Athens Institute for Education and Research
ATINER

ATINER's Conference Paper Series
ECO2018-2699

Style of Question Matters: An Experiment with Questions on Gender Violence

Manoranjan Pal
Professor
Indian Statistical Institute
India

C. S. Biswas
Indian Statistical Institute
India

S. Bannerjee
Indian Statistical Institute
India

A. Ghosh
Indian Statistical Institute
India

S. Chakraborti
Indian Statistical Institute
India

S. Guha
Indian Statistical Institute
India

P. Bharati
Indian Statistical Institute
India

An Introduction to ATINER's Conference Paper Series

Conference papers are research/policy papers written and presented by academics at one of ATINER's academic events. ATINER's association started to publish this conference paper series in 2012. All published conference papers go through an initial peer review aiming at disseminating and improving the ideas expressed in each work. Authors welcome comments

Dr. Gregory T. Papanikos
President
Athens Institute for Education and Research

This paper should be cited as follows:

Pal, M., Biswas, C. S., Bannerjee, S., Ghosh, A., Chakraborti, S., Guha, S., Bharati, P. (2018). **"Style of Question Matters: An Experiment with Questions on Gender Violence",** Athens: ATINER'S Conference Paper Series, No: **ECO2018-2699.**

Athens Institute for Education and Research
8 Valaoritou Street, Kolonaki, 10671 Athens, Greece
Tel: + 30 210 3634210 Fax: + 30 210 3634209 Email: info@atiner.gr URL: www.atiner.gr
URL Conference Papers Series: www.atiner.gr/papers.htm
Printed in Athens, Greece by the Athens Institute for Education and Research.
ISSN: 2241-2891
05/02/2020

Style of Question Matters: An Experiment with Questions on Gender Violence

Manoranjan Pal
C. S. Biswas
S. Bannerjee
A. Ghosh
S. Chakraborti
S. Guha
P. Bharati

Abstract

To see whether style of questions matters, we have, in this present paper carried out an experiment, as a part of a large project on gender violence, which is being conducted in Meghalaya and West Bengal in India with two sets of questions on gender violence – Positive and Negative. The positive and negative questions were canvassed to separate sets of adult people randomly drawn from the selected regions of the two States. A clear picture emerges from the analysis of data using a statistical model. The respondents hesitate to disagree with a statement even if it is not acceptable to them. Out of 23 questions we got 18 cases with significant difference between the number of persons who think that the statement is acceptable and the number of persons who agree to the statement. The designing of questionnaires and conducting the survey, thus, call for a thorough revision. Instead of a single set of questionnaires we 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.

Keywords: Response error in survey, Style of question, Gender violence, Chi square test, India.

Introduction

It is well known that the response to a question depends on how it is posed. The approach of the investigator to the respondent is crucial. The background, for example, presence of other persons during the interview also matters. It also depends on many other factors like the availability of time of the respondent, religious and social dictums, etc. But it is less known that the statement of the questions like “should” vs. “should not”, “is” vs. “is not”, etc. matter in a significant manner. We had an opportunity to verify it. We set nine questions to get family related views and fourteen questions to get social views on different aspects of gender violence from adult males and females. The questions were in the form of statements. The respondents were asked to choose any one of the five alternatives as “strongly agree”, “somewhat agree”, “somewhat disagree”, “strongly disagree”, “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, Wooldridge 2010) provided enough explanatory variables are available. However, in our case, we do not have any explanatory variable other than sex of the respondent. Thus, this situation calls for innovative approach.

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 in the results (Kalton and Schuman 1982). People were aware of it as early as in fifties (Hansen et al. 1951), but no satisfactory solutions evolved so far. Besides non-response, there may be deliberate concealment/distortion of facts leading to errors-in-variables. Some of the response errors are not intentional. But, when it comes to opinion survey, errors in response take a special dimension.

It may be noted that the presence of explanatory variables is not essential in our analysis. Our aim in this paper is not to analyse the response as a function of the explanatory variables. Rather we want to see whether responses differ due to style of the questions. To make it simple we have grouped the responses into dichotomous variable. The responses are transformed into dichotomous responses as “agree” (coded as 1) for “strongly agree” or “somewhat agree”, and “disagree” (coded as 0) for “somewhat disagree”, “strongly disagree”. The answer “don’t know” was not considered though we lose some degrees of freedom. There was another set of questions in which same set of statements with different style giving opposite meaning was put. For example, corresponding to the statement “Women have the right to express their opinion if they disagree with their partner”, we put the opposite statement “Women do not have the right to express their opinion if they disagree with their partner.” The first set may be termed as positive (or affirmative) set of questions and the second set may be termed as negative set of questions[1].

In the next section, we summarise the answers into 2×2 contingency table for each question. To see whether the answers differ due to the styles of the questions, we perform Chi square test for this purpose. The results show that the answers differ in most of the cases. Consequently, in the next section, we build up a model with minimal assumptions and try to find the actual proportion of persons who agreed to the question. The next section tries to generalize the model. But the model turns out to be unidentifiable. However, it gives some insight to the results of the simple model. Finally, we discuss how the questionnaires should be formulated to achieve the correct proportion in the light of the results found in the paper.

[1]“Positive” (“negative”) set of questions does not necessarily give ethically/legally/socially accepted “positive” (“negative”) views for the family or society.

Data Summary and the Results of Chi-squares Tests

The following six tables give the frequency of persons who agreed and disagreed to the statement for both positive and negative statements. 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, we have shown the frequencies for the negative question in a reverse manner. 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, style of questions matters.

Table 1. *Results of χ^2 Tests of Significance of Male Responses to Positive & Negative Statements: Family Related 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**		
DF_1	Agree	49	Disagree	36	**0.000**	**Yes**
	Disagree	2	Agree	16		
	Total	51	Total	52		
DF_2	Agree	46	Disagree	37	**0.000**	**Yes**
	Disagree	6	Agree	13		
	Total	52	Total	50		
DF_3	Agree	29	Disagree	38	**0.002**	**Yes**
	Disagree	22	Agree	14		
	Total	51	Total	52		
DF_4	Agree	36	Disagree	27	**0.020**	**Yes**
	Disagree	15	Agree	25		
	Total	51	Total	52		
DF_5	Agree	29	Disagree	26	0.375	No
	Disagree	23	Agree	23		
	Total	52	Total	49		
DF_6	Agree	35	Disagree	26	0.060	Yes
	Disagree	17	Agree	25		
	Total	52	Total	51		
DF_7	Agree	39	Disagree	21	**0.027**	**Yes**
	Disagree	12	Agree	26		
	Total	51	Total	47		
DF_8	Agree	20	Disagree	22	0.109	Yes
	Disagree	29	Agree	29		
	Total	49	Total	51		
DF_9	Agree	13	Disagree	21	**0.001**	**Yes**
	Disagree	39	Agree	30		
	Total	52	Total	51		

*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 2. *Results of χ^2 Tests of Significance of Female Responses to Positive & Negative Statements: Family Related 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**		
DF_1	Agree	55	Disagree	31	**0.000**	**Yes**
	Disagree	0	Agree	19		
	Total	55	Total	50		
DF_2	Agree	49	Disagree	27	**0.000**	**Yes**
	Disagree	6	Agree	23		
	Total	55	Total	50		
DF_3	Agree	43	Disagree	28	**0.015**	**Yes**
	Disagree	12	Agree	22		
	Total	55	Total	50		
DF_4	Agree	28	Disagree	28	0.602	No
	Disagree	27	Agree	22		
	Total	55	Total	50		
DF_5	Agree	22	Disagree	19	0.834	No
	Disagree	33	Agree	31		
	Total	55	Total	50		
DF_6	Agree	33	Disagree	18	**0.018**	**Yes**
	Disagree	22	Agree	31		
	Total	55	Total	49		
DF_7	Agree	35	Disagree	16	**0.004**	**Yes**
	Disagree	18	Agree	28		
	Total	53	Total	44		
DF_8	Agree	18	Disagree	28	**0.028**	**Yes**
	Disagree	33	Agree	21		
	Total	51	Total	49		
DF_9	Agree	11	Disagree	19	**0.035**	**Yes**
	Disagree	44	Agree	30		
	Total	55	Total	49		

*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 3. *Results of χ^2 Tests of Significance of Responses of All Sampled Persons to Positive & Negative Statements: Family Related 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**		
DF_1	Agree	104	Disagree	67	**0.000**	**Yes**
	Disagree	2	Agree	35		
	Total	106	Total	102		
DF_2	Agree	95	Disagree	64	**0.000**	**Yes**
	Disagree	12	Agree	36		
	Total	107	Total	100		
DF_3	Agree	72	Disagree	66	0.623	No
	Disagree	34	Agree	36		
	Total	106	Total	102		
DF_4	Agree	64	Disagree	55	0.347	No
	Disagree	42	Agree	47		
	Total	106	Total	102		
DF_5	Agree	51	Disagree	45	0.751	No
	Disagree	56	Agree	54		
	Total	107	Total	99		
DF_6	Agree	68	Disagree	44	**0.005**	**Yes**
	Disagree	39	Agree	56		
	Total	107	Total	100		
DF_7	Agree	74	Disagree	37	**0.000**	**Yes**
	Disagree	30	Agree	54		
	Total	104	Total	101		
DF_8	Agree	38	Disagree	50	0.087	Yes
	Disagree	62	Agree	50		
	Total	100	Total	100		
DF_9	Agree	24	Disagree	40	**0.006**	**Yes**
	Disagree	83	Agree	60		
	Total	107	Total	100		

*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 4. *Results of χ^2 Tests of Significance of Male Responses to Positive & 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**		
DS_1	Agree	42	Disagree	39	0.478	No
	Disagree	10	Agree	13		
	Total	52	Total	52		
DS_2	Agree	51	Disagree	35	**0.000**	**Yes**
	Disagree	1	Agree	17		
	Total	52	Total	52		
DS_3	Agree	45	Disagree	22	**0.000**	**Yes**
	Disagree	6	Agree	29		
	Total	51	Total	51		
DS_4	Agree	43	Disagree	24	**0.000**	**Yes**
	Disagree	8	Agree	23		
	Total	51	Total	47		
DS_5	Agree	33	Disagree	25	0.300	No
	Disagree	19	Agree	22		
	Total	52	Total	47		
DS_6	Agree	24	Disagree	18	0.153	No
	Disagree	22	Agree	30		
	Total	46	Total	48		
DS_7	Agree	43	Disagree	26	**0.001**	**Yes**
	Disagree	8	Agree	23		
	Total	51	Total	49		
DS_8	Agree	45	Disagree	34	**0.016**	**Yes**
	Disagree	5	Agree	14		
	Total	50	Total	48		
DS_9	Agree	45	Disagree	31	**0.002**	**Yes**
	Disagree	3	Agree	14		
	Total	48	Total	45		
DS_{10}	Agree	36	Disagree	29	0.163	No
	Disagree	12	Agree	18		
	Total	48	Total	47		
DS_{11}	Agree	45	Disagree	29	**0.001**	**Yes**
	Disagree	7	Agree	22		
	Total	52	Total	51		
DS_{12}	Agree	43	Disagree	31	**0.013**	**Yes**
	Disagree	9	Agree	20		
	Total	52	Total	51		
DS_{13}	Agree	29	Disagree	4	**0.000**	**Yes**
	Disagree	23	Agree	48		
	Total	52	Total	52		
DS_{14}	Agree	27	Disagree	25	0.617	No
	Disagree	23	Agree	26		
	Total	50	Total	51		

*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 5. *Results of χ^2 Tests of Significance of Female Responses to Positive & 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		
DS_1	Agree	51	Disagree	27	**0.000**	**Yes**
	Disagree	4	Agree	19		
	Total	55	Total	46		
DS_2	Agree	50	Disagree	22	**0.000**	**Yes**
	Disagree	5	Agree	28		
	Total	55	Total	50		
DS_3	Agree	42	Disagree	22	**0.000**	**Yes**
	Disagree	11	Agree	27		
	Total	53	Total	49		
DS_4	Agree	36	Disagree	31	0.430	No
	Disagree	14	Agree	17		
	Total	50	Total	48		
DS_5	Agree	26	Disagree	21	0.590	No
	Disagree	28	Agree	28		
	Total	54	Total	49		
DS_6	Agree	36	Disagree	19	**0.015**	**Yes**
	Disagree	18	Agree	26		
	Total	54	Total	45		
DS_7	Agree	48	Disagree	27	**0.000**	**Yes**
	Disagree	6	Agree	20		
	Total	54	Total	47		
DS_8	Agree	42	Disagree	29	**0.035**	**Yes**
	Disagree	10	Agree	18		
	Total	52	Total	47		
DS_9	Agree	35	Disagree	27	0.084	No
	Disagree	9	Agree	16		
	Total	44	Total	43		
DS_{10}	Agree	42	Disagree	26	**0.020**	**Yes**
	Disagree	10	Agree	18		
	Total	52	Total	44		
DS_{11}	Agree	45	Disagree	26	**0.000**	**Yes**
	Disagree	6	Agree	21		
	Total	51	Total	47		
DS_{12}	Agree	47	Disagree	28	**0.002**	**Yes**
	Disagree	8	Agree	20		
	Total	55	Total	48		
DS_{13}	Agree	34	Disagree	8	**0.000**	**Yes**
	Disagree	19	Agree	42		
	Total	53	Total	50		
DS_{14}	Agree	35	Disagree	23	0.105	No
	Disagree	17	Agree	22		
	Total	52	Total	45		

*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 & 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**		
DS_1	Agree	93	Disagree	66	**0.001**	**Yes**
	Disagree	14	Agree	32		
	Total	107	Total	98		
DS_2	Agree	101	Disagree	57	**0.000**	**Yes**
	Disagree	6	Agree	45		
	Total	107	Total	102		
DS_3	Agree	87	Disagree	44	**0.000**	**Yes**
	Disagree	17	Agree	56		
	Total	104	Total	100		
DS_4	Agree	79	Disagree	55	**0.002**	**Yes**
	Disagree	22	Agree	40		
	Total	101	Total	95		
DS_5	Agree	59	Disagree	46	0.271	No
	Disagree	47	Agree	50		
	Total	106	Total	96		
DS_6	Agree	60	Disagree	37	**0.005**	**Yes**
	Disagree	40	Agree	56		
	Total	100	Total	93		
DS_7	Agree	91	Disagree	53	**0.000**	**Yes**
	Disagree	14	Agree	43		
	Total	105	Total	96		
DS_8	Agree	87	Disagree	63	**0.002**	**Yes**
	Disagree	15	Agree	32		
	Total	102	Total	95		
DS_9	Agree	80	Disagree	58	**0.001**	**Yes**
	Disagree	12	Agree	30		
	Total	92	Total	88		
DS_{10}	Agree	78	Disagree	55	**0.008**	**Yes**
	Disagree	22	Agree	36		
	Total	100	Total	91		
DS_{11}	Agree	90	Disagree	55	**0.000**	**Yes**
	Disagree	13	Agree	43		
	Total	103	Total	98		
DS_{12}	Agree	90	Disagree	59	**0.000**	**Yes**
	Disagree	17	Agree	40		
	Total	107	Total	99		
DS_{13}	Agree	63	Disagree	12	**0.000**	**Yes**
	Disagree	42	Agree	90		
	Total	105	Total	102		
DS_{14}	Agree	62	Disagree	48	0.127	No
	Disagree	40	Agree	48		
	Total	102	Total	96		

*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.

In all the six Tables, most of questions show significant differences of proportions of responses between positive and negative type questions. If we inspect frequencies minutely then we 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. 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 of 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 believes the statement to be true. To find answers to these questions we build up a model. Responses to the statement may be symbolically written as:

Table 7. *Responses to Positive& 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}

The Model of Submissiveness and the Results

Since there is a tendency of not going against a statement even if it may not be acceptable, we assume 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,

$$\alpha = P(\text{agreeing} \mid \text{unacceptable}) \qquad \dots (01)$$

α may be regarded as the probability of a person, who does not believe the statement to be true, 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 may also be regarded as the probability of deviation from truth. We assume that α is same for both positive and negative statements.

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. Thus,

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

i.e., the probability that a positive statement is acceptable is same as the probability that a negative statement is unacceptable. Evidently, $P(A \mid \mathcal{P}) + P(A^c \mid \mathcal{P}) = 1$. Our 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, we can 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, we have,

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

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

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

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

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

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

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

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

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

Note 1.1: The other way of estimating α is $\alpha = P(a \mid \mathcal{P}) - P(a^c \mid \mathcal{P}^c)$. … (07)
This means that $\alpha = 0$, when proportion of agreeing to the positive statement is same as the proportion of disagreeing to the negative statement.

Note 1.2: Observe that, $P(A \mid \mathcal{P})$ is not same as that we get from common sense as $(P(a \mid \mathcal{P}) + P(a^c \mid \mathcal{P}^c))/2$. This is because we have made a non-symmetric assumption that only those respondents who think that the statement is unacceptable can express it otherwise. 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.

$$\begin{aligned} P(A \mid \mathcal{P}) &= (P(a \mid \mathcal{P}) - \alpha)/(1-\alpha) \\ &= (P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) + 1)/(1 - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c) + 1) \\ &= (1 - P(a \mid \mathcal{P}^c))/(2 - P(a \mid \mathcal{P}) - P(a \mid \mathcal{P}^c)) \\ &= P(a^c \mid \mathcal{P}^c)/(P(a^c \mid \mathcal{P}) + P(a^c \mid \mathcal{P}^c)) \quad \ldots (08) \\ &= (N_{12}/N_{02})/(N_{21}/N_{01} + N_{12}/N_{02}) \quad \ldots (09) \end{aligned}$$

= (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

$$P(A \mid \mathcal{P}) - (P(a \mid \mathcal{P}) = \alpha(P(A \mid \mathcal{P}) - 1). \quad \ldots (10)$$

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

It is clear that there are variations in the values of α and P(A│$\mathcal{P}$) 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 our 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, we have also shown the percentage changes from P(a│P) to P(A│P). In many of the cases these percentage changes were more than 10%, especially from statements 6 through 9 of family related issues and statements 3, 6 and 13 of social issues. 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 respect 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.

We have seen in Tables 10 and 11, 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). Percentage changes were 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. We have thus shown the differences in the form of percentage in separate columns. These differences were 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.

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*

	Male				Female				All			
Qn. No.	**A**	**P(A│P)**	**P(a│P)**	**% Change**	**A**	**P(A│P)**	**P(a│P)**	**% Change**	**α**	**P(A│P)**	**P(a│P)**	**% Change**
DF_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
DF_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
DF_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
DF_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
DF_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
DF_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**
DF_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**
DF_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**
DF_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*

	Male				Female				All			
Qn. No.	**A**	**P(A│P)**	**P(a│P)**	**% Change**	**A**	**P(A│P)**	**P(a│P)**	**% Change**	**α**	**P(A│P)**	**P(a│P)**	**% Change**
DS_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
DS_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
DS_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**
DS_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
DS_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
DS_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**
DS_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
DS_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
DS_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
DS_{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
DS_{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
DS_{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
DS_{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**
DS_{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

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 (%)
DF_1	0.268	0.380	-29.5	11.2	0.946	1.000	-5.4	5.4
DF_2	0.145	0.351	-58.7	20.6	0.865	0.832	4.0	-3.3
DF_3	-0.162	0.222	-173.0	38.4	**0.629**	0.720	-12.6	9.1
DF_4	0.187	-0.051	-466.7	-23.8	0.638	0.533	19.7	-10.5
DF_5	0.027	0.020	35.0	-0.7	0.545	0.388	40.5	-15.7
DF_6	0.163	0.233	-30.0	7.0	0.609	**0.479**	27.1	-13.0
DF_7	0.318	0.297	7.1	-2.1	**0.655**	**0.517**	26.7	-13.8
DF_8	-0.023	-0.218	-89.4	-19.5	0.422	**0.469**	-10.0	4.7
DF_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 (%)
DS_1	0.058	0.340	-82.9	-28.2	0.796	0.890	-10.6	-9.4
DS_2	0.308	0.469	-34.3	-16.1	0.972	0.829	17.2	14.3
DS_3	0.451	0.343	31.5	10.8	**0.786**	**0.684**	14.9	10.2
DS_4	0.332	0.074	348.6	25.8	0.765	0.698	9.6	6.7
DS_5	0.103	0.053	94.3	5.0	0.593	0.453	30.9	14.0
DS_6	0.147	0.244	-39.8	-9.7	**0.439**	**0.559**	-21.5	-12.0
DS_7	0.313	0.314	-0.3	-0.1	0.772	0.838	-7.9	-6.6
DS_8	0.192	0.191	0.5	0.1	0.876	0.762	15.0	11.4
DS_9	0.249	0.168	48.2	8.1	0.917	0.754	21.6	16.3
DS_{10}	0.133	0.217	-38.7	-8.4	0.712	0.754	-5.6	-4.2
DS_{11}	0.297	0.329	-9.7	-3.2	0.809	0.825	-1.9	-1.6
DS_{12}	0.219	0.271	-19.2	-5.2	0.778	0.800	-2.8	-2.2
DS_{13}	0.481	0.482	-0.2	-0.1	**0.148**	**0.309**	-52.1	-16.1
DS_{14}	0.050	0.162	-69.1	-11.2	0.516	0.610	-15.4	-9.4

The Model of Submissiveness and Assertiveness

To get a better view of the data, let us assume that, in addition to α, we have β, the proportion of persons, among persons who thinks that the statement is acceptable, disagrees with the statement. We also assume that both α and β are same for both positive and negative statements. Thus,

α = P(agreeing│unacceptable): same for both positive and negative questions, and

β = P(disagreeing│acceptable): same for both positive and negative questions.

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

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

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

$P(A^c \mid \mathcal{P})$, $P(A^c \mid \mathcal{P}^c)$ 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 we have,

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

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

Adding equations (13) and (14), we get,

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

Our earlier model is a special case of this model, i.e., when β = 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)

We can solve equations (15) and (18) to get α and β. On the other hand, if any one of α and β is known then we can get the other one from equation (15) and then get $P(A \mid \mathcal{P})$ from equation (18) as

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

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

---

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

In any case we 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).

Discussions

One should put positive as well as negative questions in the questionnaire. But the problem is that the same question with two different styles cannot be put to the same person. What we can do is that two sets of questionnaires should be prepared. 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.

The study done through these types of questions can give a clear picture about the impact of style of questions on an individual. Besides, we will be able to get the actual opinion of a group of individuals in a summary form by canvasing questions of both the styles. This may help us to locate the root of these problems and thus enable us to prescribe appropriate policies addressing these problems, which, in turn, will enable us to eradicate this evil crime from the society.
The methodology described in this paper is not applicable to gender violence only, but also to other phenomena of the society needing the views of the members in the society. Thus, it has a much broader spectrum of applications than one can think of initially.

References

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

Gelman, Andrew and Hill, Jennifer (2007). Data analysis using regression and multilevel/hierarchical models / Cambridge ; New York : Cambridge University Press, 2007

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

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). Survey methodology (Chap. 2). New York: John Wiley.

Hansen, Morris H.; Hurwitz, William N.; Marks, Eli S. and Mauldin, W. Parker (1951): *Journal of the American Statistical Association*, Vol. 46, No. 254, pp. 147-190.

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

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

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

Lessler, J. T., & Kalsbeek, W. D. (1992). Nonsampling error in surveys. New York: John Wiley.

Lyberg, L., and Kasprzyk, D. (1991): Data collection methods and measurement error: An Overview. In *Measurement Error in Surveys,* P. Biemer, S.Sudman, and R.M. Groves, eds. New York: Wiley and Sons.

Menard, Scott (2002). *Applied Logistic Regression Analysis*. SAGE
Wooldridge, Jeffrey (2010). *Econometric Analysis of Cross Section and Panel Data* (Second ed.). Cambridge: MIT Press.

Appendix A

Table A1. *Family Related Type A and Type B Statements*

Statements (Type A)	Statements (Type B)
DF1: Women have the right to express their opinion if they disagree with their partner.	DF1: Women do not have the right to express their opinion even if they disagree with their partner.
DF2: Couples should share equally in household chores if they are both working outside the home.	DF2: Couples should not share equally in household chores if they are both working outside the home.
DF3: Women can work outside the home even if the family does not need the money.	DF3: Women should not work outside the home even if the family needs the money.
DF4: Physical violence between couples is a private matter and should be handled within the family.	DF4: Physical violence between couples is not a private matter and should be handled even with outside intervention, if needed.
DF5: It is the duty of women to obey their partner always.	DF5: It is not the duty of women to obey their partner always.
DF6: If women want to see their relatives or friends, they should take permission from their partner.	DF6: If women want to see their relatives or friends, it is not necessary to take permission from their partner.
DF7: Violence against women happens more in the families with low incomes.	DF7: Violence against women happens less in the families with low incomes.
DF8: Women should be submissive.	DF8: Women should not be submissive.
DF9: Men are superior to women in all respect of life	DF9: Men are not superior to women in all respects of life

Table A2. *Type A and Type B Statements Related to Social Views*

DS1: Abuses to women have increased significantly now-a-days.	DS1: Women abuses have not increased significantly now-a-days.
DS2: Women have equal rights as men.	DS2: Women do have equal rights as men.
DS3: Men and women are equally responsible for violence against women.	DS3: Men and women are not equally responsible for violence against women.
DS4: Violence against women occurs due to belief in the traditionally defined roles of women and men.	DS4: Violence against women occurs not due to belief in the traditionally defined roles of women and men.
DS5: Abuse against women has increased due to more freedom in their mobility.	DS5: Abuse against women has increased not due to more freedom in their mobility.
DS6: Violence against women happens more among women with physical and or mental disabilities.	DS6: Violence against women happens more not necessarily among women with physical and or mental disabilities.
DS7: Too much alcohol consumption of men results in more violence against women.	DS7: Too much alcohol consumption of men does not result into more violence against women
DS8: Recent increase in the violence against women is due to the decline in the moral values in our society.	DS8: Recent increase in the violence against women is not due to the decline in the moral values in our society.
DS9: Gap between aspirations and means among "Prospect-less Young Men" is responsible for increasing incidence of women abuse.	DS9: Gap between aspirations and means among "Prospect-less Young Men" is not necessarily responsible for increasing incidence of women abuse.
DS10: Gender violence is the outcome of the natural attitude of living beings to hold dominating power.	DS10:Gender violence is not the outcome of the natural attitude of living beings to hold dominating power.
DS11: Lack of proper enforcement of existing policies/laws to protect women is the reason for increasing abuse on them.	DS11: Lack of proper enforcement of existing policies/laws to protect women is not the reason for increasing abuse on them.
DS12: Insufficient infrastructure (street light, absence of ladies toilet, in-secured public transport, etc.) is one of the main reasons for violence against women.	DS12: Insufficient infrastructure (street light, absence of ladies toilet, in-secured public transport, etc.) is not one of the main reasons for violence against women.
DS13: Protest by women leads to increase in the degree of violence against them outside home.	DS13: Protest by women leads to decrease in the degree of violence against them outside home.
DS14: Protest by women leads to decrease in the degree of violence against them at home.	DS14: Protest by women leads to increase in the degree of violence against them at home

Appendix B

1.1. 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

We have 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

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

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

We can also find the bivariate distribution of (x_i, X_i) 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–α)(1–p)	α(1–p)	1-p
X = 1	βp	(1–β)p	P
Subtotal	βp+(1–α)(1–p)	(1–β)p +α(1–p)	1

We can write 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, we have $x \sim Bin(N, (1 - \beta) p + \alpha (1 - p))$. For estimation of the parameters we can use,

$\bar{x} = N[(1 - \beta) p + \alpha (1 - p)]$ 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

We have 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, we can wite

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

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

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

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 we have the bivariate distribution of (y,Y) as follows:

y / **Y**	**y = 0**	**y = 1**	**Subtotal**
Y = 0	(1–α)p	α p	p
Y = 1	β(1–p)	(1–β)(1–p)	1-p
Subtotal	β(1–p)+(1–α)p	(1–β)(1–p)+αp	1

Thus, we can write in reverse direction as

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

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

Suppose $y = \Sigma y_i$. Since y_i's are iid, we have $y \sim Bin(N, (1 - \beta) p + \alpha (1 - p))$. For estimation of the parameters we can use,

$\bar{y} = N[(1 - \beta)(1 - p) + \alpha p]$ 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 we cannot use it, because we have only one observation for each. If we assume 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 that, because we can estimate (β – α) and the values of (β – α) differ from question to question.