Anvesak
Vol. 50 (2), pp. 1-21
Received: 27 Nov. 2020; accepted: 19 Mar. 2021

STYLE OF QUESTION MATTERS: AN EXPERIMENT WITH QUESTIONS ON GENDER VIOLENCE

MANORANJAN PAL[1]*, CHAITI SHARMA-BISWAS[1], SRIPARNA BANERJEE[2], ANJALI GHOSH[1], SUBHENDU CHAKRABARTI[1], SUMANA GUHA[3] AND PREMANANDA BHARATI[1]

[1]*Indian Statistical Institute, Kolkata, India*
[2]*West Bengal State University, Barasat, India*
[3]*St. Xavier's College, Kolkata, India*

Abstract: Reliable data in surveys is a pre-requisite to correct estimates. There are various kinds of 'response errors' in surveys which lead to biased or inconsistent estimates of the population parameters. Some of the response errors are not intentional but when it comes to opinion survey, it might often lead to asymmetric distribution of errors. This paper experimented with positive and negative styles of questions in the forms of statements on gender violence, which were canvassed among randomly drawn adult individuals. The analysis of the data using a statistical model revealed that, instead of a single set of questionnaires, one should make two sets – one positive and one negative and canvass the two sets to two independent samples in the population. The model can then be used to estimate the exact proportion of persons who accept the statement.

Key Words: Response error in survey, Style of question, Gender violence, Chi-square test, Statistical model

1. Introduction

A clear picture of the prevalence of intimate partner violence and non-partner sexual violence emerged from the report of WHO (WHO, 2013) in which it was shown that the lifetime prevalence of physical and/or sexual intimate partner violence among ever-partnered women is the highest in South-East Asia (37.7%). In other regions also this percentage is not much less. In Europe, Western Pacific and America, this accounts for 25.4%, 24.6% and 29.8% respectively, whereas in Africa it is 36.6%, very close to that of South-East Asia. In the same report, lifetime prevalence of non-partner sexual violence is found to be the lowest (4.9%) in the South-East Asian region.

In the Organization for Economic Co-operation and Development report, violence against women is defined by taking some kind of average of three components: Attitude, Actual Incidences

*Correspondence to: Professor Manoranjan Pal, Economic Research Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India. Email: manoranjan.pal@gmail.com

and Legal Protection (OECD, 2019). In that report, the violence against women in India is shown as 22.1%, which is a bit low when compared to Bangladesh (28.3), Sri Lanka (53.2%) and many other countries. However, in most of the European countries, it is around 10%.

By the term *gender violence*, it is usually meant to be against women and girl children. But if one looks at the 2010-2012 data or even more recent gender breakdown data on detected victims of trafficking for forced labour by region, then one can see that incidence of violence against men is not negligible. In fact, the incidence of (gender) violence against men is found to be higher than that against women in European countries and in America (UNODC, 2014). Generally speaking, gender violence remains a global issue. However, since the incidence of violence against women and small girls is much higher than that against men and small boys in large part of the world including India, this paper thus is concerned with violence against women. The most disappointing fact is that still now there are some countries in the regions of Northern Africa, Western Asia and sub-Saharan Africa, which have not enacted appropriate laws for addressing the issue of gender violence (United Nations, 2020).

Since it is a sensitive issue, reliable and timely data is critically needed. There is no doubt that quality of this type of data needs to be improved. Very delicate approach is needed to collect data on this issue. One should maintain two things while collecting data on gender violence – openness about the objective of data collection and confidentiality about the respondents. Especially, if one can ensure confidentiality, then the respondents may be convinced about their safety and thus reliable data may be obtained. Otherwise, it is almost impossible to extract the truth.

There are many precautions to be made while collecting sensitive data. Presence of other persons should be forbidden while collecting such data. For example, if one wants to find out whether husband beats wife and husband is present at the spot of data collection, then reliable data cannot be obtained.

Response to a question depends on many factors like the availability of time of the respondent, religious and social dictums, etc. But it is less known whether the statements of the questions like "should" vs. "should not", "is" vs. "is not", etc. matter in a significant manner. Present paper is an endeavour to deal with this matter. To verify this, a set of nine questions regarding family related matters and another set of fourteen questions regarding the various aspects of social views on gender violence have been canvassed among adult men and women (for the complete lists of questions see Tables A1 and A2 in Appendix A). The questions were in the form of statements. The respondents were asked to choose any one of the five alternatives such as "strongly agree", "somewhat agree", "somewhat disagree", "strongly disagree", and "don't know". This type of data can be analysed by using categorical response models like multinomial logistic regression, or ordered logistic regression models (Agresti, 2007; Gelman and Hill, 2007; Hilbe, 2009; Hosmer and Lemeshow, 2013; Menard, 2002; and Wooldridge, 2010), provided enough explanatory variables are available. However, in this paper, only the information on the sex of the respondent is available. The best that can be done is to prepare contingency table and test for independence of factors.

There are numerous kinds of "Errors in Response" in surveys (Lessler and Kalsbeek, 1992; Groves et al., 2004; Groves, 1989; and Lyberg and Kasprzyk, 1991), many of which lead to serious consequences like getting biased and/or inconsistent estimates of the population parameters (Kalton and Schuman, 1982). Scholars were aware of this as early as in the 1950s (Hansen et al., 1951). Apart from non-response to certain questions, there may be deliberate concealment/distortion of facts leading to errors-in-measurements. Some of the response errors are not intentional. But, when it comes to opinion survey, errors in response might lead to some consequences like asymmetric distribution of errors.

2. Literature Review

Literature with positive and negative questions as taken by the authors of this paper is rare, but there are many experiments that have been done with types of questions, number of categories in Likert scale, ordering of questions, etc.

The book by Payne (1951) still remains a classic in the art of asking questions. Afterwards, there have been some comprehensive reviews made by some authors on this topic (Krosnick and Presser 2009; and Schaeffer and Presser 2003, among others). Experiments in the wording of questions had started as early as 1940 with the paper by Cantril (1940). Next year, the author in an article showed that public opinion differs due to wording of questions (Rugg, 1941).

Krosnick and Presser (2009) have made a comprehensive review on how questions should be designed to get correct response. They are of the opinion that the questions should use simple and familiar words avoiding technical terms. Besides, the question should avoid words with ambiguous meanings, make response options mutually exclusive and exhaustive, and avoid leading or loaded questions that push respondents toward a particular answer. Krosnick and Presser have suggested that Likert scaling using 5 points should be taken by the investigators. In fact, the study of Lissitz and Green (1975) shows that 5-point scale is most effective. Later, Alwin and Krosnick (1991) and Alwin (1992) have found evidence that the reliability of individual rating scales appeared to increase as the number of categories grew, up to approximately seven or nine categories. Schaeffer and Presser (2003) have also explored literature on the issues of differences in responses due to changes in the reference periods and response categories.

Holleman et al. (2016) and Kamoen et al. (2011) have made some experiments which are similar to some of the questions taken up in the present study. Holleman et al. (2016) have found that positive and negative questions systematically affect the answers. So, there is some asymmetry in answers. The experimental evidence of Kamoen et al. (2011) has proven that respondents are more likely to disagree with negative questions than to agree with positive ones.

3. Collection and Analysis of Data

It may be noted that the presence of explanatory variables is not essential in this analysis. The aim of this paper is to see whether responses differ due to style of the questions. The original five-point responses have been grouped into dichotomous variable. The original responses are "strongly agree", "somewhat agree", "don't know or indifferent", "somewhat disagree", and "strongly disagree". The responses are transformed into dichotomous responses as "agree" (coded as 1) for both "strongly agree" and "somewhat agree"; and "disagree" (coded as 0) for both "somewhat disagree" and "strongly disagree". The answer "don't know" is not considered though some degrees of freedom are lost by not considering this answer in the analysis. There is another set of questions in which same set of statements with different styles giving opposite meaning (e.g. negative meaning) is put. For example, corresponding to the statement "women have the right to express their opinion if they disagree with their partner" (affirmative expression), the opposite statement is "women do not have the right to express their opinion if they disagree with their partner" (negative expression). The first set may be termed as positive set of questions and the second set may be termed as negative set of questions. "Positive" ("negative") set of questions does not necessarily give ethically/legally/socially accepted "positive" ("negative") views for the family or society. It is better to name them as Style A and Style B sets of questions.

In section 4, the answers of the respondents are summarised into 2×2 contingency table for each question. To see whether the answers differ due to the styles of the questions, Chi-square test is performed. The results show that the answers differ in most of the cases, i.e., the proportion of respondents saying 'agree' to positive question differs from the proportion of respondents saying 'disagree' to corresponding negative question. In section 5, a statistical model is built up with some plausible assumptions to find the actual proportion of persons who think the statement to be true. Section 6 tries to generalize the model. But the unknown parameters cannot be estimated uniquely. However, it gives some insight to the results of the earlier model. Finally, the formulation of questionnaires to achieve the correct proportion is discussed in the light of the results found from the first model in the paper.

4. Data Summary and the Results of Chi-squares Tests

Tables 1 through 6 give the frequency of persons who agreed and disagreed to the statement for both positive and negative styles. Ideally the proportion of persons who agreed to the positive style of question should be same as the proportion of persons who disagreed to the negative style of question. Hence, the frequencies for the negative question are shown in a reverse manner, i.e. frequency of "Disagree" is put first and frequency of "Agree" is put next, so that frequency of "Disagree" in the negative style of question appears in the same line as frequency of "Agree" in the positive style of question. The χ^2 test of the contingency table is then equivalent to the test of equality of proportions. The results show that the answers differ in most of the cases. Thus, the style of questions matters for responses.

Most of the questions in Tables 1 to 6 show significant differences of proportions of responses between positive and negative type questions. If the frequencies are inspected minutely, then one would observe that the pattern of responses is similar for almost all the statements, i.e., proportions of agreeing to positive questions were more than the proportions of disagreeing to the corresponding negative questions. Thus, by comparing the responses to the positive and negative questions it becomes clear that there is a tendency of agreeing to the statement even if the respondent does not believe it to be true. As already said, ideally the proportion of persons agreeing to the positive question should be same as the proportion of persons disagreeing to the negative question. Let us, for example, take responses to the statement number 1 in Table 3 of family related views. The proportion of persons who agree to the positive type question is 104 /106 = 0.981, whereas the proportion of persons who disagree to the negative question is 67 /102 = 0.657, which is much less than 0.981. One may naturally want to know the percentage of people who have agreed to the negative type statement, but thinks that it is not acceptable, i.e., believes that the statement is not true. One may also want to know the proportion of persons who actually believe the statement to be true. To find answers to these questions a model is built up, which is given in the next section.

5. The Model of Submissiveness and the Results

Responses to the statement may be symbolically written as given in Table 7. Thus,

N_{11} is the number of persons who agreed to the positive question (i.e., Type A question);

N_{21} is the number of persons who disagreed to the positive question (i.e., Type A question);

N_{12} is the number of persons who disagreed to the negative question (i.e., Type B question);

N_{22} is the number of persons who agreed to the negative question (i.e., Type B question).

N_{01}, N_{02}, N_{10} and N_{20} are marginal totals. Specifically,

$N_{11} + N_{21} = N_{01}$;
$N_{12} + N_{22} = N_{02}$;
$N_{11} + N_{12} = N_{10}$; and
$N_{21} + N_{22} = N_{20}$.

Lastly, N_{00} is the overall total frequency, i.e. total number of positive and negative questions. Also, $N_{00} = N_{01} + N_{02} = N_{10} + N_{20}$.

Table 1: Results of χ^2 Tests of Significance of Male Responses to Positive and Negative Statements: Family Related Views

Qn. No.	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
	Opinion	No. of persons	Opinion	No. of persons		
QF_1	Agree	49	Disagree	36	**0.000**	**Yes**
	Disagree	2	Agree	16		
	Total	51	Total	52		
QF_2	Agree	46	Disagree	37	**0.000**	**Yes**
	Disagree	6	Agree	13		
	Total	52	Total	50		
QF_3	Agree	29	Disagree	38	**0.002**	**Yes**
	Disagree	22	Agree	14		
	Total	51	Total	52		
QF_4	Agree	36	Disagree	27	**0.020**	**Yes**
	Disagree	15	Agree	25		
	Total	51	Total	52		
QF_5	Agree	29	Disagree	26	0.375	No
	Disagree	23	Agree	23		
	Total	52	Total	49		
QF_6	Agree	35	Disagree	26	0.060	Yes
	Disagree	17	Agree	25		
	Total	52	Total	51		
QF_7	Agree	39	Disagree	21	**0.027**	**Yes**
	Disagree	12	Agree	26		
	Total	51	Total	47		
QF_8	Agree	20	Disagree	22	0.109	Yes
	Disagree	29	Agree	29		
	Total	49	Total	51		
QF_9	Agree	13	Disagree	21	**0.001**	**Yes**
	Disagree	39	Agree	30		
	Total	52	Total	51		

Note: The details of QF_1 to QF_9 are given in Table A1 of Appendix A.

*Since "disagree" in positive question is equivalent to "agree" in negative question and *vice versa*, the numbers of cases "agree" and "disagree" in negative type of questions have been interchanged to conform with the positive style of question in the $\chi 2$ test.

Table 2: Results of χ^2 Tests of Significance of Female Responses to Positive and Negative Statements: Family Related Views

Qn. No.	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
	Opinion	No. of persons	Opinion	No. of persons		
QF_1	Agree	55	Disagree	31	**0.000**	**Yes**
	Disagree	0	Agree	19		
	Total	55	Total	50		
QF_2	Agree	49	Disagree	27	**0.000**	**Yes**
	Disagree	6	Agree	23		
	Total	55	Total	50		
QF_3	Agree	43	Disagree	28	**0.015**	**Yes**
	Disagree	12	Agree	22		
	Total	55	Total	50		
QF_4	Agree	28	Disagree	28	0.602	No
	Disagree	27	Agree	22		
	Total	55	Total	50		
QF_5	Agree	22	Disagree	19	0.834	No
	Disagree	33	Agree	31		
	Total	55	Total	50		
QF_6	Agree	33	Disagree	18	**0.018**	**Yes**
	Disagree	22	Agree	31		
	Total	55	Total	49		
QF_7	Agree	35	Disagree	16	**0.004**	**Yes**
	Disagree	18	Agree	28		
	Total	53	Total	44		
QF_8	Agree	18	Disagree	28	**0.028**	**Yes**
	Disagree	33	Agree	21		
	Total	51	Total	49		
QF_9	Agree	11	Disagree	19	**0.035**	**Yes**
	Disagree	44	Agree	30		
	Total	55	Total	49		

Note: *Since "disagree" in positive question is equivalent to "agree" in negative question and *vice versa*, the numbers of cases "agree" and "disagree" in negative type of questions have been interchanged to conform with the positive style of question in the χ^2 test.

Since there is a tendency of not going against a statement even if it may not be acceptable, it is assumed that if a person believes the statement to be true then the person "agrees to" the statement. But a portion of the people, who believe that the statement is false, will agree to the statement. Let this portion be α. Thus,

$$\alpha = P(\text{agreeing} \mid \text{unacceptable}) \quad \ldots (01)$$

Table 3: Results of χ^2 Tests of Significance of Responses of All Sampled Persons to Positive and Negative Statements: Family Related Views

Qn. No.	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
	Opinion	No. of persons	Opinion	No. of persons		
QF_1	Agree	104	Disagree	67	**0.000**	**Yes**
	Disagree	2	Agree	35		
	Total	106	Total	102		
QF_2	Agree	95	Disagree	64	**0.000**	**Yes**
	Disagree	12	Agree	36		
	Total	107	Total	100		
QF_3	Agree	72	Disagree	66	0.623	No
	Disagree	34	Agree	36		
	Total	106	Total	102		
QF_4	Agree	64	Disagree	55	0.347	No
	Disagree	42	Agree	47		
	Total	106	Total	102		
QF_5	Agree	51	Disagree	45	0.751	No
	Disagree	56	Agree	54		
	Total	107	Total	99		
QF_6	Agree	68	Disagree	44	**0.005**	**Yes**
	Disagree	39	Agree	56		
	Total	107	Total	100		
QF_7	Agree	74	Disagree	37	**0.000**	**Yes**
	Disagree	30	Agree	54		
	Total	104	Total	101		
QF_8	Agree	38	Disagree	50	0.087	Yes
	Disagree	62	Agree	50		
	Total	100	Total	100		
QF_9	Agree	24	Disagree	40	**0.006**	**Yes**
	Disagree	83	Agree	60		
	Total	107	Total	100		

Note: *Since "disagree" in positive question is equivalent to "agree" in negative question and *vice versa*, the numbers of cases "agree" and "disagree" in negative type of questions have been interchanged to conform with the positive style of question in the χ^2 test.

α may be regarded as the probability of a person, who does not believe the statement to be true, but agrees to the statement. In other words, it may be regarded as the **degree of submissiveness** of a person in regard to the statement. It is assumed that α is same for both positive and negative statements. The null hypothesis is $\alpha = 0$. The parameter α is estimated and then tested whether it is zero.

Let us now define A to be the event that the proposal is acceptable and $\mathcal{P}$ be the event that the statement is positive. A^c is the event that the statement is unacceptable. Similarly, $\mathcal{P}^c$ is the event that the statement is negative, i.e. the statement opposite to P. Thus,

$$P(A \mid \mathcal{P}) = P(A^c \mid \mathcal{P}^c), \qquad \ldots (02)$$

Table 4: Results of χ^2 Tests of Significance of Male Responses to Positive and Negative Statements: Social Views

Qn. No.	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
	Opinion	No. of persons	Opinion	No. of persons		
QS_1	Agree	42	Disagree	39	0.478	No
	Disagree	10	Agree	13		
	Total	52	Total	52		
QS_2	Agree	51	Disagree	35	**0.000**	**Yes**
	Disagree	1	Agree	17		
	Total	52	Total	52		
QS_3	Agree	45	Disagree	22	**0.000**	**Yes**
	Disagree	6	Agree	29		
	Total	51	Total	51		
QS_4	Agree	43	Disagree	24	**0.000**	**Yes**
	Disagree	8	Agree	23		
	Total	51	Total	47		
QS_5	Agree	33	Disagree	25	0.300	No
	Disagree	19	Agree	22		
	Total	52	Total	47		
QS_6	Agree	24	Disagree	18	0.153	No
	Disagree	22	Agree	30		
	Total	46	Total	48		
QS_7	Agree	43	Disagree	26	**0.001**	**Yes**
	Disagree	8	Agree	23		
	Total	51	Total	49		
QS_8	Agree	45	Disagree	34	**0.016**	**Yes**
	Disagree	5	Agree	14		
	Total	50	Total	48		
QS_9	Agree	45	Disagree	31	**0.002**	**Yes**
	Disagree	3	Agree	14		
	Total	48	Total	45		
QS_{10}	Agree	36	Disagree	29	0.163	No
	Disagree	12	Agree	18		
	Total	48	Total	47		
QS_{11}	Agree	45	Disagree	29	**0.001**	**Yes**
	Disagree	7	Agree	22		
	Total	52	Total	51		
QS_{12}	Agree	43	Disagree	31	**0.013**	**Yes**
	Disagree	9	Agree	20		
	Total	52	Total	51		
QS_{13}	Agree	29	Disagree	4	**0.000**	**Yes**
	Disagree	23	Agree	48		
	Total	52	Total	52		
QS_{14}	Agree	27	Disagree	25	0.617	No
	Disagree	23	Agree	26		
	Total	50	Total	51		

Note: *"Disagree" in positive question is equivalent to "agree" in negative question and *vice versa.*
The details of QS_1 to QS_{14} are given in Table A2 of Appendix A.

Table 5: Results of χ^2 Tests of Significance of Female Responses to Positive and Negative Statements: Social Views

	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
Qn. No.	Opinion	No. of persons	Opinion	No. of persons		
QS_1	Agree	51	Disagree	27	**0.000**	**Yes**
	Disagree	4	Agree	19		
	Total	55	Total	46		
QS_2	Agree	50	Disagree	22	**0.000**	**Yes**
	Disagree	5	Agree	28		
	Total	55	Total	50		
QS_3	Agree	42	Disagree	22	**0.000**	**Yes**
	Disagree	11	Agree	27		
	Total	53	Total	49		
QS_4	Agree	36	Disagree	31	0.430	No
	Disagree	14	Agree	17		
	Total	50	Total	48		
QS_5	Agree	26	Disagree	21	0.590	No
	Disagree	28	Agree	28		
	Total	54	Total	49		
QS_6	Agree	36	Disagree	19	**0.015**	**Yes**
	Disagree	18	Agree	26		
	Total	54	Total	45		
QS_7	Agree	48	Disagree	27	**0.000**	**Yes**
	Disagree	6	Agree	20		
	Total	54	Total	47		
QS_8	Agree	42	Disagree	29	**0.035**	**Yes**
	Disagree	10	Agree	18		
	Total	52	Total	47		
QS_9	Agree	35	Disagree	27	0.084	No
	Disagree	9	Agree	16		
	Total	44	Total	43		
QS_{10}	Agree	42	Disagree	26	**0.020**	**Yes**
	Disagree	10	Agree	18		
	Total	52	Total	44		
QS_{11}	Agree	45	Disagree	26	**0.000**	**Yes**
	Disagree	6	Agree	21		
	Total	51	Total	47		
QS_{12}	Agree	47	Disagree	28	**0.002**	**Yes**
	Disagree	8	Agree	20		
	Total	55	Total	48		
QS_{13}	Agree	34	Disagree	8	**0.000**	**Yes**
	Disagree	19	Agree	42		
	Total	53	Total	50		
QS_{14}	Agree	35	Disagree	23	0.105	No
	Disagree	17	Agree	22		
	Total	52	Total	45		

Note: *Since "disagree" in positive question is equivalent to "agree" in negative question and *vice versa*, the numbers of cases "agree" and "disagree" in negative type of questions have been interchanged to conform with the positive style of question in the χ^2 test.

Table 6: Results of χ^2 Tests of Significance of Responses of Sampled Persons to Positive and Negative Statements: Social Views

	Positive Statement		Negative Statement		P value of χ^2 test*	Whether Responses to Positive & Negative Questions Differ Significantly
Qn. No.	Opinion	No. of persons	Opinion	No. of persons		
QS_1	Agree	93	Disagree	66	**0.001**	**Yes**
	Disagree	14	Agree	32		
	Total	107	Total	98		
QS_2	Agree	101	Disagree	57	**0.000**	**Yes**
	Disagree	6	Agree	45		
	Total	107	Total	102		
QS_3	Agree	87	Disagree	44	**0.000**	**Yes**
	Disagree	17	Agree	56		
	Total	104	Total	100		
QS_4	Agree	79	Disagree	55	**0.002**	**Yes**
	Disagree	22	Agree	40		
	Total	101	Total	95		
QS_5	Agree	59	Disagree	46	0.271	No
	Disagree	47	Agree	50		
	Total	106	Total	96		
QS_6	Agree	60	Disagree	37	**0.005**	**Yes**
	Disagree	40	Agree	56		
	Total	100	Total	93		
QS_7	Agree	91	Disagree	53	**0.000**	**Yes**
	Disagree	14	Agree	43		
	Total	105	Total	96		
QS_8	Agree	87	Disagree	63	**0.002**	**Yes**
	Disagree	15	Agree	32		
	Total	102	Total	95		
QS_9	Agree	80	Disagree	58	**0.001**	**Yes**
	Disagree	12	Agree	30		
	Total	92	Total	88		
QS_{10}	Agree	78	Disagree	55	**0.008**	**Yes**
	Disagree	22	Agree	36		
	Total	100	Total	91		
QS_{11}	Agree	90	Disagree	55	**0.000**	**Yes**
	Disagree	13	Agree	43		
	Total	103	Total	98		
QS_{12}	Agree	90	Disagree	59	**0.000**	**Yes**
	Disagree	17	Agree	40		
	Total	107	Total	99		
QS_{13}	Agree	63	Disagree	12	**0.000**	**Yes**
	Disagree	42	Agree	90		
	Total	105	Total	102		
QS_{14}	Agree	62	Disagree	48	0.127	No
	Disagree	40	Agree	48		
	Total	102	Total	96		

Note: *Since "disagree" in positive question is equivalent to "agree" in negative question and *vice versa*, the numbers of cases "agree" and "disagree" in negative type of questions have been interchanged to conform with the positive style of question in the χ^2 test.

Table 7: Responses to Positive and Negative Styles of Statements of a Question

Positive Statement		Negative Statement		All
Opinion	**No. of persons**	**Opinion**	**No. of persons**	**No. of persons**
Agree	N_{11}	Disagree	N_{12}	N_{10}
Disagree	N_{21}	Agree	N_{22}	N_{20}
Total	N_{01}	Total	N_{02}	N_{00}

i.e., the probability that a positive statement is acceptable is same as the probability that a negative statement is unacceptable. Evidently, $P(A \mid \mathcal{P}) + P(A^c \mid \mathcal{P}) = 1$, since the answer "I do not know" has not been considered for the analysis. The aim is to find $P(A \mid \mathcal{P})$.

Suppose a (a^c) is the event that the person agrees to (disagrees with) the proposal. An estimate of $P(a \mid \mathcal{P})$, in this case, is N_{11} / N_{01}. Similarly, an estimate of $P(a^c \mid \mathcal{P}^c)$ is N_{12} / N_{02}.

Theorem 1: $\alpha = P(a \mid \mathcal{P}) + P(a \mid \mathcal{P}^c) - 1$; and $P(A \mid \mathcal{P}) = (P(a \mid \mathcal{P}) - \alpha) / (1-\alpha)$.

Proof: Since α is the proportion of persons (among persons who thinks that the statement is acceptable) who agrees, it can be symbolically write for the positive statement,

$$\alpha * P(A^c \mid \mathcal{P}) + P(A \mid \mathcal{P}) = P(a \mid \mathcal{P}). \quad \ldots (03)$$

Similarly, for the negative statement,

$$\alpha * P(A^c \mid \mathcal{P}^c) + P(A \mid \mathcal{P}^c) = P(a \mid \mathcal{P}^c). \quad \ldots (04)$$

From the equation (03) it follows that $\alpha * (1 - P(A \mid \mathcal{P})) + P(A \mid \mathcal{P}) = P(a \mid \mathcal{P})$.

$\Rightarrow (1-\alpha) * P(A \mid \mathcal{P}) + \alpha = P(a \mid \mathcal{P}) \Rightarrow P(A \mid \mathcal{P}) = (P(a \mid \mathcal{P}) - \alpha) / (1-\alpha)$.

From the equation (04)

$\alpha * P(A^c \mid \mathcal{P}^c) + (1 - P(A^c \mid \mathcal{P}^c)) = P(a \mid \mathcal{P}^c)$.

$\Rightarrow (1-\alpha) * P(A^c \mid \mathcal{P}^c) = P(a^c \mid \mathcal{P}^c) \Rightarrow (1-\alpha) * P(A^c \mid \mathcal{P}^c) = 1 - P(a \mid \mathcal{P}^c)$.

$\Rightarrow P(A^c \mid \mathcal{P}^c) = (1 - P(a \mid \mathcal{P}^c)) / (1-\alpha)$

Now the restriction $P(A \mid \mathcal{P}) = P(A^c \mid \mathcal{P}^c)$ is imposed to get

$(P(a \mid \mathcal{P}) - \alpha) / (1-\alpha) = (1 - P(a \mid \mathcal{P}^c)) / (1-\alpha) \Rightarrow (P(a \mid \mathcal{P}) - \alpha) = (1 - P(a \mid \mathcal{P}^c))$

$$\Rightarrow \alpha = P(a \mid \mathcal{P}) + P(a \mid \mathcal{P}^c) - 1. \quad \ldots (05)$$

It is now a routine work to get the $P(A \mid \mathcal{P})$ as

$$P(A \mid \mathcal{P}) = (P(a \mid \mathcal{P}) - \alpha) / (1-\alpha). \text{ Q.E.D.} \quad \ldots (06)$$

Note 1.1: This theorem gives us the way of estimating the actual proportion of persons who think that the (positive) statement is acceptable regardless of whether the persons agree or disagree to the statement provided the assumptions are true.

Note 1.2: Observe that, $P(A \mid \mathcal{P})$ is not same as $(P(a \mid \mathcal{P}) + P(a^c \mid \mathcal{P}^c)) / 2$ as the common sense suggests. This is because a non-symmetric assumption that "only those respondents who think that the statement is unacceptable can express it otherwise" is made. This is not because the total numbers of observations in both the positive and negative statements are not same. To get an insight to $P(A \mid \mathcal{P})$, let us further break it up to individual cell frequencies.

$$\begin{aligned} P(A \mid \mathcal{P}) &= (P(a \mid \mathcal{P}) - \alpha) / (1-\alpha) \\ &= (P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) + 1) / (1 - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) + 1) \\ &= (1 - P(a \mid \mathcal{P}^c)) / (2 - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c)) \\ &= P(a^c \mid \mathcal{P}^c) / (P(a^c \mid \mathcal{P}) + P(a^c \mid \mathcal{P}^c)) \quad \ldots (07) \end{aligned}$$

$= (N_{12} / N_{02}) / (N_{21} / N_{01} + N_{12} / N_{02})$ … (08)

= (Proportion of persons who disagree to the negative statement) / (Proportion of persons who disagree to the positive statement + Proportion of persons who disagree to the negative statement)

Note 1.3: If $\alpha > (<) 0$ then $P(A \mid \mathcal{P}) \leq (\geq) (P(a \mid \mathcal{P})$, equality holding only if $P(A \mid \mathcal{P}) = 1$. If $\alpha = 0$, then $P(A \mid \mathcal{P}) = (P(a \mid \mathcal{P})$. This follows from the fact that $P(A \mid \mathcal{P}) = (P(a \mid \mathcal{P}) - \alpha) / (1-\alpha)$ can be rewritten as

$$P(A \mid \mathcal{P}) - (P(a \mid \mathcal{P}) = \alpha(P(A \mid \mathcal{P}) - 1). \quad \ldots (09)$$

The values of α and $P(A \mid \mathcal{P})$ are calculated for all the statements taking all the respondents views and also separately for men and women. The results are given in the Tables 8 and 9.

It is clear that there are variations in the values of α and $P(A \mid \mathcal{P})$ in both family-related and social statements. The values of α ranged from -21.8% to 38.0% in family-related views, whereas it ranged from 5.0% to 48.2% in social views. Also observe that, in social views, there were no negative values of α. This justifies the assumption that the respondents, who think that the statement is true, will fully agree to the statement. In family related views, this assumption fails only in a few cases. In Tables 8 and 9, the percentage changes from $P(a \mid P)$ to $P(A \mid P)$ are shown. In many of the cases these percentage changes were more than 10%. Statements 8 and 9 of family related issues deserve special mention. These two statements are "Women should be submissive" and "Men are superior to women in all respects of life".

Proportions of persons, who believe that the statements are true, are more than the corresponding proportions of persons who agree to these statements, and this is more pronounced among women. In each of the social issues, the actual proportion of believers is less than the proportion of assenters.

Gender differences in the probability of deviation from truth and probability of believing the statement to be true for nine family related statements (Table 10) and for fourteen social statements (Table 11) are seen. Percentage changes are too high in some statements, which may be because either the values of α were too small for males or females or the sign differed. The differences are shown in the form of percentage in separate columns. These differences are also too high for many of the statements for both family and social views. Thus, men and women differed much in their expressions through deviations from truth or in their actual beliefs.

6. The Model of Submissiveness and Assertiveness Taken Together

To get a better view of the data, let us assume that, in addition to α, there is β, the proportion of persons, among persons who think that the statement is acceptable, disagree with the statement. Thus, β can be viewed as the probability of assertiveness. It is also assumed that both α and β are same for both positive and negative statements. Thus,

$$\left.\begin{array}{l}\alpha = P(\text{agreeing} \mid \text{unacceptable}): \text{same for both positive and negative questions, and} \\ \beta = P(\text{disagreeing} \mid \text{acceptable}): \text{same for both positive and negative questions.}\end{array}\right\} \quad \ldots(10)$$

Given the assumptions, both the parameters α and β are inestimable from the contingency table. This is because three parameters, namely, $P(A \mid \mathcal{P})$, α and β, are to be estimated from two equations. The two equations are as follows.

$$\alpha * P(A^c \mid \mathcal{P}) + (1 - \beta)\, P(A \mid \mathcal{P}) = P(a \mid \mathcal{P}) \quad \ldots (11)$$

$$\alpha * P(A^c \mid \mathcal{P}^c) + (1 - \beta)\, P(A \mid \mathcal{P}^c) = P(a \mid \mathcal{P}^c)^2 \quad \text{... (12)}$$

$P(A^c \mid \mathcal{P})$, $P(A^c \mid \mathcal{P}^c)$ and $P(A \mid \mathcal{P}^c)$ are known once $P(A \mid \mathcal{P})$ is known, since $P(A^c \mid \mathcal{P}) = 1 - P(A \mid \mathcal{P})$, $P(A^c \mid \mathcal{P}^c) = P(A \mid \mathcal{P})$, $P(A \mid \mathcal{P}^c) = 1 - P(A^c \mid \mathcal{P}^c) = 1 - P(A \mid \mathcal{P})$.

Thus,
$$\alpha * (1 - P(A \mid \mathcal{P}) + (1 - \beta) * P(A \mid \mathcal{P}) = P(a \mid \mathcal{P}) \quad \text{... (13)}$$
$$\alpha * P(A \mid \mathcal{P}) + (1 - \beta) *(1 - P(A \mid \mathcal{P}) = P(a \mid \mathcal{P}^c) \quad \text{... (14)}$$

Table 8: Probability of Believing the Statement to be True and Percentage Deviation from Truth for Nine Family Related Statements Separately for Male, Female and All People

Qn. No.	Male				Female				All			
	α	P(A | P)	P(a | P)	% Change	α	P(A | P)	P(a | P)	% Change	α	P(A | P)	P(a | P)	% Change
QF_1	0.268	0.946	0.961	-1.6	0.380	1.000	1.000	0.0	0.324	0.972	0.981	-0.9
QF_2	0.145	0.865	0.885	-2.3	0.351	0.832	0.891	-6.6	0.248	0.851	0.888	-4.2
QF_3	-0.162	**0.629**	**0.569**	**10.5**	0.222	0.720	0.782	-7.9	0.032	0.669	0.679	**-1.5**
QF_4	0.187	0.638	0.706	-9.6	-0.051	0.533	0.509	4.7	0.065	0.576	0.604	-4.6
QF_5	0.027	0.545	0.558	-2.3	0.020	0.388	0.400	-3.0	0.022	0.465	0.477	-2.5
QF_6	0.163	0.609	0.673	-9.5	0.233	**0.479**	**0.600**	**-20.2**	0.196	**0.547**	**0.636**	**-14.0**
QF_7	0.318	**0.655**	**0.765**	**-14.4**	0.297	**0.517**	**0.660**	**-21.7**	0.246	**0.559**	**0.712**	**-21.5**
QF_8	-0.023	0.422	0.408	3.4	-0.218	**0.469**	**0.353**	**32.9**	-0.120	**0.446**	**0.380**	**17.4**
QF_9	-0.162	**0.354**	**0.250**	**41.6**	-0.188	**0.326**	**0.200**	**63.0**	-0.176	**0.340**	**0.224**	**51.8**

Table 9: Probability of Believing the Statement to be True and Percentage Deviation from Truth for Fourteen Social Statements Separately for Male, Female and All People

Qn. No.	Male				Female				All			
	α	P(A | P)	P(a | P)	% Change	α	P(A | P)	P(a | P)	% Change	α	P(A | P)	P(a | P)	% Change
QS_1	0.058	0.796	0.808	-1.5	0.340	0.890	0.927	-4.0	0.196	0.837	0.869	-3.7
QS_2	0.308	0.972	0.981	-0.9	0.469	0.829	0.909	-8.8	0.385	0.909	0.944	-3.7
QS_3	0.451	**0.786**	**0.882**	**-10.9**	0.343	**0.684**	**0.792**	**-13.6**	0.397	**0.729**	**0.837**	**-12.9**
QS_4	0.332	0.765	0.843	-9.3	0.074	0.698	0.720	-3.1	0.203	0.727	0.782	-7.0
QS_5	0.103	0.593	0.635	-6.6	0.053	0.453	0.481	-5.8	0.077	0.519	0.557	-6.8
QS_6	0.147	**0.439**	**0.522**	**-15.9**	0.244	**0.559**	**0.667**	**-16.2**	0.202	**0.499**	**0.600**	**-16.8**
QS_7	0.313	0.772	0.843	-8.4	0.314	0.838	0.889	-5.7	0.315	0.805	0.867	-7.2
QS_8	0.192	0.876	0.900	-2.7	0.191	0.762	0.808	-5.7	0.190	0.818	0.853	-4.1
QS_9	0.249	0.917	0.938	-2.2	0.168	0.754	0.795	-5.2	0.210	0.835	0.870	-4.0
QS_{10}	0.133	0.712	0.750	-5.1	0.217	0.754	0.808	-6.7	0.176	0.733	0.780	-6.0
QS_{11}	0.297	0.809	0.865	-6.5	0.329	0.825	0.882	-6.5	0.313	0.816	0.874	-6.6
QS_{12}	0.219	0.778	0.827	-5.9	0.271	0.800	0.855	-6.4	0.245	0.790	0.841	-6.1
QS_{13}	0.481	**0.148**	**0.558**	**-73.5**	0.482	**0.309**	**0.642**	**-51.9**	0.482	**0.227**	**0.600**	**-62.2**
QS_{14}	0.050	0.516	0.540	-4.4	0.162	0.610	0.673	-9.4	0.108	0.560	0.608	-7.9

[2]Observe that the equations $\beta * P(A \mid \mathcal{P}) + (1 - \alpha)\, P(A^c \mid \mathcal{P}) = P(a^c \mid \mathcal{P})$ and $\beta * P(A \mid \mathcal{P}^c) + (1 - \alpha)\, P(A^c \mid \mathcal{P}^c) = P(a^c \mid \mathcal{P}^c)$, are just complementary to equations (1) and (2), respectively. Thus these two equations are redundant.

Table 10: Gender Differences in the Probability of Deviation from Truth and Probability of Believing the Statement to be True for Nine Family Related Statements

Qn. No.	Male α	Female α	Percentage Change	Difference (%)	Male P(A | P)	Female P(A | P)	Percentage Change	Difference (%)
QF_1	0.268	0.380	-29.5	11.2	0.946	1.000	-5.4	5.4
QF_2	0.145	0.351	-58.7	20.6	0.865	0.832	4.0	-3.3
QF_3	-0.162	0.222	-173.0	38.4	0.629	0.720	-12.6	9.1
QF_4	0.187	-0.051	-466.7	-23.8	0.638	0.533	19.7	-10.5
QF_5	0.027	0.020	35.0	-0.7	0.545	0.388	40.5	-15.7
QF_6	0.163	0.233	-30.0	7.0	0.609	0.479	27.1	-13.0
QF_7	0.318	0.297	7.1	-2.1	0.655	0.517	26.7	-13.8
QF_8	-0.023	-0.218	-89.4	-19.5	0.422	0.469	-10.0	4.7
QF_9	-0.162	-0.188	-13.8	-2.6	0.354	0.326	8.6	-2.8

Table 11: Gender Differences in the Probability of Deviation from Truth and Probability of Believing the Statement to be True for Fourteen Social Statements

Qn. No.	Male α	Female α	Percentage Change	Difference (%)	Male P(A | P)	Female P(A | P)	Percentage Change	Difference (%)
QS_1	0.058	0.340	-82.9	-28.2	0.796	0.890	-10.6	-9.4
QS_2	0.308	0.469	-34.3	-16.1	0.972	0.829	17.2	14.3
QS_3	0.451	0.343	31.5	10.8	0.786	0.684	14.9	10.2
QS_4	0.332	0.074	348.6	25.8	0.765	0.698	9.6	6.7
QS_5	0.103	0.053	94.3	5.0	0.593	0.453	30.9	14.0
QS_6	0.147	0.244	-39.8	-9.7	0.439	0.559	-21.5	-12.0
QS_7	0.313	0.314	-0.3	-0.1	0.772	0.838	-7.9	-6.6
QS_8	0.192	0.191	0.5	0.1	0.876	0.762	15.0	11.4
QS_9	0.249	0.168	48.2	8.1	0.917	0.754	21.6	16.3
QS_{10}	0.133	0.217	-38.7	-8.4	0.712	0.754	-5.6	-4.2
QS_{11}	0.297	0.329	-9.7	-3.2	0.809	0.825	-1.9	-1.6
QS_{12}	0.219	0.271	-19.2	-5.2	0.778	0.800	-2.8	-2.2
QS_{13}	0.481	0.482	-0.2	-0.1	0.148	0.309	-52.1	-16.1
QS_{14}	0.050	0.162	-69.1	-11.2	0.516	0.610	-15.4	-9.4

Adding equations (13) and (14), one gets

$$\alpha - \beta = P(a \mid \mathcal{P}) + P(a \mid \mathcal{P}^c) - 1. \quad \text{... (15)}$$

The earlier model is a special case of this model, i.e., when $\beta = 0$.

If $P(A \mid \mathcal{P})$ is known, then $(1 - P(A \mid \mathcal{P})) / P(A \mid \mathcal{P})$ (= D, say) is also known. Dividing both (13) and (14) by $P(A \mid \mathcal{P})$,

$$\alpha * D + (1 - \beta) = P(a \mid \mathcal{P}) / P(A \mid \mathcal{P}) \quad \text{... (16)}$$

$$\alpha + (1 - \beta) * D = P(a \mid \mathcal{P}^c) / P(A \mid \mathcal{P}) \quad \text{... (17)}$$

Eqn. (13) – Eqn. (14) is

$$\alpha * (D - 1) + (1 - \beta) * (1 - D) = P(a \mid \mathcal{P}) / P(A \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) / P(A \mid \mathcal{P})$$

or, $$(1 - \alpha - \beta) * (1 - D) = P(a \mid \mathcal{P}) / P(A \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) / P(A \mid \mathcal{P})$$

or, $$(1 - \alpha - \beta) = [P(a \mid \mathcal{P}) / P(A \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) / P(A \mid \mathcal{P})] / (1 - D)$$

or, $(1 - \alpha - \beta) = [P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c)] \; / \; [2 * P(A \mid \mathcal{P}) - 1]$ ... (18)

The parameters α and β can be estimated by solving equations (15) and (18). On the other hand, if any one of α and β is known then one can get the other one from equation (15) and then get $P(A \mid \mathcal{P})$ from equation (18) as

$$P(A \mid \mathcal{P}) = [(P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c)) + (1 - \alpha - \beta)] \; / [2*(1 - \alpha - \beta)] \quad \text{... (19)}$$

$$= \tfrac{1}{2}[(P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c)) \; / (1 - \alpha - \beta) + 1] \quad \text{... (20)}$$

In any case one should have one more equation to know the complete picture. Thus, this model is not identified.

The above model can also be reformulated as a binomial distribution model, but the identification problem remains the same (see Appendix B for details).

7. Discussions

This paper depicts a methodology of finding out the correct proportions of persons who accepts a statement by carrying out an appropriate opinion survey. In this paper, the word "accepting" is taken synonymous to "thinking it to be true", whereas the word "agreeing" is taken synonymous to "expressing that the person agrees it to be true". Accepting (or not accepting) a statement and agreeing (or disagreeing) to the statement are thus different. In this paper, it is assumed that if a person accepts a statement then the person also agrees to the statement, but if a person does not accept the statement then the person may or may not disagree to the statement.

As an illustration, nine statements on family related issues and fourteen statements on social issues on gender violence are canvassed. The methodology described in this paper is not applicable to gender violence only, but also to other phenomena of the society requiring the views of the members in the society. Thus, it has a much broader spectrum of applications than one can initially think of.

8. Conclusion and Policy Implication

The study depicts a clear picture about the impact of style of questions on an individual. Besides, one would be able to get the actual belief of a group of individuals in a summary form by canvasing questions of both the styles. The questions dealt with in this paper are related to gender violence. While researching a sensitive issue such as gender violence, an extra-ordinary care is to be taken about methodology. This may help one to locate the root of the problems of gender violence and thus enable to prescribe appropriate administrative as well as social reform policies to eradicate this evil crime from the society.

The implication of this result is that one should put positive as well as negative questions on the same issue, but there should be two sets of questionnaires, as the same question with two different styles cannot be asked to the *same person*. Of course, this does not mean that one set contains only positive questions, and the other set contains only negative questions. Let us make it clearer. One should prepare two sets of questionnaires. The first set should have some positive and some negative styles of questions. The other set should contain the same set of questions with styles changed, i.e., opposite to that of the first style.

This study concludes with the quotations from Schaeffer and Presser (2003): "Although asking questions will always involve an element of art, future research is likely to provide guidance for decisions about many other features of wording. The resulting improvements in survey measurement should facilitate progress in all areas of social science that make use of questionnaires."

Acknowledgement

This paper is an outcome of the project entitled "Gender Violence in India: Its Roots, Nature and Extent" funded by the Indian Council of Social Science Research (ICSSR), New Delhi.

About the authors

Manoranjan Pal is a professor at the Economic Research Unit, Indian Statistical Institute, Kolkata, India. **Chaiti Sharma-Biswas** is an associate scientist at the Economic Research Unit, Indian Statistical Institute, Kolkata, India. **Sriparna Banerjee** is a Ph.D. student at the Anthropology Department, West Bengal State University, Barasat, India. **Anjali Ghosh** is a professor at the Psychology Research Unit, Indian Statistical Institute, Kolkata, India. **Subhendu Chakrabarti** has retired as an associate scientist from the Economic Research Unit, Indian Statistical Institute, Kolkata, India. **Sumana Guha** is an assistant professor at the Department of Commerce (Morning), St. Xavier's College, Kolkata, India. **Premananda Bharati** is a professor at the Biological Anthropology Unit, Indian Statistical Institute, Kolkata, India.

References

Agresti, Alan (2007), *An Introduction to Categorical Data Analysis*, New York: Wiley-Interscience.

Alwin D.F., and J.A. Krosnick (1991), "The reliability of survey attitude measurement: The influence of question and respondent attributes", *Sociological Methods & Research*, 20(1): 139–181.

Alwin, D.F. (1992), "Information transmission in the survey interview: Number of response categories and the reliability of attitude measurement". In *Sociological Methodology*, P.V. Marsden (ed.), Vol. 22, pp. 83–118.

Cantril, H. (1940), "Experiments in the Wording of Questions", *Public Opinion Quarterly*, 4(2): 330-332.

Groves, R.M. (1989), *Survey errors and survey costs*, New York: John Wiley.

Groves, R.M., F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau (2004), *Survey methodology*, New York: John Wiley.

Hansen, Morris H., William N. Hurwitz, Eli S. Marks, and W. Parker Mauldin (1951), "Response Errors in Surveys", *Journal of the American Statistical Association,* 46(254): 147-190.

Hilbe, Joseph M. (2009), *Logistic Regression Models*, London: Chapman and Hall/CRC.

Holleman, B., N. Kamoen, A. Krouwel, J. Van de Pol, and Claes de Vreese (2016), "Positive vs. Negative: The Impact of Question Polarity in Voting Advice Applications", *PLoS ONE*, Open Access, pp. 1-17.

Hosmer, David, and Stanley Lemeshow (2013), *Applied Logistic Regression* (Third Edition), New York: Wiley.

Kalton, G., and H. Schuman (1982), "The Effect of the Question on Survey Responses: A Review", *Journal of the Royal Statistical Society*, Series A, 145 (Part 1): 42-73.

Kamoen, N., B. Holleman, P. Mak, T. Sanders, and H. van den Bergh (2011), "Agree or Disagree? Cognitive Processes in Answering Contrastive Survey Questions", *Discourse Processes*, 48(5): 355–385.

Krosnick, Jon A., and S. Presser (2009), "Question and Questionnaire Design". In *Handbook of Survey Research* (2nd Edition), James D. Wright and Peter V. Marsden (eds.), San Diego: Elsevier.

Lessler, J.T., and W.D. Kalsbeek (1992), *Non-sampling error in surveys*, New York: John Wiley.

Lissitz, R.W., and S.B. Green (1975), "Effect of the number of scale points on reliability: A Monte Carlo approach", *Journal of Applied Psychology*, 60(1): 10-13.

Lyberg, L., and D. Kasprzyk (1991), "Data collection methods and measurement error: An Overview". In *Measurement Errors in Surveys*, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, S. Sudman, (eds.), pp. 237-257, New York: Wiley and Sons.

Menard, Scott (2002), *Applied Logistic Regression Analysis*, California: SAGE.

OECD (2019), "Gender, Institutions and Development", available at: https://data.oecd.org/inequality/violence-against-women.htm.

Payne, S.L. (1951), *The Art of Asking Questions*, Princeton: Princeton University Press.

Rugg, Donald (1941), "Experiments in Wording Questions: II", *Public Opinion Quarterly*, 5(1): 91–92.

Schaeffer, N.C., and S. Presser (2003), "The Science of Asking Questions", *Annual Review of Sociology*, 29(1): 65–88.

United Nations (2020), *The World's Women 2020: Trends and Statistics*, New York: United Nations Statistics Division.

UNODC (2014), *Global Report on Trafficking in Persons*, Vienna: United Nations Office on Drugs and Crime (UNODC).

WHO (2013), *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*, Geneva: World Health Organisation (WHO) [in association with London School of Hygiene and Tropical Medicine and South African Medical Research Council].

Wooldridge, Jeffrey (2010), *Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.

Appendix A

Table A1: Family-Related Type A and Type B Statements

Statements (Type A)	Statements (Type B)
QF_1: Women have the right to express their opinion if they disagree with their partner.	QF_1: Women do not have the right to express their opinion even if they disagree with their partner.
QF_2: Couples should share equally in household chores if they are both working outside the home.	QF_2: Couples should not share equally in household chores if they are both working outside the home.
QF_3: Women can work outside the home even if the family does not need the money.	QF_3: Women should not work outside the home even if the family needs the money.
QF_4: Physical violence between couples is a private matter and should be handled within the family.	QF_4: Physical violence between couples is not a private matter and should be handled with external intervention, if needed.
QF_5: It is the duty of women to obey their partner always.	QF_5: It is not the duty of women to obey their partner always.
QF_6: If women want to see their relatives or friends, they should take permission from their partner.	QF_6: If women want to see their relatives or friends, it is not necessary to take permission from their partner.
QF_7: Violence against women happens mostly in the families with low incomes.	QF_7: Violence against women happens rarely in the families with low incomes.
QF_8: Women should be submissive.	QF_8: Women should not be submissive.
QF_9: Men are superior to women in all respects of life	QF_9: Men are not superior to women in all respects of life

Table A2: Type A and Type B Statements Related to Social Views on Gender Violence

Statements (Type A)	Statements (Type B)
QS_1: Abuses to women have increased significantly now-a-days.	QS_1: Women abuses have not increased significantly now-a-days.
QS_2: Women have equal rights as men.	QS_2: Women do not have equal rights as men.
QS_3: Men and women are equally responsible for violence against women.	QS_3: Men and women are not equally responsible for violence against women.
QS_4: Violence against women occurs due to belief in the traditionally defined roles of women and men.	QS_4: Violence against women occurs not due to belief in the traditionally defined roles of women and men.
QS_5: Abuse against women has increased due to greater freedom in their mobility.	QS_5: Abuse against women has increased not due to greater freedom in their mobility.
QS_6: Violence happens mostly against those women who have physical and/or mental disabilities.	QS_6: Not necessarily violence happens mostly against those women who have physical and/or mental disabilities
QS_7: Too much alcohol consumption by men results in increased violence against women.	QS_7: Too much alcohol consumption by men does not result in increased violence against women.
QS_8: Recent increase in the violence against women is due to the decline in the moral values in our society.	QS_8: Recent increase in the violence against women is not due to the decline in the moral values in our society.
QS_9: Gap between aspirations and means among "Prospect-less Young Men" is responsible for increasing incidence of women abuse.	QS_9: Gap between aspirations and means among "Prospect-less Young Men" is not necessarily responsible for increasing incidence of women abuse.
QS_{10}: Gender violence is the outcome of the natural attitude of human beings to hold dominating power.	QS_{10}: Gender violence is not the outcome of the natural attitude of human beings to hold dominating power.

Statements (Type A)	Statements (Type B)
QS_{11}: Lack of proper enforcement of existing policies /laws to protect women is the reason for increasing abuse on them.	QS_{11}: Lack of proper enforcement of existing policies /laws to protect women is not the reason for increasing abuse on them.
QS_{12}: Insufficient infrastructure (i.e. inadequate street light, absence of ladies toilet, insecured public transport, etc.) is one of the main reasons for violence against women.	QS_{12}: Insufficient infrastructure (i.e. inadequate street light, absence of ladies toilet, insecured public transport, etc.) is not one of the main reasons for violence against women.
QS_{13}: Protest by women leads to increase in the degree of violence against them outside home.	QS_{13}: Protest by women leads to decrease in the degree of violence against them outside home.
QS_{14}: Protest by women leads to increase in the degree of violence against them at home.	QS_{14}: Protest by women leads to decrease in the degree of violence against them at home.

Appendix B

The Binomial Distribution of the Submissive and Assertive Model

Suppose X_i is a binary random variable taking values as

$X_i = 1$ if the statement is acceptable

$= 0$ if the statement is not acceptable

It is assumed that

$$P(X_i = 1) = p$$

and $$P(X_i = 0) = 1 - p$$

x_i is another random variable which takes values as

$x_i = 1$ if the ith respondent agrees to the statement

$= 0$ if the ith respondent does not agree to the statement

$$P(x_i = 1) = P(x_i = 1 \,\&\, X_i = 1) + P(x_i = 1 \,\&\, X_i = 0)$$
$$= P(x_i = 1 \mid X_i = 1)\, P(X_i = 1) + P(x_i = 1 \mid X_i = 0)\, P(X_i = 0)$$
$$= (1 - \beta)\, p + \alpha\, (1 - p)$$
$$P(x_i = 0) = P(x_i = 0 \,\&\, X_i = 1) + P(x_i = 0 \,\&\, X_i = 0)$$
$$= P(x_i = 0 \mid X_i = 1)\, P(X_i = 1) + P(x_i = 0 \mid X_i = 0)\, P(X_i = 0)$$
$$= \beta\, p + (1 - \alpha)\, (1 - p)$$

The bivariate distribution of (x_i, X_i) can be found and hence the conditional distribution of X given x values. The bivariate distribution of (x, X) is as follows:

X \ x	x = 0	x = 1	Subtotal
X = 0	$(1-\alpha)(1-p)$	$\alpha(1-p)$	1-p
X = 1	βp	$(1-\beta)p$	P
Subtotal	$\beta p+(1-\alpha)(1-p)$	$(1-\beta)p +\alpha(1-p)$	1

This can be written in reverse direction as

$P(X_i=0 \mid x_i=0) = \frac{(1-\alpha)(1-p)}{\beta p+(1-\alpha)(1-p)}$ and $P(X_i=1 \mid x_i=0) = \frac{\beta p}{\beta p+(1-\alpha)(1-p)}$. Also,

$P(X_i=0 \mid x_i=1) = \frac{\alpha(1-p)}{(1-\beta)p+\alpha(1-p)}$ and $P(X_i=1 \mid x_i=1) = \frac{(1-\beta)p}{(1-\beta)p+\alpha(1-p)}$.

Suppose $x = \Sigma x_i$. Since x_i's are iid, one has $x \sim Bin(N, (1-\beta)p + \alpha(1-p))$. For estimation of the parameters the following can be used.

$\bar{x} = N[(1-\beta)p + \alpha(1-p)]$ and $\overline{x^2} - \bar{x}^2 = N[(1-\beta)p + \alpha(1-p)][\beta p + (1-\alpha)(1-p)]$.

Suppose the corresponding random variables for negative statement are Y_i and y_i, i.e.

Y_i is a binary random variable taking values as

$Y_i = 1$ if the statement is acceptable

$= 0$ if the statement is not acceptable

It is assumed that

$P(Y_i=1) = 1-p$

and $P(Y_i=0) = p$.

[Remember that $Y_i = 1 \equiv X_i = 0$ and $Y_i = 0 \equiv X_i = 1$.]

y_i is another random variable which takes values as

$y_i = 1$ if the ith respondent agrees to the statement

$= 0$ if the ith respondent does not agree to the statement

Thus,

$P(y_i=1 \mid Y_i=1) = 1-\beta$, $P(y_i=0 \mid Y_i=1) = \beta$, $P(y_i=1 \mid Y_i=0) = \alpha$, $P(y_i=0 \mid Y_i=0) = (1-\alpha)$

Similarly,

$P(y_i=1) = P(y_i=1 \& Y_i=1) + P(y_i=1 \& Y_i=0)$

$= P(y_i=1 \mid Y_i=1)\,P(Y_i=1) + P(y_i=1 \mid Y_i=0)\,P(Y_i=0)$

$= (1-\beta)(1-p) + \alpha p$,

Since $P(X_i=1) = P(Y_i=0) = 1 - P(Y_i=1)$ and Since $P(X_i=0) = P(Y_i=1) = 1 - P(Y_i=0)$

$P(y_i=0) = P(y_i=0 \& Y_i=1) + P(y_i=0 \& Y_i=0)$

$= P(y_i=0 \mid Y_i=1)\,P(Y_i=1) + P(y_i=0 \mid Y_i=0)\,P(Y_i=0)$

$= \beta(1-p) + (1-\alpha)p$

Thus, the bivariate distribution of (y,Y) is as follows:

Y \ y	y = 0	y = 1	Subtotal
Y = 0	$(1-\alpha)p$	αp	p
Y = 1	$\beta(1-p)$	$(1-\beta)(1-p)$	1-p
Subtotal	$\beta(1-p)+(1-\alpha)p$	$(1-\beta)(1-p)+\alpha p$	1

It can be written in a reverse direction as

$P(Y_i = 0 \mid y_i = 0) = \dfrac{(1-\alpha)p}{\beta(1-p)(1-\alpha)p}$ and $P(Y_i = 1 \mid y_i = 0) = \dfrac{\beta(1-p)}{\beta(1-p)+(1-\alpha)p}$. Also,

$P(Y_i = 0 \mid y_i = 1) = \dfrac{\alpha p}{(1-\beta)(1-p)+\alpha p}$ and $P(Y_i = 1 \mid y_i = 1) = \dfrac{(1-\beta)(1-p)}{(1-\beta)(1-p)+\alpha p}$.

Suppose $y = \Sigma y_i$. Since y_i's are iid, it follows that $y \sim \text{Bin}(N, (1-\beta)\, p + \alpha\,(1-p))$. For estimation of the parameters the following equations can be used.

$\bar{y} = N\,[(1-\beta)\,(1-p) + \alpha\, p]$ and $\overline{y^2} - \bar{y}^2 = N[(1-\beta)\,(1-p) + \alpha\, p][\,\beta\,(1-p) + (1-\alpha)\, p]$.

$\bar{x} = N\,[(1-\beta)\, p + \alpha\,(1-p)]$ and $\overline{x^2} - \bar{x}^2 = N[(1-\beta)\, p + \alpha\,(1-p)][\,\beta\, p + (1-\alpha)\,(1-p)]$.

But it cannot be used because there is one observation for each. If it is assumed that the parameters have same values for each question, then it may be possible to arrive at some solution. But there is no way to assume it; because one can only estimate $(\beta - \alpha)$ and the values of $(\beta - \alpha)$ differ from question to question.