



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

Deliverable Number: 7.12

Deliverable Title: Improving web panel respondent behaviour: The effect of encouragement messages throughout the course of the survey

Work Package: 7 'A survey future online'

Deliverable type: Report

Dissemination status: Public

Submitted by: ESS ERIC (UPF)

Authors: Oriol J. Bosch (UPF), Wiebke Weber (UPF), Melanie Revilla (UPF)

Date Submitted: July 2018

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.





 @SERISS_EU

www.seriss.eu

SERISS (Synergies for Europe's Research Infrastructures in the Social Sciences) aims to exploit synergies, foster collaboration and develop shared standards between Europe's social science infrastructures in order to better equip these infrastructures to play a major role in addressing Europe's grand societal challenges and ensure that European policymaking is built on a solid base of the highest-quality socio-economic evidence.

The four year project (2015-19) is a collaboration between the three leading European Research Infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey of Health Ageing and Retirement in Europe (SHARE ERIC) and the Consortium of European Social Science Data Archives (CESSDA AS) – and organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and the WageIndicator Survey.

Work focuses on three key areas: Addressing key challenges for cross-national data collection, breaking down barriers between social science infrastructures and embracing the future of the social sciences.

Please cite this deliverable as: Bosch, O.J., Weber, W., and M. Revilla (2018) Improving web panel respondent behaviour: The effect of encouragement messages throughout the course of the survey. Deliverable 7.12 of the SERISS project funded under the *European Union's Horizon 2020 research and innovation programme* GA No: 654221. Available at: www.seriss.eu/resources/deliverables?

Summary

This report analyses an experiment conducted in waves 2, 4 and 6 of the CROss-National Online Survey (CRONOS) panel in the three participating countries, Estonia, Slovenia and Great Britain. The second wave was conducted between April and June 2017. The fourth and the sixth wave were open, respectively, from September to October 2017, and January to February 2018.

This experiment is intended to establish the impact of motivational messages in web surveys on data quality. Respondents were randomly assigned into a control group and two treatment groups in Wave 2 and were kept in the same group for the remaining two waves. Four to five other messages (depending on wave) were placed throughout the questionnaire. The control group respondents were only displayed a neutral introduction, while those in the treatment groups were also presented additional motivational text, emphasizing the positive or the negative consequences of their behaviours. In order to establish the impact on data quality, we compare control and treatment groups in terms of break-offs, item nonresponse, non-differentiation, completion time, efforts reported and evaluation of the survey experience.

Contents

1.	Introduction.....	5
2.	Motivational messages in web surveys.....	5
2.1	Background.....	5
2.2	Contribution.....	6
3.	Method and data.....	7
3.1	The CRONOS panel.....	7
3.2	Waves.....	8
3.3	Questionnaire.....	8
3.4	The experiment.....	9
3.5	Data quality indicators.....	9
3.6	Analyses.....	10
3.6.1	Break off.....	10
3.6.2	Item nonresponse.....	10
3.6.3	Non-differentiation.....	11
3.6.4	Completion time.....	12
3.6.5	Effort reported.....	12
3.6.6	Survey evaluation.....	13
4.	Findings.....	13
4.1	Break off.....	13
4.2	Item nonresponse.....	14
4.3	Non-differentiation.....	15
4.4	Completion time.....	16
4.5	Effort reported.....	17
4.6	Survey evaluation.....	18
5.	Discussion and Conclusions.....	18

1. Introduction

One of the biggest challenges for online panels is the balancing act between high data quality and representativeness. Non-participation, breakoffs, or losing participation over waves are a risk for the representativeness of the data, while missing answers or satisficing response behaviour endangers the data quality. Measurement accuracy in survey estimates relies on respondents paying attention to questions and tasks, and making an effort to provide accurate answers. However, respondent motivation during the survey response process can differ across individuals and situations, thus leading to suboptimal responding behaviours (Cannell, Miller, and Oksenberg 1981; J. A. Krosnick 1991), also known as ‘satisficing’ (Krosnick, 1991). On the advent of web surveys, it was soon recognised that certain aspects and features of web data collection may damage motivation and thus trigger satisficing behaviours. Survey methodologists have been searching for strategies to identify and reduce this kind of behaviour among online respondents, but studies tend to suffer from low effect sizes and research has been limited to a small set of western countries.

Motivating, praising or telling respondents off might affect respondent’s willingness to participate and increase the number of answers. However, if these answers are given without appropriate thinking, this can actually lead to a decrease of the data quality. While in interviewer-administrated surveys, respondents can receive this encouragement spontaneously from the interviewer, in web surveys this feedback has to be decided in advance and be introduced as a design feature of the survey.

The present deliverable describes an experiment conducted in waves 2, 4 and 6 of the CROSS- National Online Survey (CRONOS) panel. This assessment is intended to establish the impact of motivational messages in web surveys. Thus, we evaluated a number of indicators related to data quality, survey evaluation and completion time.

2. Motivational messages in web surveys

2.1 Background

Measurement error refers to the difference between a recorded response and the actual value of the variable of interest (Groves 1989). This difference can be rooted in various aspects of the measurement context, including the mode of administration, the questionnaire, and the respondent. As researchers, we have little to no control over the characteristics of those respondents who participate in the study, but we can try to use methodological approaches to maximise respondent engagement in the survey responding process. A number of theoretical models have been proposed to explain respondent-related measurement error (Tourangeau, Rips and Rasinski, 2000), leading

to different approaches to reduce it. In web surveys, the main strategies have focused on improving respondent's motivation and engagement by using interactive features, providing feedback when respondents answer too fast or presenting messages that highlight the importance of answering accurately.

Previous literature has studied the possibility of motivating respondents to participate in web surveys and answer the questions in the best way they can, by providing specific messages in the introduction of the survey (see e.g. Revilla, 2016) or within the course of the surveys. We focus on this second possibility in this deliverable.

Since panellists have different motivations to sign up for the panel in the first place (enjoyment, curiosity, belief that sharing their opinion can have an impact, interest in the reward, etc; see Brügger et al., 2011) for more details), ideally panellists could be encouraged according to their motivation. However, this can only be done if initially panellists answer a survey about their motivation to participate in the panel. This is a rather expensive endeavour which is not possible for all surveys.

Thus, Sakshaug and Crawford (2010) studied a different kind of respondent tailored motivations. They tested different messages of encouragement and their effect on deterring or delaying breakoffs. They conducted an experiment with college students where they tested three conditions: 1) no messages of encouragement (control), (2) three brief messages of encouragement with generic content, and (3) three brief messages of encouragement that contained content tailored to their specific school. They did not find any association between displaying the messages and deterring or delaying breakoffs. Kapelner and Chandler (2010) did not differentiate between their respondents and tried to motivate all of them by showing an alarming red text at the bottom of each question page: "Please answer accurately. Your responses will be used for research." They used Amazon's MTurk and did not find an effect on breakoffs either. More recently, Al Baghal and Lynn (2015) conducted an experiment in the Innovation Panel, part of the Great Britain Household Longitudinal Study which is based on a stratified, multi-stage sample of persons and households in England, Scotland and Wales. They found that prompting a respondent with a motivational message after a missed question that states "If possible please provide an answer to this question as this is one of the key questions in this study. Please be assured that the information you give us will be treated confidentially." increases item response. However, this cannot be done for any missed question as this would add to respondent's burden and dilute the effectiveness of the motivational statement.

2.2 Contribution

The literature review shows mixed results: some studies found an impact of the motivational messages, whereas others do not. In addition, the limited empirical evidence existing is mainly based on studies within a single country, in English speaking countries and on opt-in panels or convenience samples. Only the study of Al

Baghal and Lynn (2015) was based on a probability-based panel in three UK countries while the others were based on opt-in or student samples in a single country. However, this research has been carried out in a limited set of countries, and largely before smartphones became a popular device for web survey completion.

Our goal is to study the impact of motivational messages in web surveys, by extending the existing literature 1) using a probability-based web panel, which should provide results more easily generalizable to the general population; 2) collecting data in three different countries spread over Europe and using different languages, which will allow to test if there are differences across cultures and languages; 3) collecting data in several waves of the panel, which will allow see if there is an effect of the survey content; 4) considering more indicators to assess the effect of the motivational, not only break-off rates but also data quality, and 5) testing different motivational messages by following one of the recommendations for further research Sakshaug and Crawford (2010: 58) gave. As they did not find an effect of positive formulated motivational messages, they suggested emphasizing the negative consequence of a breakoff intermittently throughout the questionnaire. However, bearing in mind that stressing the importance of completing the survey might increase satisficing behaviour, we adjusted the message to emphasize the negative consequence of breakoff and satisficing. Moreover, we tested not only the negative but also the positive motivational message compared to no message at all.

To do this, we conducted an experiment in waves 2, 4 and 6 from the CRONOS survey in the three participating countries, Estonia, Slovenia and Great Britain.

3. Method and data

3.1 The CRONOS panel

The CRONOS panel was set up in Estonia, Great Britain and Slovenia with three goals: 1) evaluating the feasibility of conducting web surveys on probability samples of the general population in a cross-national context, 2) establishing the foundations for building efficient cross-national infrastructures for web survey data collection, and 3) developing a blueprint for comparative web surveys using probability samples¹.

Respondents were invited to become panel members after they participated in the European Social Survey (ESS) Round 8 face-to-face interview, which took place between September 2016 and February 2017. After the ‘welcome survey’ was launched, data collection was conducted bi-monthly for 12 months, starting in February 2017. Panel members received an unconditional incentive with their invitation to each wave of

¹ For more information please visit:

http://www.europeansocialsurvey.org/methodology/methodological_research/modes_of_data_collection/cronos.html

the survey. To allow for their participation, the project provided internet-enabled tablets to 182 panel members who did not have internet access for personal use. Recruitment rates, calculated as the proportion of individuals in the gross ESS sample who initially agreed to join CRONOS (including hesitant respondents), ranged from around 30% in Great Britain to around 40% in Estonia and Slovenia.

3.2 Waves

This experiment was conducted in CRONOS waves 2, 4 and 6. The second wave was launched the 27th of April 2017 and was open until the 10th of June 2017. The fourth and the sixth wave were open, respectively, from the 7th of September 2017 to the 12th of October 2017, and the 9th of January 2018 to the 18th of February 2018. CRONOS response rates² respectively for waves 2, 4 and 6 were, in Estonia: 23% (664), 20% (581) and 21% (600); in Great Britain: 16% (692), 14% (610) and 14% (641); and in Slovenia: 21% (482), 25% (561) and 25% (571). Furthermore, the participation rates³, were respectively for waves 2, 4 and 6, in Estonia: 82%, 74% and 77%; in Great Britain: 60%, 53% and 56%; and in Slovenia: 63%, 78% and 80% (Villar et al., 2018). A total of 1,763 (wave 2), 1,702 (wave 4) and 1,702 (wave 6) panellists completed the survey. Those were distributed, respectively for control, treatment 1 (negative, see section 3.4) and treatment 2 (positive, see section 3.4), in wave 2: 584, 591 and 588; in wave 4: 565, 553 and 584; and in wave 6: 583, 548 and 571. In terms of electronic devices, a majority of respondents participated in the survey using a computer, accounting for respectively 59.5%, 59.2% and 59.5% of the total respondents for waves 2, 4 and 6.

3.3 Questionnaire

Wave 2 comprised 99 questions about the stages of life, family and the Internet use, among other issues. Wave 4 was shorter than the other questionnaires, containing 64 questions about, for instance, cultural dimensions, the nature and the environment, and personality traits. Finally, Wave 6 comprised 84 questions about income fairness, social and institutional trust and political efficacy, for example.

In line with the main objectives of the CRONOS panel, data quality assessment was an important component of each data collection wave. Questions at the end of each wave asked respondents to evaluate the survey completion process. Paradata describing the response process (such as response latencies) were collected. Questionnaires also

² Response rate is calculated as the sum of complete and partial interviews over total number of issued sample units eligible for CRONOS excluding those identified as ineligible by interviewers (e.g. , ineligible addresses in Great Britain , sample units no longer living in the country), and those who did not meet the CRONOS eligibility criteria (sample members aged 15-17 and/or living in Northern Ireland).

³ Number of participants (partial+complete) as a proportion of sample units invited to participate in CRONOS

included methodological experiments to investigate the performance of various approaches to question design and other questionnaire-related factors that may influence data quality.

For each wave, the source questionnaire was translated from English into Estonian, Russian and Slovenian. When designing the CRONOS panel, the aim was to minimise questionnaire-related measurement error by paying close attention to question wording and to the visual design of questions across devices, following web questionnaire design guidelines as far as possible.

3.4 The experiment

Respondents were randomly assigned into a control group and two treatment groups in Wave 2 and were kept in the same group for the remaining two waves. The first motivational texts were displayed to respondents immediately after the introductory (welcome) page. Four to five other messages (depending on wave) were placed throughout the questionnaire. Some messages were presented as additions to standard introductions to questionnaire modules. The control group respondents were only shown a neutral introduction, while those in the treatment groups were also presented additional motivational text. Other messages were displayed as standalone texts on separate pages without a specific introduction to the topic; such pages were skipped for control group respondents.

The first motivational message was displayed immediately after the welcome page of the questionnaire and was the same in all waves. The contents of other messages and their placement varied between the waves. This allowed tailoring the content to a specific topic and selecting the most appropriate locations to display the messages. All motivational texts are listed in the Appendix A.

3.5 Data quality indicators

Various indicators can be used to operationalise data quality, but none can cover the concept adequately if used in isolation. For variables that can be externally observed, such as behaviours (e.g., number of visits to the hospital) and certain facts (e.g., marital status), record data may be available that can be compared to the observed survey value as an estimate of measurement error. For attitudinal questions, however, such external checks are not possible. Instead, a combination of indicators is typically used including item-nonresponse, response latency, non-differentiation, and length and quality of answers to open-ended questions (Berzelak, Weber and Revilla, 2018), .

Data quality indicators of attitudinal variables do have limitations, which relate to their level of precision and accuracy (their own measurement error), the ease of implementation and the ease of interpretation. For example, non-differentiation (see definition below) requires the inclusion of questions formulated in ‘opposite’ directions, so that giving the exact same answer to all questions would be incompatible. Response

latencies are sometimes difficult to analyse, due to missing timing information related to recording problems as well as extreme values related to respondent multitasking. Interpretation is also challenging: short response times can indicate that a question is easy or the topic very salient for the respondent, but it may in turn indicate that respondents are not engaging sufficiently in the response process, perhaps not reading the full question. Revilla and Ochoa (2015) found a significant positive correlation between response latency and other indicators of data quality, but the correlation was not large at .30 in absolute value. Moreover, some indicators can only be computed for certain types of questions, which may not be available in a specific questionnaire. Due to these limitations, in the present study we evaluated as many different data quality indicators as possible to assess the effect of the motivational messages.

3.6 Analyses

Analyses were done using Stata 14 (StataCorp., 2015). In order to see if there was an effect of the positive and negative motivational messages on the quality of the answers, we compared each treatment group to the control group, in terms of the different data quality indicators described in the subsections below. All analyses were done separately for each country. Comparisons were done for each block of questions that come after one of the messages, and before the next one (except for survey evaluation, see 3.6.6). T-tests were used when comparing means, and Z-tests ('prtest' in Stata) when comparing proportions. Differences were considered to be statistically significant if the p-values associated with the tests were lower than .05.

3.6.1 Break off

Break off occurs when respondents stop completing the survey before the end of the questionnaire due to motivational, technical or other reasons. A high incidence of survey break-offs can indicate excessive burden of survey participation, high sensitivity of questions, lack of motivation or technical problems, among other things (see e.g. Peytchev, 2009).

In this study, we computed the proportion of respondents leaving the survey in each set of questions of the different waves, per country and experimental group. Then, we tested if the differences of those proportions across groups were significant.

3.6.2 Item nonresponse

Item nonresponse occurs when respondents do not answer one (or more) question(s). This may occur for various reasons such as inability or unwillingness to provide an answer, failure to adequately comprehend the question or to form an appropriate response, lack of knowledge, or lack of motivation to devote sufficient effort to the response process. In such cases, respondents may not answer the question at all or select a specific non- substantive answer, such as 'Don't know' or 'Prefer not to answer' (Berzelak, Weber and Revilla, 2018).

We consider item non-response non-answers but also all “don’t know” answers and refusals. Proportions were calculated for those who completed the survey. We first computed the proportions of missing values for each question in a specific block, and then took the average of these proportions over all questions in a block. That average was compared across groups (using t-tests, being aware that the N was small).

3.6.3 Non-differentiation

Non-differentiation occurs if a respondent, when offered a similar response scale across a set of questions, fails to differentiate between the questions, choosing for each (almost) the same answers (Jon A. Krosnick, 1991), which leads to a low variance of his/her answers. In its extreme form, the respondent selects the exact same answer repeatedly: this pattern is also referred to as ‘straight-lining’. These two indicators (low variance and pure straight-lining) have often been used as indicators of satisficing (Jon A. Krosnick 1991).

There are difficulties in operationalising non-differentiation in a way that is meaningful as an indicator of data quality. For low response variance across questions to really indicate low data quality, it is important to include items that one would expect to be negatively correlated. Indeed, if all questions are strongly related in topic and are expressed in the ‘same direction’ (where a given scale point represents one direction of the attitude for all questions) observing low variance in answers may not mean that the respondent is not putting effort into the survey process, and may instead be a sign of coherence across answers and good data quality.

To deal with this issue, in this study, we first identified sets of questions using the same scale at least five times in a row and with the items not all in the same direction or not measuring the same. We could identify one for wave 2, four for wave 4 and two for wave 6 (see Appendix B for the question wording).

For wave 2:

- Set 1: 13 questions about respondent’s opinion about parenting (w2q55-w2q78).

For wave 4:

- Set 1: 10 questions about how the respondent characterizes him/herself regarding own opinions (w4q5-w4q14)
- Set 2: 20 questions about how the respondent characterises him/herself regarding problem-solving capacities and interpersonal relationships (w4q15-w4q35).
- Set 3: 12 questions about respondent’s opinion about the environment (w4q36-w4q47).
- Set 4: 6 questions about how the respondent characterizes him/herself regarding own opinions and problem-solving capacities (w4q54-w4q59).

For wave 6:

- Set 1: 7 questions about respondent's opinion regarding the influence of factors in finding a job (w6q22-w6q28).
- Set 2: 9 questions about respondent's opinion about the internet (w6q48-w6q56).

Then, to measure non-differentiation, first, for each of these sets, we computed the variance of answers of each respondent, focusing on respondents who provided substantive responses, i.e. excluding 'don't know' and 'prefer not to answer'. Then, we computed the mean for each set. We compared the significance of the differences of the means across groups.

Second, we calculated the proportion of pure straight-liners (respondents with a variance of 0: the respondent selects always exactly the same answer) for each set and compared across groups.

3.6.4 Completion time

In this study, we considered completion times to be indicators of data quality. Shorter completion times can indicate that respondents are not making enough effort to carefully answer the questions, whereas they can also indicate that the survey questions is easy to process or that respondents have strong opinions in their minds (Revilla and Couper, 2017). Completion times are measured using paradata. However, there are different ways to measure them (see Couper and Peterson, 2016; Yhan and Tourangeau, 2008). Although it is possible to account for long completion times due to lost of focus due to online multitasking, i.e. using other applications or visiting other web pages using the same device (Höhne and Schlosser, 2017), the CRONOS panel did not collect such information. Hence, completions times in the CRONOS panel were measured as the total time that the respondent spent between clicking to the "next" button page once, and clicking again. Thus, we need to deal with outliers, who would be multitasking (either on the same device, or outside of the device, e.g. watching TV or speaking with friends or relatives). We handled the presence of outliers by following the truncation procedure used by Yan and Tourangeau (2008), where we replaced observations beyond the lower and upper one percentile with the lower and upper one percentile values respectively.

After applying this correction for outliers, we compared the mean completion time for a given block across groups.

3.6.5 Effort reported

Previous indicators were based on external measures of data quality. Internal auto-evaluation of respondent about the quality of their own answers can also be collected using a measure of self-reported effort. Revilla (2016) found that motivational messages and commitment statements affect the self-reported effort of a specific kind of

respondents, for instance, those with reasonable levels of effort make a higher effort. Furthermore, Revilla and Ochoa (2015) found that more reported efforts go together with longer completion times, even if they do not find a significant relationship with other quality indicators.

In the three waves of interest for this study, a question measuring the self-reported effort was included almost at the end of the survey: "How much did you work at providing the most accurate answers you can to the questions in this survey?". The question used a unipolar 5-point fully labelled scale from "Not at all" to "A great deal." The full scale is provided in Appendix C. Since that question was only asked at the end of each wave, we compared the mean across groups by wave. Differences should be due to the cumulative effect of the different motivational messages. Results were computed focusing on respondents who provided substantive responses.

3.6.6 Survey evaluation

Respondents' perception of their experience answering the survey is also an indicator of data quality often studied in the literature. This is particularly important in a panel, because low satisfaction with the process may affect the likelihood of participation in subsequent surveys. In waves 2, 4 and 6, two questions were included almost at the end of the survey to evaluate: a) how difficult the respondents found it to understand and answer the questions in the survey, b) how much they enjoyed answering the survey. Both questions used a unipolar 5-point fully labelled scale from "Not at all" to "A great deal" or "Extremely".

The analyses for both questions were done in a similar way as the ones for the effort reported.

4. Findings

4.1 Break off

The first indicator used to operationalize data quality was the proportion of break-offs. Table 1 presents the proportions of respondents leaving the survey in each set of questions for the different waves, per country and group.

Overall, the proportion of break-offs was small, similar to the welcome CRONOS survey (Berzelak, Weber and Revilla, 2018). Comparing across groups, there were no significant differences in the proportion of break-offs. Hence the motivational messages did not reduce the proportion of break-offs, even when emphasizing the importance of continuing filling in the questionnaire. However, this might be linked to the very low overall break-offs (already in the control group).

Table 1. Proportion of break-offs per set of questions, country and group (in %)

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave 2	Set 1	1.4	1.3	0.0	1.3	2.1	0.8	1.2	0.0	0.6
	Set 2	0.0	0.4	0.0	0.9	0.4	2.1	0.0	0.0	1.3
	Set 3	1.4	1.3	0.0	1.3	0.0	0.0	0.0	0.6	0.6
	Set 4	1.4	2.6	2.7	0.4	0.4	0.9	0.0	0.0	1.8
	Set 5	0.0	0.4	0.0	0.0	0.0	0.4	0.0	0.0	0.0
Wave 4	Set 1	0.5	2.0	1.5	1.0	2.1	1.8	1.5	0.6	2.2
	Set 2	1.0	1.5	0.5	0.0	0.5	0.0	0.0	0.0	0.6
	Set 3	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
	Set 4	0.0	1.0	0.0	0.5	1.6	1.3	0.5	0.0	0.0
	Set 5	0.0	0.5	0.0	0.0	0.0	0.4	0.0	0.0	0.0
Wave 6	Set 1	1.5	2.9	3.5	2.9	6.9	3.9	4.3	4.2	2.9
	Set 2	0.5	1.5	0.0	0.5	1.5	0.9	0.0	0.5	0.0
	Set 3	0.5	0.5	1.0	0.0	1.0	0.4	0.0	0.0	0.0
	Set 4	0.0	1.0	0.0	1.0	0.0	0.4	0.0	0.0	0.0
	Set 5	1.0	1.0	0.5	1.4	2.5	0.9	1.0	1.1	0.6
	Set 6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: No significant differences

4.2 Item nonresponse

Next, we explore the proportions of missing values for each of the questions in a specific block. Table 2 presents the average item non-response for each set, per country and wave.

There are 18 significant differences out of 96 comparisons across groups.

First, looking at Estonia, for wave 2 and set 3, the treatment group exposed to messages about the negative consequences had a significantly higher proportion of item non-response than the control group. However, for the first set of wave 4 the control group had a higher proportion of item non-response than both treatment groups. Hence, results for Estonia are not clear.

Similarly, Slovenia presents mixed results. The control group had a slightly lower proportion of missing item for set 3 in waves 4 and 6, whereas for set 2 in wave 4 the proportion was significantly higher than for the group exposed to messages about the negative consequences.

Conversely, results for Great Britain show a clearer pattern. The control group had a significantly higher proportion of item non-response than both negative and positive treatment groups for sets 3 and 4 in wave 2 and sets 1, 2 and 3 in wave 4. Additionally, in wave 6 the control group had a higher proportion than the group exposed to messages about the positive consequences for set 2 and than the group exposed to messages about the negative consequences for set 3.

Table 2. Average proportion of item non-response per set of questions, country and group (in %)

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave 2	Set 1	3.3	2.8	3.0	2.9	2.6	2.6	2.6	2.0	1.7
	Set 2	0.4	0.2	0.5	0.4	0.1	0.2	0.3	0.3	0.1
	Set 3	0.1	0.3*	0.0	0.4	0.0*	0.2*	0.1	0.1	0.1
	Set 4	0.3	0.5	0.2	0.8	0.1*	0.1*	0.3	0.3	0.2
	Set 5	0.1	0.0	0.5	0.6	0.2	0.2	0.2	0.0	0.0
Wave 4	Set 1	0.35	0.0*	0.0*	0.4	0.2*	0.0*	0.1	0.0	0.2
	Set 2	0.1	0.1	0.1	0.6	0.0*	0.1*	0.34	0.0*	0.5
	Set 3	0.0	0.0	0.2	0.5	0.1*	0.0*	0.0	0.1	0.2*
	Set 4	1.5	1.5	0.7	2.5	2.4	1.2	1.1	1.4	1.4
	Set 5	0.2	0.0	0.4	0.7	0.0	0.0	0.0	0.0	0.0
Wave 6	Set 1	0.5	0.9	0.5	1.4	0.8	1.1	0.7	0.9	0.9
	Set 2	0.1	0.4	0.3	0.2	0.1	0.0*	0.1	0.2	0.0
	Set 3	1.1	0.9	0.9	1.1	0.3*	0.5	0.5	1.0*	1.0
	Set 4	0.6	0.8	0.4	1.8	0.8	1.4	0.7	0.6	0.7
	Set 5	2.3	2.1	1.8	4.4	2.6	3.4	2.7	2.0	3.0
	Set 6	0.4	0.7	0.7	0.8	0.4	1.1	0.7	0.6	0.4

Note: * in column “Negative” indicates significant differences between control and negative (5% level);

* in column “Positive” indicates significant differences between control and positive (5% level)

To sum up, for Estonia and Slovenia the impact of exposing respondents to motivational messages is low and without a clear pattern. However, for Great Britain nearly half of the sets present significant positive impacts, meaning that exposing respondents to motivational messages reduces the proportion of item nonresponse.

4.3 Non-differentiation

Moving to non-differentiation, Table 3 presents the mean variance of each set for all waves per country and group.

Table 3. Variance of answers per set of questions, country and group

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave 2	Set 1	1.0	1.0	1.0	1.2	1.2	1.2	1.0	1.1	1.0
Wave 4	Set 1	3.8	3.7	3.9	4.6	3.8*	4.4	3.9	3.9	3.6
	Set 2	5.1	5.1	5.3	5.8	5.3	5.7	5.8	5.6	5.5
	Set 3	1.3	1.5*	1.3	1.6	1.6	1.6	1.6	1.5	1.6
	Set 4	0.7	0.6	0.6	0.6	0.6	0.7	0.6	0.6	0.5
Wave 6	Set 1	0.7	0.6	0.6	0.8	0.8	0.8	0.7	0.7	0.8
	Set 2	1.3	1.3	1.2	1.2	1.2	1.3	1.5	1.4	1.4

Note: * in column “Negative” indicates significant differences between control and negative (5% level);

* in column “Positive” indicates significant differences between control and positive (5% level)

Overall, differences across groups are small and do not present a clear pattern. Indeed, only two differences were statistically significant out of 42 tests, both in wave 4, but they do not go in the same direction:

- for set 3, the group exposed to messages about the negative consequences in Estonia had a higher variance than the control group
- for set 1, the group exposed to messages about the negative consequences in Great Britain had a lower variance than the control group.

Table 4 presents the proportions of pure straight-liners for each set of questions, per wave and country.

Table 4. Proportion of straight-liners per set of questions, country and group (in %)

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave										
2	Set 1	0.5	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.0
Wave	Set 1	0.6	0.0	1.6	1.0	0.0	0.5	3.1	2.2	1.7
4	Set 2	0.0	0.0	1.1	0.5	0.0	0.5	0.5	0.0	0.0
	Set 3	1.7	0.0	0.5	1.0	1.6	0.0	0.5	2.2	0.0
	Set 4	5.1	4.6	6.8	5.8	6.7	5.5	5.7	6.7	10.3
Wave	Set 1	1.0	1.6	1.6	2.5	2.2	0*	2.0	1.7	1.8
6	Set 2	0.0	0.0	0.0	1.5	0.0	0.5	0.0	1.7	1.8

Note: * in column “Negative” indicates significant differences between control and negative (5% level);

* in column “Positive” indicates significant differences between control and positive (5% level)

Overall, only a small proportion of respondents did not differentiate at all between their answers on these sets. Nevertheless, set 4 in wave 4, a repetition of some previously asked questions, presented a higher proportion of pure straight-liners, ranging from 4.6% to 10.3%. Comparing across groups, we only found a significant difference for set 1 in wave 6. The proportion of pure straight-liners for the group exposed to messages about the positive consequences was lower than for the control group.

4.4 Completion time

Table 5 presents the average completion time for each set of questions in each wave, across groups and per country.

There were 7 significant differences across groups out of 96 tests. First, in Estonia, the group exposed to messages about the negative consequences had a significantly higher completion time than the control group for three sets of questions (set 1 in wave 2 and sets 3 and 4 in wave 6). Thus, respondents exposed to messages about the negative consequences dedicated more time to answer the questions in those sets, which may be a sign of a higher effort to answer. Conversely, in Great Britain, the group exposed to messages about the negative consequences had a significantly lower completion time than the control group for sets 1 and 2 in wave 6. Finally, in Slovenia the group exposed to messages about the positive consequences had a significantly lower completion time

than the control group, in set 2 in wave 2 and set 1 in wave 6. Therefore, there is no clear patten across countries.

Table 5. Average completion time per set of questions, country and group (in seconds)

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave 2	Set 1	19.5	21.8*	20.9	17.9	17.6	18.2	18.2	19.2	17.2
	Set 2	19.0	20.8	19.5	17.5	17.4	18.0	18.8	17.4	16.8*
	Set 3	15.3	16.8	15.3	14.7	14.2	14.8	15.3	14.9	13.9
	Set 4	13.4	14.5	13.9	12.2	12.2	12.7	12.7	12.7	11.7
	Set 5	12.0	12.7	12.6	10.0	10.4	11.1	11.6	11.5	10.8
Wave 4	Set 1	16.4	18.4	17.4	14.7	15.0	15.2	14.7	14.1	13.8
	Set 2	14.2	15.2	15.1	13.4	13.5	13.7	12.7	12.3	11.9
	Set 3	12.3	13.2	12.8	11.6	11.3	11.8	11.9	11.8	11.3
	Set 4	15.0	16.5	15.9	14.5	14.4	14.8	13.5	13.1	12.8
	Set 5	9.1	9.7	9.4	9.8	9.6	10.0	9.1	9.7	9.2
Wave 6	Set 1	18.8	20.9	20.0	20.1	16.9*	18.7	19.5	18.4	16.1*
	Set 2	15.5	17.3	17.2	15.8	14.0*	15.3	14.4	14.1	13.6
	Set 3	17.0	19.4*	19.0	16.9	15.8	16.4	15.7	16.0	15.2
	Set 4	34.4	36.8*	36.0	35.1	34.5	34.5	32.6	33.3	33.1
	Set 5	19.7	21.4	21.7	19.1	18.2	18.9	17.4	18.1	18.2
	Set 6	11.7	12.6	12.2	10.3	10.3	10.3	11.1	10.1	10.1

Note: * in column “Negative” indicates significant differences between control and negative (5% level);
* in column “Positive” indicates significant differences between control and positive (5% level)

4.5 Effort reported

Table 6 presents the self-reported effort of respondents, ranging from 1 (“not at all”) to 5 (“A great deal”).

Table 6. Mean effort reported per wave, country and group

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
—	Wave 2	2.4	2.8*	2.5	3.5	3.5	3.6	3.7	3.8	3.9
—	Wave 4	2.4	2.8*	2.6*	3.4	3.5	3.5	3.9	4.0	3.9
—	Wave 6	2.7	3.1	3.0	3.6	3.7	3.7	3.9	4.0	3.9

Note: * in column “Negative” indicates significant differences between control and negative (5% level);
* in column “Positive” indicates significant differences between control and positive (5% level)

In Estonia, respondents from the control group self-reported significantly less effort than those exposed to messages about the negative consequences, for wave 2, and than both treatment groups for wave 4. In the other countries, there were no significant differences.

4.6 Survey evaluation

Table 7 presents the means for each of the survey evaluation questions. Values range from 1 to 5, 1 being “not at all” and 5 being “Extremely difficult” or “A great deal”. See Appendix A for more information.

Table 7. Mean survey evaluation per wave, country and group

		Estonia			Great Britain			Slovenia		
		Control	Negative	Positive	Control	Negative	Positive	Control	Negative	Positive
Wave 2	Difficult	1.5	1.6	1.5	1.3	1.4	1.3	1.6	1.6	1.6
	Enjoy	2.9	2.9	3.0	3.1	3.3	3.3*	3.8	3.9	4.0
Wave 4	Difficult	1.7	1.8	1.8	1.6	1.7	1.7	1.9	1.8	1.8
	Enjoy	2.8	2.9	2.9	3.2	3.2	3.3	3.8	3.9	3.8
Wave 6	Difficult	1.6	1.7	1.7	1.6	1.6	1.5	1.8	1.7	1.7
	Enjoy	2.8	2.9	2.9	3.1	3.2	3.2	3.7	3.7	3.8

Note: * in column “Negative” indicates significant differences between control and negative (5% level);

* in column “Positive” indicates significant differences between control and positive (5% level)

Overall, differences across groups are small and mainly non-significant. We find only one significant difference, in Great Britain for wave 2: respondents from the group exposed to messages about the positive consequences self-reported a significantly higher enjoyment filling in the questionnaire than those from the control group.

5. Discussion and Conclusions

The main goal of this study was to explore the impact on data quality of introducing motivational messages in web surveys. Therefore, we conducted an experiment in waves 2, 4 and 6 of the CRONOS panel (probability-based) in three different countries (Estonia, Great Britain, Slovenia).

Overall, no clear pattern has been found regarding the effect of motivational messages on data quality: results depend on the data quality indicator considered and the country. Indeed, we found for all three countries that motivational messages, positive or negative, did not have any impact on the proportion of break-offs. Nevertheless, the proportion of break-offs in the control group are already low. For item non-response, results are mixed. On the one hand, in Great Britain, in almost half of the sets, the control group had a higher average proportion of item non-response than one or both treatment groups. Both positive and negative messages appear equivalent. On the other hand, results from Estonia and Slovenia are small and inconsistent, with some sets presenting positive impacts of the motivational messages while others show negative effects. Furthermore, few significant differences across groups were found when studying non-differentiation and straight-lining. Moreover, the impact on completion

time of exposing respondents to motivational messages is low and inconsistent across countries.

In terms of self-reported efforts, respondents from Estonia in the control group reported to have invested significantly less work in providing most accurate answers than those from the group exposed to messages about the negative consequences in wave 2, and from both treatment groups in wave 4. In addition, asking respondents about their survey experience, we find overall few significant differences across groups: only in Great Britain, respondents exposed to messages about the positive consequences enjoyed the survey significantly more than those from the control group.

These findings regarding data quality are similar to what other authors have found in the past. Sakshaug and Carwford (2010) and Kapelner and Chandler (2010) found no effect on break-offs. Besides, Al Baghal and Lynn (2015) found that prompting a respondent with a motivational message after a missed question increases item response. Results from our study in Great Britain show also an increase of item response. In addition, respondents from both treatment groups in Estonia considered that they spent more work on providing the answers than those from the control group for waves 2 and 4. However, this higher effort cannot be appreciated in the completion time results.

Although we have found few significant impacts of exposing respondents to motivational messages, differences across country have appeared. Estonia and Great Britain have opposite results for item non-response, variance and completion time. Besides, Slovenia presents nearly no significant differences (7 of 140 comparisons).

The impact of motivational messages also varied depending on their position in the questionnaire. In addition to the first motivational texts displayed after the introductory (welcome) page, four to five other messages were placed throughout the questionnaire. Thus, the cumulative effect could decrease across the questionnaire, i.e. respondents being tired of repeatedly being exposed to motivational messages, or increase, i.e. respondents being more aware after being exposed to more messages. Results demonstrate that, of 26 significant differences (excluding questions about self-reported effort and survey evaluation), 16 were found in the first half (sets 1 and 2 for waves 2 and 4, and from 1 to 3 for wave 6), six in the middle (set 3 for waves 1 and 2), and three in the last half (sets 4 and 5 for wave 2 and 4, and from 4 to 6 for wave 6). Hence, significant impacts, positive or negative, on data quality were mainly found on the first sets early in the questionnaire.

Finally, we have analysed two different types of motivational messages, emphasizing the positive or negative consequences. Overall, their impacts are similar in direction (positive or negative effects on data quality compared to the control group) and in frequency (how many significant differences with the control group) with both types of motivational message having relatively little effect. However, in Estonia exposing respondents to positive messages had almost no impact, whereas in Great Britain the

effects, where they occur, are always positive and in Slovenia they are negative.

Overall, as in previous studies (Sakshaug and Carwford, 2010; Kapelner and Chandler, 2010; Al Baghal and Lynn, 2015), we found that motivational messages have little impact on data quality, suggesting that using standard motivational messages to all respondents is not the correct approach. However, more research is needed to investigate if adaptative approaches (i.e. a message only showed to those individuals answering too fast), and/or targeted motivational messages (i.e. messages adapted and only showed to specific cohorts associated with low data quality) could improve the results. In addition, since there are differences across countries, strategies to motivate respondents should be country-specific. Further research in different countries is therefore necessary.

Nevertheless, these results have limits. First, as Berzelak, Weber and Revilla (2018) demonstrated, data quality for the CRONOS panel is high, as expected. Respondents not only decided voluntarily to participate in the CRONOS panel but already previously in the European Social Survey face-to-face interview. Hence, the impact of motivational messages on already motivated respondents may not be the same than for the general population. In addition, the motivational messages used for this experiment were rather specific. Indeed, simpler and more generic messages could produce different results.

References

- Al Baghal, T., & Lynn, P. (2015). Using motivational statements in web-instrument design to reduce item-missing rates in a mixed-mode context. *Public Opinion Quarterly*, 79(2), 568–579. <http://doi.org/10.1093/poq/nfv023>
- Berzelak, N., Weber, W., and M. Revilla (2018) Analysis of panellist responding behaviour in the CRONOS welcome survey. Deliverable 7.11 of the SERISS project funded under the *European Union's Horizon 2020 research and innovation programme* GA No: 654221. Available at: www.seriss.eu/resources/deliverables
- Bruggen, E., Wetzels, M., de Ruyter, K., & Schillewaert, N. (2011). Individual differences in motivation to participate in online panels: the effect on response rate and response quality perceptions. *International Journal of Market Research*, 53(3), 369. <http://doi.org/10.2501/IJMR-53-3-369-390>
- Cannell, Charles F., Peter V. Miller, and Lois Oksenberg. 1981. "Research on Interviewing Techniques." *Sociological Methodology* 12: 389. <http://www.jstor.org/stable/270748?origin=crossref> (September 8, 2017).
- Fricke, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of web and telephone surveys. *Public Opinion Quarterly*, 69(3),

370-392.

- Groves, Robert M. 1989. 48 Wiley series in probability and mathematical statistics Probability and mathematical statistics Survey Errors and Survey Costs. Wiley-Interscience. 2009. Survey Methodology. Wiley; 2 edition.
- Kapelner, A., & Chandler, D. (2010). Preventing satisficing in online surveys: A “kapcha” to ensure higher quality data. *Information Systems Journal*, (June), 1–10. Retrieved from <http://crowdconf.com/images/finalpapers/kapelner.pdf>
- Krosnick, J. A. 1991. “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Survey.” *Applied Cognitive Psychology* 5(3): 213–36.
- Krosnick, Jon A. 1991. “Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Survey.” *Applied Cognitive Psychology* 5(3): 213–36. <http://doi.wiley.com/10.1002/acp.2350050305> (September 8, 2017).
- Peytchev, A. (2009). Survey breakoff. *Public Opinion Quarterly*, 73(1), 74-97.
- Revilla, M. (2016). Impact of raising awareness of respondents on the measurement quality in a web survey. *Quality & Quantity*, 50(4), 1469-1486.
- Revilla, Melanie, and Carlos Ochoa. 2015. “What Are the Links in a Web Survey Among Response Time, Quality, and Auto-Evaluation of the Efforts Done?” *Social Science Computer Review* 33(1): 97–114. <http://journals.sagepub.com/doi/10.1177/0894439314531214> (September 8, 2017).
- Revilla, Melanie, Willem Saris, Germán Loewe, and Carlos Ochoa. 2015. “Can a Non-Probabilistic Online Panel Achieve Question Quality Similar to that of the European Social Survey?” *International Journal of Market Research* 57(3): 395–412. <http://journals.sagepub.com/doi/10.2501/IJMR-2015-034> (April 26, 2018).
- Sakshaug, J., & Crawford, S. (2010). The impact of textual messages of encouragement on web survey breakoffs: An experiment. *International Journal of Internet*. Retrieved from https://www.researchgate.net/profile/Scott_Crawford4/publication/266502915_The_Impact_of_Textual_Messages_of_Encouragement_on_Web_Survey_Breakoffs_An_Experiment/links/55a415bf08ae00cf99c8f58e.pdf
- StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.
- Villar, A. and Sommer, E. (2017) *Web recruitment design plans and experimental testing*. Deliverable 7.3 of the SERISS project funded under the *European Union’s Horizon 2020 research and innovation programme* GA No: 654221. Available at: www.seriss.eu/resources/deliverables.
- Villar, Ana, Sommer, Elena, Finnøy, Didrik, Gaia, Alessandra, Berzelak, Nejc, and Bottoni, Gianmaria. (2018) CROSS-National Online Survey (CRONOS) panel data and documentation user guide. London: ESS ERIC.

Yan, Ting, and Roger Tourangeau. 2008. "Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times." *Applied Cognitive Psychology* 22(1): 51–68. <http://doi.wiley.com/10.1002/acp.1331> (September 8, 2017).

Appendix A. Motivational messages⁴

Message # and location	Experimental group	Message text
WAVE 2		
2.1 After intro	Control	N/A
	Positive emphasis	Answering questions carefully can greatly improve the accuracy and usefulness of the findings from this study. The more accurate your answers, the better the data will describe the opinions of people living in [COUNTRY].
	Negative emphasis	Answering questions carelessly can greatly reduce the accuracy and usefulness of the findings from this study. The less accurate your answers, the worse the data will describe the opinions of people living in [COUNTRY]
2.2 Before Q16	Control	N/A
	Positive emphasis	You have completed the first section of this survey. We really appreciate your continued co-operation. When you answer all questions, you help the results of the survey be more accurate and more useful.
	Negative emphasis	You have completed the first section of this survey. We really appreciate your continued co-operation. If you stopped answering questions now, the results of the survey could be less accurate and less useful.
2.3 Before Q33	Control	This section of the survey is about your views on marriage, family and parenting.
	Positive emphasis	This section of the survey is about your views on marriage, family and parenting. It is very important for this study to include the opinions of those who are married and those who are not, those who have children and those who do not. Please continue completing the survey to help make sure that both types of people are represented in the study.
	Negative emphasis	This section of the survey is about your views on marriage, family and parenting. It would be harmful for the study if we did not include the opinions of both married and unmarried people, of those who have children and those who do not. Please continue completing the survey to avoid having an unbalanced representation of these kinds of people.
2.4 Before Q55	Control	N/A
	Positive emphasis	We are moving now onto a new section and topic. We appreciate your continued effort, because studies have shown that answering questions carefully can greatly improve the

⁴ To see screenshots of the actual surveys, please visit:
https://www.europeansocialsurvey.org/docs/cronos/CRONOS_source_questionnaire_wave0-6.zip

		accuracy and usefulness of our research findings.
	Negative emphasis	We are moving now onto a new section and topic. We appreciate your continued effort, because studies have shown that answering questions carelessly can greatly reduce the accuracy and usefulness of our research findings.
2.5 Before Q90	Control	We are almost at the end of the survey. To help us improve our questions in the future, here are a few questions which are similar to others at the beginning of this survey.
	Positive emphasis	We are almost at the end of the survey. To help us improve our questions in the future, here are a few questions which are similar to others at the beginning of this survey. Please answer carefully to make the most of your effort in completing this survey.
	Negative emphasis	We are almost at the end of the survey. To help us improve our questions in the future, here are a few questions which are similar to others at the beginning of this survey. Please answer carefully so that your effort in completing this survey is not lost.
WAVE 4		
4.1 After intro	Control	N/A
	Positive emphasis	Answering questions carefully can greatly improve the accuracy and usefulness of the findings from this study. The more accurate your answers, the better the data will describe the opinions of people living in [COUNTRY].
	Negative emphasis	Answering questions carelessly can greatly reduce the accuracy and usefulness of the findings from this study. The less accurate your answers, the worse the data will describe the opinions of people living in [COUNTRY].
4.2 Before Q15	Control	N/A
	Positive emphasis	We really appreciate your continued co-operation. When you answer all questions, you help the results of the survey be more accurate and more useful. Up next, we have a few more statements that might describe you.
	Negative emphasis	We really appreciate your continued co-operation. If you stopped answering questions now, the results of the survey could be less accurate and less useful. Up next, we have a few more statements that might describe you.
4.3 Before Q36	Control	The next section of the survey is about environmental protection.
	Positive emphasis	The next section of the survey is about environmental protection. It is very important for this study to get answers from people with different opinions about the environmental problems and how, if at all, we should tackle them. Please continue completing the survey to help make sure that various opinions about this topic are represented in the study.
	Negative	The next section of the survey is about environmental

	emphasis	protection. It is very important for this study to get answers from people with different opinions about the environmental problems and how, if at all, we should tackle them. Please continue completing the survey to avoid having an unbalanced representation of opinions about this topic.
4.4 Before Q51	Control	NA
	Positive emphasis	We appreciate your continued effort, because studies have shown that answering questions carefully can greatly improve the accuracy and usefulness of our research findings.
	Negative emphasis	We appreciate your continued effort, because studies have shown that answering questions carelessly can greatly reduce the accuracy and usefulness of our research findings.
4.5 Before Q60	Control	We are almost at the end of the survey. To help us improve our questions in the future, here are some final questions about your experience completing this survey.
	Positive emphasis	We are almost at the end of the survey. To help us improve our questions in the future, here are some final questions about your experience completing this survey. Please answer carefully to make the most of your effort in completing this survey.
	Negative emphasis	We are almost at the end of the survey. To help us improve our questions in the future, here are some final questions about your experience completing this survey. Please answer carefully so that your effort in completing this survey is not lost.
WAVE 6		
6.1 After intro	Control	N/A
	Positive emphasis	Answering questions carefully can greatly improve the accuracy and usefulness of the findings from this study. The more accurate your answers, the better the data will describe the opinions of people living in [COUNTRY].
	Negative emphasis	Answering questions carelessly can greatly reduce the accuracy and usefulness of the findings from this study. The less accurate your answers, the worse the data will describe the opinions of people living in [COUNTRY].
6.2 Before Q15	Control	N/A
	Positive emphasis	We really appreciate your continued co-operation. When you answer all questions, you help the results of the survey be more accurate and more useful.
	Negative emphasis	We really appreciate your continued co-operation. If you stopped answering questions now, the results of the survey could be less accurate and less useful.

6.3 Before Q29	Control	There are many different views as to what makes a society fair or unfair. In the next section, we would like to ask you for your personal opinion about this topic.
	Positive emphasis	There are many different views as to what makes a society fair or unfair. In the next section, we would like to ask you for your personal opinion about this topic. It is very important for this study to get answers from people with different opinions. Please continue completing the survey to help make sure that various opinions are represented in the study.
	Negative emphasis	There are many different views as to what makes a society fair or unfair. In the next section, we would like to ask you for your personal opinion about this topic. It is very important for this study to get answers from people with different opinions. Please continue completing the survey to avoid having an unbalanced representation of opinions.
6.4 Before Q48	Control	Now we have some questions about your views on the Internet.
	Positive emphasis	Now we have some questions about your views on the Internet. We appreciate your continued effort, because studies have shown that answering questions carefully can greatly improve the accuracy and usefulness of our research findings.
	Negative emphasis	Now we have some questions about your views on the Internet. We appreciate your continued effort, because studies have shown that answering questions carelessly can greatly reduce the accuracy and usefulness of our research findings.
6.5 Within Q72	Control	In which year and month did you complete your highest educational qualification?
	Positive emphasis	In which year and month did you complete your highest educational qualification? Please remember that accurate information can greatly increase the value of our research.
	Negative emphasis	In which year and month did you complete your highest educational qualification? Please remember that inaccurate information can greatly reduce the value of our research.
6.6 Before Q80	Control	N/A
	Positive emphasis	We are almost at the end of the survey. Please keep answering the remaining questions carefully, because complete data are easier for researchers to analyse.
	Negative emphasis	We are almost at the end of the survey. Please keep answering the remaining questions carefully, because incomplete data are more difficult for researchers to analyse.

Appendix B. Questions where non-differentiation was calculated

Wave 1/ Set 1: A few questions about parenting

Question number	How much do you [agree or disagree] with this statement?
w2q55	Children's needs should come before those of their parents.
w2q56	It's best that parents listen to the parenting advice of professionals rather than simply rely on family and friends.
w2q57	It is alright for parents to, now and then, be less available for their children.
w2q58	A child's successes and failures mostly reflect how well their parents are bringing them up.
w2q59	A family's daily routine should be organised around what works best for parents rather than for their children
w2q60	Parents who seek advice on parenting are admitting that they are not doing a very good job.
w2q61	Parents should always be available for their children.
w2q62	Parents need to give children the freedom to learn from their own mistakes.
w2q63	Children should be the centre of their parents' attention.
w2q64	Parents naturally know how best to bring up their children.
w2q65	To reach their full potential, it is important that children take part in a wide range of organised activities outside of their home.
w2q66	Good parents constantly worry about their child's well-being and comfort.
w2q67	It is a parents' role to discover and develop their children's special talents.
w2q68	Parents should make sure their children's basic needs are met, even if it means cutting down on essentials for themselves.

w2q69	Parents have a life of their own and should not be asked to sacrifice their own well-being for the sake of their children.
w2q70	Good parents are those who buy their children the latest toys and gadgets.
w2q71	Parenting is very stressful if you want to do it right.
w2q72	Nowadays, it is so much more difficult to be a good parent than it used to be.
w2q73	Today's parents try to control too much of their children's lives.
w2q74	Today's parents feel a lot of pressures to do everything right for their children.
w2q75	Watching children grow up is life's greatest joy.
w2q76	People who have never had children lead empty lives.
w2q77	Having children interferes too much with the freedom of parents.

Response scale:
1 Strongly agree 2 Agree 3 Neither agree nor disagree 4 Disagree 5 Strongly disagree (Don't know) (Prefer not to answer) (No answer)

Wave 4 / Set 1: How the respondent characterises him/herself regarding own opinions

Question number	How well does this statement describe you?
w4q5	You prefer to say what you are thinking, even if it is inappropriate for the situation.

w4q6	You show your inner feelings even if it disturbs the harmony in your family.
w4q7	You feel uncomfortable when you express disagreement with members of your family.
w4q8	You are comfortable expressing disagreement with friends.
w4q9	You avoid expressing disagreement in front of strangers.
w4q10	You try to adapt to people around you, even if it means hiding your inner feelings.
w4q11	You see yourself as unique and different from others.
w4q12	Being different from others makes you feel uncomfortable.
w4q13	You like being different from other people.
w4q14	You try to avoid being noticeably different from others.

Response scale	Response scale Experiment 2 for experimental group 2
Does not describe me at all	Not at all
Does not describe me very well	A little
Does describe me moderately well	Moderately
Does describes me very well	Very well
Does describes me perfectly	Exactly
(Don't know)	(Don't know)
(Prefer not to answer)	(Prefer not to answer)
(No answer)	(No answer)

Wave 4/ Set 2: how the respondent characterises him/herself regarding own problem-solving capacities and interpersonal relationships

Question number	How well does this statement describe you?
------------------------	---

w4q15	Other people have great influence over the choices you make.
w4q16	You make decisions about your life on your own.
w4q17	You prefer to rely completely on yourself rather than depend on others.
w4q18	You try to avoid being reliant on others.
w4q19	You prefer to ask other people for help rather than rely only on yourself.
w4q20	You feel uncomfortable in situations where you have to rely only on yourself.
w4q21	You always ask your family for advice before making a decision.
w4q22	You prefer to do what you want without letting your family influence you.
w4q23	Your happiness is unrelated to the happiness of your family.
w4q24	If someone in your family is sad, you feel the sadness as if it were your own.
w4q25	You value good relationships with the people close to you more than your personal achievements.
w4q26	Your own success is very important to you, even if it disrupts your friendships.
w4q27	You value personal achievements more than good relationships with the people close to you.
w4q28	You behave in the same way even when you are with different groups of people
w4q29	You always see yourself in the same way even when you are with different people.

w4q30	You act very differently at home compared to how you act in public.
w4q31	You see yourself differently in different social environments.
w4q32	You behave differently when you are with people very different from you.
w4q33	You behave the same way at home and in public.
w4q34	You are considerate and kind to others.
w4q34	You have a forgiving nature.

Response scale	Response scale Experiment 2 for experimental group 2
Does not describe me at all Does not describe me very well Does describe me moderately well Does describes me very well Does describes me perfectly (Don't know) (Prefer not to answer) (No answer)	Not at all A little Moderately Very well Exactly (Don't know) (Prefer not to answer) (No answer)

Wave 4/ Set 3: Respondent's opinion about the environment

Question number	How much do you agree or disagree with this statement?
w4q36	We worry too much about the future of the environment.
w4q37	Modern lifestyle in [COUNTRY] is harmful to the environment.
w4q38	Modern science will solve our environmental problems with little change to our way of life.

w4q39	I am willing to pay much higher taxes in order to protect the environment.
w4q40	What humans do have no negative impact on the environment.
w4q41	Economic growth always harms the environment.
w4q42	In order to protect the environment, [COUNTRY] needs economic growth.
w4q43	We need to change our lifestyles if we want to protect the environment.
w4q44	We don't pay enough attention to environmental issues.
w4q45	Global warming will be very harmful for the environment.
w4q46	Products designed to protect the environment are unnecessarily expensive.
w4q47	It is my duty to do what I can to protect the environment.

Wave 4/ Set 4: how the respondent characterises him/herself regarding own opinions and problem-solving capacities

Question number	How much do you agree or disagree with this statement?
w4q54	I really enjoy a task that involves coming up with new solutions to problems.
w4q55	I like to have the responsibility of handling a situation that requires a lot of thinking.
w4q56	I would prefer complex to simple problems.
w4q57	I have strong opinions even when I am not personally involved.
w4q58	I have many more opinions than the average person.

w4q59	It is very important to me to hold strong opinions.
-------	---

Response scale:	Response scale Experiment 2 for experimental group 2
1 Strongly agree 2 Agree 3 Neither agree nor disagree 4 Disagree 5 Strongly disagree (Don't know) (Prefer not to answer) (No answer)	1 Strongly agree 2 Somewhat agree 3 Neither agree nor disagree 4 Somewhat disagree 5 Strongly disagree (Don't know) (Prefer not to answer) (No answer)

Wave 6/ Set 1: respondent's opinion regarding the influence of factors in finding a job

Question number	In your opinion, how much influence does this factor have on the decision to recruit a person for a job in [COUNTRY]?
w6q22	The person's educational qualifications.
w6q23	The person's knowledge and skills.
w6q24	The person's on-the-job experience.
w6q25	Whether the person knows someone in the organization.
w6q26	Whether the person has an immigrant background.
w6q27	The person's gender.
w6q28	Whether the person is willing to work for a low wage.

Response scale:
1 Not much or no influence 2 Some influence 3 Quite a lot influence 4 A great deal of influence
(Don't know) (Prefer not to answer) (No answer)

Wave 6/ Set 2: respondent's opinion about the internet

Question number	In your opinion, how much influence does this factor have on the decision to recruit a person for a job in [COUNTRY]?
w6q48	Going online is an efficient way to find information.
w6q49	The Internet makes life easier.
w6q50	The Internet helps me save time.
w6q51	Going online helps me pass the time when I am bored.
w6q52	When I am online I do not feel lonely.
w6q53	Going online allows me to keep in touch with people.
w6q54	It is easier for me to meet people online than in person.
w6q55	The Internet is frustrating to use.
w6q56	There is too much inappropriate and bad material online.

Response scale:
1 Strongly agree 2 Agree 3 Neither agree nor disagree 4 Disagree 5 Strongly disagree
(Don't know) (Prefer not to answer) (No answer)

Appendix C. Survey evaluation questions

Question	Response scale
How difficult was it for you to understand and answer the questions in this survey? (w2q94/ w4q60/w6q77)	Not at all difficult Slightly difficult Moderately difficult Very difficult Extremely difficult (Don't know) (Prefer not to answer)
And how much did you work at providing the most accurate answers you can to the questions in this survey? (w2q95/w4q61/w6q78)	Not at all A little A moderate amount A lot A great deal (Don't know) (Prefer not to answer)
How much did you enjoy answering this survey? (w2q96/w4q62/w6q80)	Not at all A little A moderate amount A lot A great deal (Don't know) (Prefer not to answer)