stimulate respondents who would otherwise not have responded. Pushing respondents to smartphones instead of their tablets or pc/laptops could actually backfire.

Although the differences in login rates and response rates were not significant with these sample sizes, they would be relevant in a mixed mode setting if they would actually be this large (almost 8 percentage points difference between the highest and lowest login rate in these data). The data inspire to replication experiments. Those would create more mass to study interaction effects and subgroup differences. The answer to the question we set out to study, i.e., how long the splits should be in our SQD cannot yet be answered with the present data.

## References

Crawford, Couper & Lamias (2001) web surveys. Perception of burden. *Social science Computer Review, 19,* 146-162.

Galesic. M.. & Bosnjak. M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public opinion quarterly. 73*(2). 349-360.

Revilla, M. and Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research,* 59, 557-566.

Roberts, A. (2017). The effect of respondent characteristics on break-off in online surveys; A comparison between respondents using smartphones, tablets and pc. Paper presented at the 28th Workshop on Person and Household Nonresponse, Utrecht, the Netherlands.

Smith, P. (2017). An experimental examination of methods for increasing response rates in a push-to-web survey of sport participation. Paper presented at the 28th Workshop on Person and Household Nonresponse, Utrecht, the Netherlands.

Yan, T., Conrad, F., Tourangeau, R. and Couper, M.P. (2010). Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research,* *23*, 131-147.

1. *The views in this paper are those of the author, and do not necessarily reflect those of Statistics Netherlands.* [↑](#footnote-ref-1)
2. Response rate is percentage of responses from sample. [↑](#footnote-ref-2)