

## CONTEXT EFFECT IN SOCIAL SURVEYS: WORDING OF SENSITIVE CONCEPTS<sup>1</sup>

**Abstract.** *Many researchers confirm that survey responses are significantly influenced by the context in which a survey is conducted. In a broader sense, we are talking about the actors' social and cultural context that determines the characteristics of the interview as a social event. The validity of research findings is also influenced by the narrower (local) context, or the technical and substantive relevance of the instrument, including the wording of the questions. The paper empirically analyses the effects of the wording of a survey question taken from the International Social Survey Programme (ISSP). The survey was performed on a sample of the adult Slovenian population. We compare answers to two versions of the same survey question where, in the first version, the word "violence" was used and, in the second, the word "revolution". The starting assumption is that, when used in the Slovenian social and historical context, these terms trigger different answers from respondents. The results show that effects are not visible at a summary level of all respondents, but may be detected in individual respondent categories.*

**Keywords:** *context effect, ISSP, survey methodology*

### Introduction

Many researchers confirm that summary results of individual responses given in social surveys are to a considerable extent also an outcome of the context in which the survey questions was asked (e.g. Smith, 1987; Zaller and Feldman, 1992; Tourangeau and Smith, 1996; Cinnirella, 1998; Tourangeau et al., 2003; Cocco and Tuzzi, 2013; Hafner-Fink and Uhan, 2013; Boukes and Morey, 2018; Olson et al., 2019). The effects of the context also encompass how individual survey questions are verbalised, also known as the "wording effect". This is not only a context effect that refers to the 'technical'

---

\* Samo Uhan, PhD, Associate Professor, Faculty of Social Sciences, University of Ljubljana, Slovenia; Mitja Hafner Fink, PhD, Associate Professor, Faculty of Social Sciences, University of Ljubljana, Slovenia.

<sup>1</sup> The authors acknowledge financial support from the Slovenian Research Agency (research core funding No. P5-0151).

relevance of the instrument (the survey questionnaire) but, beyond that, it introduces the broader social and cultural context into the survey situation. Researchers find that the wording effect continues to be the most problematic and perhaps worst-developed area of social survey methodology. While writing about this already years ago, Groves (1989) found no general theory was able to predict the presence or absence of such an effect. There is also no consensus among researchers on how common these effects are, or how the method of collecting data relates to these effects. Smith (1992) illustrates the situation in methodological studies of contextual effects with a medical situation in the late nineteenth century when doctors were diagnosing a wide range of diseases such as cancer.

In this article, we deal with the problem of different wording of a survey question, which is supposed to measure the same concept. We show results of an empirical test of the wording effect within the framework of an actual survey situation with respect to the Slovenian Public Opinion survey. More precisely, it concerns the effect of two different wordings of the same survey question from the ISSP module "Role of Government". In so doing, we are particularly interested in how these effects are reflected when *sensitive questions* or concepts are being used (cf. Turangeau and Yan, 2007). We assume that when asking about sensitive concepts different verbalisation triggers (or intensifies) the effect of the (wider) socio-cultural context, such as historical experience, identity or language.

## Theoretical framework

We assume that opinion polls do not simply measure the actual opinions held by respondents, but also reflect the process of communication that arises in an individual survey situation. This includes effects of various aspects of communication that different disciplines deal with, from (social) psychology, (socio) linguistics, communication science, to methodology. Specifically, we mean both the narrower context of the survey situation (e.g. the technical characteristics of the survey instrument) and the broader context that includes the personal (psychological), cultural and social context of the actors engaged in the survey situation (see Uhan, 1998). Respondents' selection of answers in the survey situation is therefore influenced by the whole measurement process so, even in a technical sense, classical test theory is insufficient when testing the quality of the measurement result, which only takes account of the actual value and the measurement error. We hence consider a more appropriate extended measurement model which, on top of the actual value and measurement error, takes account of the content of the measured concept, the measurement instrument's characteristics and the specificity (uniqueness) of the individual measurement.

As mentioned, our interest is in the effect of the survey instrument not in a technical sense, but as a trigger for the broader socio-cultural context. Namely, respondents' responses are not solely affected by the questionnaire's formal structure, but also by the semantic level of the questions (the question wording). Many studies show that relatively minor changes in the way questions are formulated can lead to extensive changes in respondents' responses (Schuman and Presser, 1981; Schwarz and Strack, 1991; Hollbrook et al., 2007; Olson et al., 2019). Although the phenomenon is recorded relatively often (Hippler et al., 1987), methodological literature is unable to properly explain its occurrence. Esser (in: Alwin, 1991: 17) suggests, for example, that survey data should be understood and interpreted according to the social situation that determines the responses given by motivating respondents to respond to social stimuli. In our case, the social situation should not simply be understood as the survey situation (communication between interviewer and respondent), but also as the context or a frame defined by the specific wording of questions. Empirical data are the result of the "bounded rationality" (Esser, 1993) of respondents. While interpreting the data, it must be assumed that people are at any time (including a survey interview) striving to achieve personal satisfaction or reduce the psychological tensions associated with an action. In the survey situation, respondents seek to evaluate the consequences of possible responses and opt for alternatives that bring the most 'social benefits'. Therefore, if researchers wish to anticipate the behaviour of respondents, they must first understand their 'motivational structure'. For example, a respondent may provide valid data or merely meet the researchers' expectations according to the modalities the research instrument offers. The respondent selects alternatives in accordance with their interests in the survey situation, making their choice rational, even if their actions do not actually meet the criteria of objective reality. The respondent is, in short, motivated to express an opinion that will reduce the accumulated tension that inevitably occurs as they form opinions.

The 'answering questions' process in the survey situation context is more than just a simple question-and-answer procedure. We can only start to analyse the content of the answers by knowing the context that shapes the respondents' beliefs, which also requires an understanding of the characteristics of public communication and the social environment in which the respondents' attitudes are formed.

In this regard, a 'contextual approach' is taken in public opinion research that takes account of the *wider socio-cultural ("global") context* and 'collective' beliefs held by individuals. What opinion polls mostly show are projections of internal psychic needs or characteristics that are only indirectly related to public affairs. At the same time, the results measured are 'external'

factors which are not subject to critical judgement and assessment. The results of research in this sense represent individual or social views of the world, which are fundamentally different from aggregated and generalised research results. The results do not summarise the views held by the individuals and thus do not predict specific actions. Therefore, in this case, we cannot talk about any transfer but the results of opinion polls may therefore simply be understood as consolidated and articulated 'public attitudes'.

The validity of research findings, to the same extent as for the broader social context, is also influenced by the *narrower ("local") context*, or the technical and substantive relevance of the instrument, which includes the standardisation of procedures, the survey mode, the questionnaire's formal structure, and the wording of questions and modalities. Schuman and Presser (1981) note that respondents' answers are largely influenced by how questions are specifically formulated. Various studies point to the problem of the instability of responses and the important influence of the characteristics of the survey instrument (survey questions) on this lack of stability (e.g. Zaller and Feldman, 1992; Olson et al., 2019).

In the process of answering questionnaires, the two contexts mentioned above intertwine. We believe this is especially pronounced when questions raise the problem of so-called social desirability.

## Research problem

The problem of the effects of the wording on answering the questionnaire means that our research problem (explicitly or implicitly) also includes the following social survey topics, typically dealt with separately: conceptual and functional equivalence, context of the survey (local and global), cognitive processing of information, and problem of social desirability.

Within the framework of the international ISSP 2016 Role of Government study, as part of the issues of tolerance and extreme forms of political activity, two questions were posed to the Slovenian respondents that allowed them to express their views on the actions of people whose views most people would regard as extreme. Starting with a hypothetical situation, we asked the respondents two versions of one question: (a) "Consider people who want to overthrow the government using *violence*. Do you think such people should be allowed to hold public meetings to express their views?"; and (b) "Consider people who want to overthrow the government through *revolution*. Do you think such people should be allowed to hold public meetings to express their views?" We assume that the use of different wording ("violence" vs. "revolution") created different contexts that triggered different respondent reactions.

The background of our experiment is not merely a theoretical reflection,

but an actual survey situation. The Slovenian version of the ISSP question differs from the original regarding use of the words “overthrow the government through revolution” found in the original ISSP questionnaire and instead uses the words “overthrow the government using violence”.

Although the research team’s decision to modify the question is not documented, an attempt to reconstruct the circumstances that led Slovenian researchers to use the term “violence” reveals motives/justifications for their decision.

The question was first included in the Slovenian Public Opinion (SJM) survey as part of the ISSP Role of Government module in 1989, when Slovenia was formally part of the Yugoslav Federation with a socialist system, and not yet a member of the ISSP research group. The reason the researchers replaced the term “revolution” with “violence” lies in the mentioned historical and political context of the then socialist society, when the word “revolution” held a specific connotation that the Slovenian researchers believed did not coincide with the ISSP module’s research intentions. During socialism, “revolution” held primarily positive and emancipatory connotations, without clearly suggesting “violent overthrow” as stated in the original ISSP question. In the context of the former Yugoslav socialist system, which essentially emerged by way of revolution, use of the word “revolution” in this sense is paradoxical because the word “counterrevolution” was used to denote the opponents of the socialist regime. At the same time, in 1989, transitional processes were already underway that legitimated positions problematising socialism and (socialist) revolution. In this context, use of the word “revolution” would therefore have had unpredictable and inconsistent effects on the answers. Using the original wording would thus induce semantic noise in this respect. We assume this explains why use of the more neutral term “violence” seemed more appropriate. (Probably) due to the tendency to use equivalent measurements over time, the SJM survey continued to use this word before the latest implementation of the Role of Government module (ISSP, 2016).

The problem is noted by Cicourel (1982) in a discussion on the validity of public opinion data. Cicourel contended the research results are not valid unless contextual (ecological) validity is provided, understood as the synchronisation of the ‘artificial’ survey situation with real life. This can also be shown as a problem of functional equivalence, which is particularly relevant in cross-cultural comparative research. This kind of validity can be achieved by matching the formulations prepared by researchers and the spontaneous understanding the respondents develop (Cicourel, 1982).

In this context, we must mention the “indexicality” concept used by Harold Garfinkel in the setting of ethnomethodological research (Garfinkel, 1987). With use of this term, Garfinkel wished to stress that terms or symbols

hold different meanings in different contexts (see Giddens, 1989: 40). This understanding built on Bar-Hillel's use of the term "indexical expressions" by which Bar-Hillel explained that the meaning of words depends on the circumstances/contexts in which they were pronounced (Bar-Hillel, 1954).

The decision made by the SJM group researchers can therefore be explained by understanding the concept of validity: from the researchers' point of view, the data are valid if they reflect the conceptual idea. Compared to exact measurements, the survey remains the interaction event, regardless of the degree of instrument standardisation. Thus, the key difference between the two question variants is their use of words with different semantic dimensions.

Here we recall the circumstances of our experiment, (1) in the first version, respondents answered the question of whether people who want to overthrow the government using violence should be allowed to hold public meetings; while (2) in the second version whether people who want to overthrow the government through revolution should be allowed to hold such meetings. Starting from the wider historical, political and social context, i.e. post-socialist society (Slovenia), we conclude that the question containing the word "revolution" with the expected (mentioned) connotations falls in the 'sensitive question' category. As a reference, we rely on the following definition of a sensitive question:

*A question is sensitive when it asks for a socially undesirable answer, when it asks in effect that respondent admits he or she has violated a social norm.* (Tourangeau and Yan, 2007: 860)

Tourangeau and Yan (2007) believe sensitive questions are often discussed as part of the broader problem of social desirability (Näher and Krumpal, 2012; Blair et al., 1977; Tourangeau et al., 2000; Schwarz and Bless, 1992 and 2007; DeMaio, 1985). When answering questions about "revolution" during a period of socialism (such as an authoritarian regime), respondents were more likely to think about the desirability of possible answers – e.g. by thinking about which kinds of answers would not cause them problems, and which kind might cause them problems.

Thus, the specific purpose of this research is to test the 'wording effect' thesis. We wished to test how the use of different semantic aspects of the same question triggers changes in respondents' reactions to the question. We are in fact interested in the influence of characteristics of the measurement instrument (question wording) on a respondent's *cognitive processes*, which leads to the selection of a specific response category. Our goal is to discover two types of effect (by answering two research questions). We ask: (a) whether the question wording (of 'sensitive concepts') can affect

the consistency of respondents' answering (whether different wording may produce a different answer); and (b) whether the use of 'sensitive concepts' can bring differences in how the concept being measured is perceived among different social groups with different social norms.

We observe these effects by examining whether respondents holding different values and political orientations respond differently to the change in wording. We therefore wonder whether the change in wording reinforces (or weakens) the impact of the broader socio-cultural context in the form of collective consciousness, which we observe here through the current *ideological political divisions* as significantly determined by events during the Second World War, the socialist revolution, and in the period of socialism after the said War. We expect that use of the word "revolution" strengthens the effect of this context such that differences in responses between respondents with left and right political affiliations will increase. We also believe that differences between respondents with varying attitudes to the past (the period mentioned above) will become more pronounced. Below, we refer to this hypothetical starting point as an 'ideological thesis'<sup>2</sup>.

### Data, method and analytical model

880

To test our expectations about the wording effect, we introduced an experiment-like situation in the framework of the regular social survey situation of the Slovenian Public Opinion 2015 survey. The survey was performed on a probability sample of adult inhabitants (aged 18 years or older) of the Republic of Slovenia (N=1024). Data were collected between November 2015 and February 2016 (Hafner Fink and Malešič, 2016).

The question using the word "violence" was found in the first part of the questionnaire (within the ISSP Role of Government module) while the question using the word "revolution" came at the end of the questionnaire (after the demographic questions). In the second version of the question ("revolution"), respondents were reminded that this question was being asked again. We therefore asked all respondents to answer both questions. Between these two versions, more than 100 questions were asked, making it safe to assume the first version of the question had no (direct) influence on answers to the second version of the question. All respondents were asked in the same manner. In order to approach an experimental situation, we replaced "control group" with "order of questions" (the considerable distance between the two versions), so as to neutralise potential effects of the

---

<sup>2</sup> Therefore, we are not interested in how individuals' personality traits influence the way they respond to a change in wording, but are concerned with the question of whether different wording triggers a change in the broader socio-cultural context that influences responses.



first version of the question. This approach enabled us to observe the effects of wording at the individual level and the (expected) change in response between the first and second versions of the question, which is in line with our research question. A pure experimental situation was not ensured in that the 'micro location' of the two versions of the questions is different – the first is at the start of the questionnaire and the second at the end. We also did not use a control group.

The exact wording of the two questions is as follows:

Version using the word “*violence*” (in the first part of the questionnaire):

*There are some people whose views are considered extreme by the majority. Consider people who want to overthrow the government using violence. Do you think such people should be allowed to ...  
... hold public meetings to express their views?*

The original ISSP version using the word “*revolution*” (at the end of the questionnaire) read:

*There are some people whose views are considered extreme by the majority. Consider people who want to overthrow the government through revolution. Do you think such people should be allowed to ...  
... hold public meetings to express their views?*

Answer categories for both versions are as follows:

- 1 – definitely allowed
- 2 – allowed
- 3 – not allowed
- 4 – definitely not allowed
- 8 – don't know, can't choose

We test the effect of using the different terms (“*violence*” and “*revolution*”) in three steps:

- The *first step* was conducted at the “*aggregate*” or “*summary*” level. We compare the frequency distribution of answers to both question versions. Namely, we wish to check whether a different level of (in)tolerance towards “*extreme*” groups can be measured with a different wording of the question for the whole sample (representing the entire Slovenian adult population). We also identify the matching of answers by calculating the correlation between answers to both versions of the question.
- In the *second step*, we check whether, regardless of the possible matching of frequencies at the population level, some individuals choose different answers when different terms are used – e.g., in the first version they are tolerant and in the second version they are intolerant of extreme



groups. We simply cross-tabulate the answers to both questions to identify those individuals who changed their answer between the first and second versions of the question.

- In the *third step*, we try to determine individual factors that may influence how the two versions of the question were answered and the change in response at the individual level. We used three models of logistic regression analysis where the dependent variables are as follows: 1) version using the word “violence”; 2) version using the word “revolution”; 3) change versions one and two of the question. The following factors are included in the model: attitudes towards the Slovenian ‘partisans’ (members of a Slovenian WW2 resistance movement), left-right political orientation, religiosity, age, education, and interest in politics.

## Results

The results are shown separately for each step. We first observe possible net changes at the aggregate level (the level of the whole sample). In the next step, we observe changes in individual responses and, finally, the influence of the factors on the answers given to different versions of the question and on the change in answers between the first and second question versions.

### *Wording effect on the population level (net ‘changes’ of frequency distribution)*

When simply comparing the frequency distributions of answers between the two question versions, we can only detect minor differences (slightly smaller proportion of ‘tolerance’ when the word “revolution” is used) that are not statistically significant – in both cases, approximately 60% of respondents choose the answer “not allowed” and approximately one-third the answer “allowed” (see Table 1). Yet, the correlation between answers to the first (“violence”) and second (“revolution”) question version is not perfect (Spearman’s  $\rho = 0.667$ ), indicating that individual respondents did not consistently answer the two versions of the question. This result already supports our hypothesis about the effect of using two different terms. Therefore, in the next step, we look more closely at what is happening at the level of individual respondents.

*Table 1: FREQUENCY DISTRIBUTIONS OF ANSWERS TO QUESTIONS WITH DIFFERENT WORDING (IN %)*

	Definitely allowed	Allowed	Not allowed	Definitely not allowed	Don't know	No answer
<i>Version 1 (first half of the questionnaire):</i>						
Consider people who want to overthrow the government <b>by violence</b> . Do you think such people should be allowed to... hold public meetings to express their views?	5.8	29.7	34.9	24.3	5.3	0.0
merged categories	35.5		59.2		5.3	
<i>Version 2 (end of the questionnaire):</i>						
Consider people who want to overthrow the government <b>through revolution</b> . Do you think such people should be allowed to... hold public meetings to express their views?	2.9	30.4	36.4	24.7	5.1	0.5
merged categories	33.3		61.1		5.6	

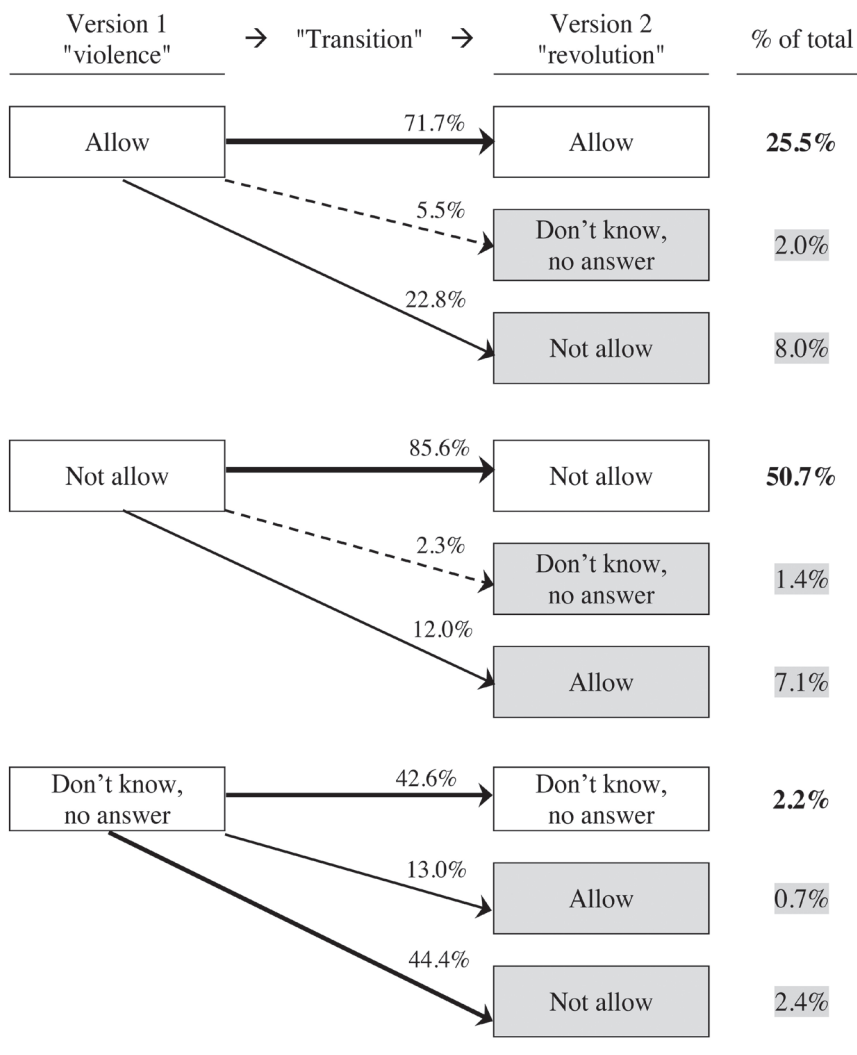
n = 1024

Source: Authors' own analyses based on SJM 2015 data.

### *Differences at the level of individuals*

When we cross-tabulate answers given to version 1 (“violence”) and version 2 (“revolution”), we find a substantial proportion of individuals changed their answer from the first to the second question version. If we understand changes only as a switch between three main positions/categories (“allow”, “not allow” or “don’t know/no answer”), then we establish that 21.6% of respondents changed their position between the first (at the start of the questionnaire) and second question versions (at the questionnaire’s end) (see Diagram 1). This proportion becomes much greater if we also consider changes in intensity within the same position – it reaches a value of 38.1%. A more detailed look shows that more than 15% of respondents adopted the opposite positions – they moved from “allow” to “not allow”, or vice versa (Diagram 1). We can also see that respondents who in the first version chose “not allow” were more ‘consistent’ (85.6% remained in the same position) than those who chose “allow” (just 71.7% of them stayed in the same position) (Diagram 1).

Diagram 1: TRANSITIONS (CHANGES) OF ANSWERS BETWEEN QUESTION VERSION 1 ("VIOLENCE") AND QUESTION VERSION 2 ("REVOLUTION") (SHADED AREAS SHOW CHANGES)



Source: Authors' own analyses based on the SJM 2015 data.

For further analysis, we also prepared a quantitative measure of the change in tolerance shown towards anti-government groups that came with the change in the question's wording – from using the word "violence" in the first version to using the word "revolution" in the second version of the question. For this purpose, we developed an '*index of change*' to measure

the change in tolerance on a scale from -4 to +4.<sup>3</sup> The lowest value (-4) measures the maximum shift in the direction of not allowing rallies to be held by anti-government groups, the value of 0 means stability (remaining in the same position), and the highest value (4) is the biggest shift in the direction of allowing anti-government groups to hold rallies (see Table 2). To use it in the logistic regression model, this variable is dichotomised in the following way: values above 0 (changes in the direction of greater tolerance – allowing rallies) are coded as 1 (16.5%) while all other values are coded as 0 (83.5%).

*Table 2: INDEX OF CHANGE OF TOLERANCE FOR ANTI-GOVERNMENT RALLIES – THE ‘DIRECTION’ OF CHANGING ANSWER FROM THE FIRST VERSION (“VIOLENCE”) TO THE SECOND VERSION OF THE QUESTION (“REVOLUTION”)*

	n	%
-4 direction “not allow”	2	0.2
-3	23	2.2
-2	65	6.3
-1	132	12.9
0 no change	634	61.9
1	94	9.2
2	56	5.5
3	15	1.5
4 direction “allow”	3	0.3
Total	1024	100

Source: Authors’ own analyses based on the SJM 2015 data.

### *Factors in answering and/or mediators of the wording effect*

We narrowed the analysis to explain the ‘tolerant’ response to the question. Therefore, we are interested in: a) which factors best explain the choice of “(definitely) allowed” answer for both versions of the question; and b) which factors influence the change in the level of tolerance between the first (“violence”) and second versions (“revolution”) of the question. For the analysis, we prepared three dummy variables presenting a ‘tolerant’ response to the survey question:

- tolerance for anti-government rallies where the word “violence” is used;
- tolerance for anti-government rallies where the word “revolution” is used; and
- change in the direction of greater tolerance (or direction of less intolerance).

<sup>3</sup> When preparing the index, the answer “don’t know /can’t choose” was treated as the middle of the ‘scale’.

Thus, we tested three models based on the 'ideological thesis', where we assumed the following two variables are the main predictors of how the question on tolerance regarding actions against a government (regardless of the wording) is answered:

- *Left-right* political orientation as an indicator of current political (ideological) divisions is our first predictor based on the 'ideological thesis'. For the analysis, we used two dichotomous variables: a) voting for left parties at the last elections; and b) voting for right parties at the last elections.
- *Attitudes to the expression "Slovenian partisans"* (members of a Slovenian WW2 resistance movement) is important from the historical point of view: "Slovenian partisans" are understood as those who not only fought against the occupation during WW2 but also fought on the side of the communist *revolution*. It was prepared as a binary variable such that a positive attitude was coded with 1.

We also included the following four control variables in the model whose relevance is not so much due to the ideological thesis but to other aspects related to the context of the survey:

- *Religiosity* (measured by attendance at religious services) – an indicator of traditionalism, but also of *ideological* position, which is relevant when the word "revolution" is considered. We expected that religious respondents react less tolerantly when the word revolution is used. In the analysis, we used the dichotomous variable: 1 for attendance at religious services (special occasions like a wedding or funeral are not included) and 0 for no attendance.
- *Education level*, which may be understood as an indicator of social status and also an indicator of *cognitive sophistication* (see Uhan and Hafner-Fink, 2013). We expected that people with a higher level of cognitive sophistication (assuming that means those with a higher education) were more likely to perceive a difference in meaning and adjust their response accordingly (in line with their political value orientations). However, we generally expected that, regardless of the wording, the more educated respondents would express greater tolerance of anti-government rallies than the less educated. In the multivariate analysis, we used a 5-point scale from primary to university level.
- *Age*, which is understood here in two ways: 1) it may simply indicate differences between younger and older population due to the process of growing up (problem of maturing, which might be related with conservatism); or 2) it may be seen as an indicator of generations that each have a different (historical) experience (problem of socialisation). For use in the multivariate analysis, the variable was prepared as a 7-point scale, with each category denoting 10 years (1 – up to 30 years ... 7 – more than 80 years).

- *Interest in politics*, which may also be understood as an indicator of involvement in politics. Since attitudes towards anti-government groups is clearly a political issue, we estimate it makes sense to include this variable in the analytical model. It was measured on a scale from 1 (not interested) to 5 (very interested).

We also expect that these variables ‘mediate’ the wording effect by influencing the ‘consistency’ of answering the two versions of the question. We thus predict these variables influence the ‘direction’ of changing the answer from the first (“violence”) to the second version of the question (“revolution”). Left-right orientation, attitudes towards the partisans, and religiosity were included in the model to test our ‘ideological thesis’ about the effect of concept contextualisation by using different wording – in the first version “violence” and in the second version “revolution”. In general, we expect greater tolerance towards anti-government rallies among the left-oriented, among those with positive attitudes towards the partisans, and among non-religious respondents. The variable “age” was also included for the same reason – as a factor of concept contextualisation. Namely, we expect that older generations are more likely to react to different wording because of their historical experience (socialisation) with socialism when “revolution” held a mainly positive and emancipatory meaning. However, we also expect that different reactions to changed wording might also be attributed to the ageing process: it is anticipated that conservative values (e.g. reluctance to change) are stronger among the elderly (e.g. Schwartz, 2007), which means that young people are expected to be more tolerant than older ones when faced with use of the word “revolution”.

To investigate the influence of the mentioned factors (variables), we applied three different analytical methods – univariate analysis of variance, linear regression, and binary logistic regression (this article presents the results of this method) – to test three separate models for all three dependent variables<sup>4</sup>:

- tolerance for anti-government groups when the word “violence” was used (model 1);
- tolerance for anti-government groups when the word “revolution” was used (model 2); and

<sup>4</sup> We included in the models both the direct (main) effects of individual factors (independent variables) and two-way interaction effects. Results of these three methods did not differ substantially and also showed that interaction effects only contributed a small share to the whole model’s explained variance (e.g. in model 3, when all two-way interaction effects were included, the Adjusted R Squared reached a value of 0.292, whereas with only the main effects the value was 0.272). Multicollinearity test (for high correlations between independent variables) showed no problems: values of the Tolerance statistic within multiple linear regression models in all cases exceed 0.78.

- change in the direction of greater tolerance between the “violence” and “revolution” versions (model 3).

In the first, step we conducted some basic bivariate analysis of association of all three dependent variables with the independent variables (factors) mentioned above. In fact, we used a standard bivariate cross-tabulation with a Chi-squared test and Cramer’s V as a measure of association (see Table 3). Looking at Model 1 and Model 2, we see that all factors are statistically significant ( $p < 0.05$ ) in at least one case (Table 3). Yet, when it comes to Model 3 (changes leading to more tolerance), only *left-party preferences* shows a statistically significant association (Table 3).

It is evident that answers (tolerance for anti-government rallies) differ regarding *left-right political orientation*:

- *left-party supporters show more sympathy* for anti-government protests than others (only) when the term “revolution” is used (42.2% vs. 31.6%); and
- *right-party supporters show less tolerance* for anti-government protests (compared to all others) in both cases, although the differences are slightly larger when the term “violence” is used (26.2% vs. 37.4%) (Table 3).

We also find confirmation of the *relevance of the ‘ideological thesis’*: changes between the “violence” mentioned and “revolution” mentioned in the direction of greater tolerance for anti-government protests is significantly associated only with left-party preferences. Namely, there is a higher probability that left-oriented respondents would change their answer in the direction of showing greater support for protests than other respondents would (24.8% vs. 12.4%) (Table 3).

The ‘ideological thesis’ is further supported by the fact that answers also differ with respect to *religiosity* and *attitudes towards the partisans*:

- *respondents with positive attitudes to the partisans strongly support anti-government rallies* more than other respondents in both cases and, as expected, differences between the two groups are slightly bigger when the word “revolution” is used (39.4% vs. 28.1%); and
- *respondents who (regularly) attend religious services are less in favour of anti-government rallies* than other respondents in both cases, especially when the word “revolution” is used (28.2% vs. 44.6%) (Table 3).

Differences among the age groups support our thesis that this variable should be interpreted as an indicator of conservatism: *younger respondents show greater sympathy for anti-government protests than older ones*, which is less obvious with the word “violence”, yet with word “revolution” the differences are clear (50.5% vs. 26.0%) (Table 3).

Also statistically significant are the associations of both question versions with the variable *education*: in both cases, *more educated respondents are*



Table 3: OF ATTITUDES TOWARDS ANTI-GOVERNMENT RALLIES (BIVARIATE ANALYSIS)

		Ver. 1 ("violence") – allow rallies (%) (Model 1)	Ver. 2 ("revolution") – allow rallies (%) (Model 2)	Change from ver. 1 to ver. 2 – towards allow (%) (Model 3)
partisans	<i>negative</i>	31.7	28.1	13.5
	<i>positive</i>	40.1	39.4	15.4
		Chi-Sq. = 7.779** Cramer's V = 0.087	Chi-Sq. = 14.710** Cramer's V = 0.120	Chi-Sq. = 0.699 Cramer's V = 0.026
left-party preference	<i>no</i>	35.8	31.6	12.4
	<i>yes</i>	34.2	42.2	24.8
		Chi-Sq. = 0.160 Cramer's V = 0.013	Chi-Sq. = 6.867** Cramer's V = 0.082	Chi-Sq. = 17.096** Cramer's V = 0.129
right-party preference	<i>no</i>	37.4	34.9	14.7
	<i>yes</i>	26.2	25.0	12.5
		Chi-Sq. = 7.697** Cramer's V = 0.087	Chi-Sq. = 6.234* Cramer's V = 0.078	Chi-Sq. = 0.563 Cramer's V = 0.023
religious services	<i>no</i>	44.6	44.6	17.5
	<i>yes</i>	31.4	28.2	13.0
		Chi-Sq. = 16.826** Cramer's V = 0.128	Chi-Sq. = 26.651** Cramer's V = 0.162	Chi-Sq. = 3.745 Cramer's V = 0.061
age	<i>18 to 30</i>	44.1	50.5	17.0
	<i>31 to 50</i>	34.8	34.1	15.7
	<i>51 to 70</i>	33.5	27.2	13.7
	<i>above 70</i>	32.0	26.0	10.0
		Chi-Sq. = 7.657 Cramer's V = 0.087	Chi-Sq. = 35.181** Cramer's V = 0.186	Chi-Sq. = 3.992 Cramer's V = 0.063
education	<i>primary</i>	24.5	18.2	9.9
	<i>secondary</i>	36.4	33.1	14.5
	<i>tertiary</i>	41.6	44.1	17.1
		Chi-Sq. = 14.967** Cramer's V = 0.121	Chi-Sq. = 34.460** Cramer's V = 0.184	Chi-Sq. = 4.817 Cramer's V = 0.069

\* p &lt; 0.05; \*\* p &lt; 0.01

Source: Authors' own analyses based on the SJM 2015 data.

*more inclined to protests.* This association is stronger when the term "revolution" is used: among respondents who had completed primary school (or less) only 18.2% of respondents support anti-government protests, while in the group with some college or higher education this share exceeds 44% (Table 3).

To test the pure (direct) effects (controlled for other model variables) of the mentioned factors, we prepared three separate (multivariate) binary logistic regression models for three dependent variables (see Table 4). Overall results of testing these three models show the same pattern already observed on the bivariate level and presented in Table 3.

In the multivariate situation of logistic regression models, only *age* and *religiosity* show a consistent pattern of influence:

- a. *Older respondents* are less likely to give a 'tolerant' answer with respect to anti-government rallies than younger ones (regardless of the wording) (with odds ratios of 0.878 and 0.766) (see Table 3, Models 1 and 2). Older respondents are also less likely (Exp (B) = 0.829) to change their answer in the direction of tolerance when "revolution" is used compared to when "violence" is used (Table 4, Model 3).
- b. *Respondents who attend religious services* are also less likely to give a 'tolerant' answer in relation to anti-government rallies (odds ratios of 0.646 and 0.569) (see Table 3, Models 1 and 2). They are also less likely (Exp (B) = 0.548) to change their answer in the direction of tolerance when "revolution" is used compared to when "violence" is used (Table 4, Model 3).

*Table 4: PREDICTORS OF (CHANGE OF) ANSWERS ON TWO VERSIONS OF QUESTIONS ON TOLERANCE TO ANTI-GOVERNMENT GROUPS ("VIOLENCE" AND "REVOLUTION") – RESULTS OF BINARY LOGISTIC REGRESSION*

	Allow rallies <sup>a</sup> Version 1 – "violence" (Model 1)		Allow rallies <sup>a</sup> Version 2 – "revolution" (Model 2)		Change towards higher tolerance from version 1 to version 2 <sup>b</sup> (Model 3)	
Predictors:	B	Exp (B)	B	Exp (B)	B	Exp (B)
<i>Slovenian partisans (binary)</i>	0.291	*1.337	0.571	**1.771	0.325	1.384
<i>Voting left party (binary)</i>	-0.464	*0.629	0.055	1.056	0.708	**2.031
<i>Voting right party (binary)</i>	-0.483	*0.617	-0.192	0.825	0.031	1.031
<i>Religion services (binary)</i>	-0.437	**0.646	-0.564	**0.569	-0.602	**0.548
<i>Age (7 categories)</i>	-0.130	**0.878	-0.266	**0.766	-0.187	**0.829
<i>Education (5 categories)</i>	0.061	1.063	0.131	*1.140	0.184	*1.202
<i>Interest for politics (1-5)</i>	0.271	**1.311	0.184	**1.202	-0.149	0.862
<i>Allow rallies – version 1 (control variable)</i>	–	–	–	–	1.281	**3.602
<i>Constant</i>	-0.706	0.494	-0.519	0.595	-2.486	0.083
Hosmer-Lemeshow test:	$\chi^2 = 5.039$ ; df = 8; p = 0.753		$\chi^2 = 7.375$ ; df = 8; p = 0.497		$\chi^2 = 3.392$ ; df = 8; p = 0.907	
Model Summary:	Nagelkerke R <sup>2</sup> = 0.073		Nagelkerke R <sup>2</sup> = 0.127		Nagelkerke R <sup>2</sup> = 0.303	

<sup>a</sup> The dependent variable was prepared as follows: both answers presenting tolerance ('definitely allowed' and 'allowed') are coded as value 1, all other answers were coded as 0.

<sup>b</sup> Binary variable: 1 – change towards (higher) tolerance (index of change presented in Table 2 is higher than 0); 0 – other

\* p < 0.05; \*\* p < 0.01

Source: Authors' own analyses based on SJM 2015 data.

Left-oriented respondents regard the concept of “revolution” as more legitimate, seen in the fact that the probability of shifting towards greater tolerance for protests between version 1 (“violence”) and version 2 (“revolution”) is higher among left-oriented respondents (Exp (B) = 2.031) (Table 4, Model 3). The same pattern emerges for attitudes towards the Slovenian partisans: it is statistically significant that respondents with positive attitudes towards the partisans are more likely to support anti-government protests only when “revolution” is used (Exp (B) = 1.771) (Table 4, Model 2). We may observe the reverse pattern in the case of religiosity: the probability of a shift towards greater tolerance is substantially lower among respondents who regularly attend religious services (Exp (B) = 0.548) (Table 4, Model 3). All of this is in line with our ‘ideology thesis’.

The effect of age shows that older respondents are more conservative, meaning they do not support groups trying to overthrow the government. We cannot not explain the effect of age as the impact of social (historical) context, or by way of a nostalgia thesis that refers to the idealised view of life in former socialist Yugoslavia as established by socialist revolution. We instead understand the effect of age as an effect of conservatism.

## Conclusion

891

The results of our experiment show important findings concerning the use of different wording. The first of these relates to the first research question, whether or not the question wording (of ‘sensitive concepts’) can affect the consistency of respondents’ answers (whether different wording can lead to a different answer). At the summary level, it seems we are unable to confirm such an effect. An in-depth analysis at the level of responses of individual respondents reveals that a significant share of the respondents apparently reacted to the different wording. Namely, we see that only just over 60% of those did not change their answer and that in the second version more than 15% of all respondents gave the opposite answer to the first version.

We may also positively respond to another research question where we asked whether use of ‘sensitive concepts’ can create differences in the perception of the measurement concept between different social groups with different social norms. These differences become apparent when using the term “revolution”, which we understand as constituting a ‘sensitive term’. It emerges that there is increased differentiation in the answers of respondents holding different ideological and political value orientations. The results also confirm the relevance of cognitive sophistication since education is seen as a relevant factor in changing responses. Among the higher educated, the proportion of respondents who changed their answers was

higher than among the less educated. While this association is statistically significant, it is quite weak, suggesting that in our case the effect of wording (a change in responses to questions with different wording) is better explained by the 'ideological thesis' than by the hypothesis of the effect of cognitive sophistication.

How to explain the above effects? Our experiment shows that by altering the context in the sense of changed wording, we also trigger the functioning of a wider socio-cultural context in the form of collective consciousness, represented here by ideological political affiliation (left-right, attitude to recent history). Therefore, we can explain the different answers to the two question versions by stating that use of the word "revolution" (unlike in the question version using "violence") can trigger different reactions among respondents holding a different historical experience or different ideological and political affiliations. Namely, the word "revolution" (and its understanding) is strongly connected to historical circumstances and developments in Slovenia – the positive sense attributed to the term during socialism started to degrade during the process of the transition to a market economy.

What do our findings imply for the design of survey research? The first implication relates to the conceptualisation of survey research, especially international research. The second one concerns the methodological adequacy of survey design.

With regard to the conceptualisation of survey research, the problem of the validity of measurements is very important. As shown, when 'sensitive terms' are being used, the relevant influences of the wider socio-cultural context surface. In this way, there may be a different understanding of a certain question and, thus, a different 'identification' of the concept being measured among different social groups with a different value background. This problem is even more pronounced in international surveys where very dissimilar socio-cultural contexts can act differently on an identical question. In this setting, it is important to achieve the so-called functional (and conceptual) equivalence of translations, which is especially evident while translating the original version of a questionnaire into the languages of the participating countries. The problem is not simply linguistic, but also socio-linguistic.

Another important implication relates to the issue of testing the effects of using different wording. Although at the summary level we did not find significant differences between the answers to the different versions of the questionnaire, we established important changes at the individual level.

This test was made possible because the same respondents answered both versions of the question. This then meant that we shaped the circumstances similar to a true social science experiment. We therefore consider

that, in order to ensure the quality of research findings, the criteria of a classic experiment must be followed as far as possible. At the same time, it seems necessary to introduce mechanisms for ensuring methodological triangulation. In our case, it would be meaningful to conduct additional cognitive interviews to help reveal possible different understandings of the two alternate words “revolution” and “violence”.

#### BIBLIOGRAPHY

- Abramson, Paul R. and Charles W. Ostrom (1994): Question Wording and Partisanship: Change and Continuity in Party Loyalties during the 1992 Election Campaign. *Public Opinion Quarterly* 58 (1): 21–48.
- Alwin, Duane F. (1991): Research on Survey Quality. *Sociological Methods & Research* 20 (1): 3–29.
- Bar-Hillel, Yehoshuah (1954): Indexical Expressions. *Mind* 63 (251): 357–379. Accessible at <http://www.jstor.org/stable/2251354>, 13. 4. 2018.
- Blair, Ed, Seymour Sudman, Norman M. Bradburn and Carol Stocking (1977): How to Ask Questions about Drinking and Sex: Response Effects in Measuring Consumer Behaviour. *Journal of Marketing Research* 14 (3): 316–321.
- Boukes, Mark and Alyssa C. Morey (2018): Survey Context Effects and Implications for Validity: Measuring Political Discussion Frequency in Survey Research. *Journal of Survey Statistics and Methodology* 7 (2): 201–226.
- Cicourel, Aaron V. (1982): Interviews, Surveys, and the Problem of Ecological Validity. *The American Sociologist* 17 (1): 11–20.
- Cinnirela, Marco (1998): Manipulating Stereotype Rating Tasks: Understanding Questionnaire Context Effects on Measures of Attitudes, Social Identity, and Stereotypes. *Journal of Community & Applied Social Psychology* 8 (5): 345–362.
- Cocco, Michele and Arjuna Tuzzi (2013): New Data Collection Modes for Surveys: A Comparative Analysis of the Influence of Survey Mode on Question-Wording Effects. *Qual Quant* 47 (6): 3135–3152.
- DeMaio, Theresa J. (1985): Social Desirability and Survey Measurement: A Review. In: Charles F. Turner and Elizabeth Martin (eds.), *Surveying Subjective Phenomena* 2, 257–282. New York: Russel Sage Foundation.
- Esser, Hartmut (1993): Response Set: Habit, Frame or Rational Choice? In: Dagmar Krebs and Peter Schmidt (eds.), *New Directions in Attitude Measurement*, 293–314. Berlin, New York: Walter de Gruyter.
- Garfinkel, Harold (1987): *Studies in Ethnomethodology*. Cambridge: Polity Press.
- Giddens, Anthony (1989): *Nova pravila sociološke metode* (New rules of sociological method). Ljubljana: ŠKUC, FF.
- Groves, Robert M. (1989): *Survey Errors and Survey Costs*. New York: Wiley.
- Hafner-Fink, Mitja and Samo Uhan (2013): Bipolarity and/or Duality of Social Survey Measurement Scales and the Question-order Effect. *Qual Quant* 47 (2): 839–852.
- Hippler, Hans-J., Norbert Schwarz and Seymour Sudman (eds.) (1987): *Social Information Processing and Survey Methodology*. New York: Springer-Verlag New York Inc.

- Holbrook, Allyson L., Jon A. Krosnick, David Moore and Roger Tourangeau (2007): Response Order Effects in Dichotomous Categorical Questions Presented Orally: The Impact of Question and Respondent Attributes. *Public Opinion Quarterly* 71 (3): 325–348.
- Näher, Anatol-F. and Ivar Krumpal (2012): Asking Sensitive Questions: The Impact of Forgiving Wording and Question Context on Social Desirability Bias. *Qual Quant* 46 (5): 1601–1616.
- Olson, Kristen, Jolene D. Smyth and Amanda Ganshert (2019): The Effects of Respondent and Question Characteristics on Respondent Answering Behaviours in Telephone Interviews. *Journal of Survey Statistics and Methodology* 7 (2): 275–308.
- Schuman, Howard and Stanley Presser (1981): *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording, and Context*. New York: Academic Press.
- Schaeffer, Nora C. (1991): Conversation with a Purpose or Conversation? Interaction in the Standardized Interview. In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz and Seymour Sudman (eds.), *Measurement Errors in Surveys*, 367–391. New York: Wiley.
- Schwartz, Shalom H. (2007): Value Orientations: Measurement, Antecedents, and Consequences Across Nations. In Robert Jowell, Caroline Roberts, Rory Fitzgerald and Gillian Eva (eds.), *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, 169–203. London, Thousand Oaks, New Delhi, and Singapore: Sage.
- Schwarz, Norbert and Herbert Bless (1992): Constructing Reality and Its Alternatives: An Inclusion/exclusion Model of Assimilation and Contrast Effects in Social Judgment. In Leonard L. Martin and Abraham Tesser (eds.), *The Construction of Social Judgments*, 217–245. Hillsdale, NJ: Lawrence Erlbaum.
- Schwarz, Norbert and Herbert Bless (2007): Mental Construal Processes: The Inclusion/exclusion Model. In Diederik A. Stapel and Jerry Suls (eds.), *Assimilation and Contrast in Social Psychology*, 119–141. Philadelphia: Psychology Press.
- Schwarz, Norbert and Fritz Strack (1991): Evaluating One's Life: A Judgment Model of Subjective Well-Being. In Fritz Strack, Michael Argyle and Norbert Schwarz (eds.), *International Series in Experimental Social Psychology* 21: 27–47. Subjective Well-being: An Interdisciplinary Perspective. Elmsford, NY, USA: Pergamon Press.
- Smith, Tom W. (1987): That Which We Call Welfare by Any Other Name Would Smell Sweeter: An Analysis of the Impact of Question Wording on Response Patterns. *Public Opinion Quarterly* 51 (1): 75–83.
- Smith, Tom W. (1992): Thoughts on the Nature of Context Effects. In Norbert Schwarz and Seymour Sudman (eds.), *Context Effects in Social and Psychological Research*, 163–184. New York: Springer-Verlag.
- Tourangeau, Roger and Tom W. Smith (1996): Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly* 60 (2): 275–304.

- Tourangeau, Roger and Ting Yan (2007): Sensitive Questions in Surveys. *Psychological Bulletin* 133 (5): 859–883.
- Tourangeau, Roger, Lance J. Rips and Kenneth A. Rasinski (2000): *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.
- Tourangeau, Roger, Eleanor Singer and Stanley Presser (2003): Context Effects in Attitude Surveys: Effects on Remote Items and the Impact on Predictive Validity. *Sociological Methods & Research* 31 (4): 486–513.
- Uhan, Samo (1998): *Prava in neprava mnenja*. Ljubljana: Fakulteta za družbene vede.
- Uhan, Samo and Mitja Hafner Fink (2013): Context Effects in Social Surveys: Between Instrument and Respondent. *Teorija in praksa* 50 (1): 233–248.
- Zaller, John R. and Stanley Feldman (1992): A Simple Theory of the Survey Response: Answering Questions Versus Revealing Preferences. *American Journal of Political Science* 36 (3): 579–616.

#### SOURCES

- Hafner Fink, Mitja and Marjan Malešič (2016): *Slovenian Public Opinion 2015: Work Orientation (ISSP 2015), Role of Government (ISSP 2016), Mirror of Public Opinion and National Security Survey*. Accesible at [https://doi.org/10.17898/ADP\\_SJM15\\_V1](https://doi.org/10.17898/ADP_SJM15_V1), 20. 1. 2017.