天主教輔仁大學英國語文學系學士班畢業成果

ENGLISH DEPARTMENT, FU JEN CATHOLIC UNIVERSITY
GRADUATION PROJECT 2018

指導教授：施佑芝老師
Dr. Doris Shih

Are You Aware of the Alternatives Untold? — Impact of Expectation Bias

學　生　：　唐　　欣　撰
Carol Hsin Tang

403110125 Carol Tang

English Composition & Conversation III, B

Professor Shih

21 June 2017

Research Paper Draft 4

Are You Aware of the Alternatives Untold? — Impact of Expectation Bias

I. Topic: Discussion on how experimenters' personal expectation and attitude reflecting on questionnaire narratives may affect respondents' answers to questionnaires and thus obstruct the validity of research outcomes in terms of cognitive bias.

II. Introduction

A. Research studies are important in this era of explosive information

1. Noise in statistic process can disturb policymakers and other data users (Burger et al. 499)
2. Survey results are largely affecting people's life
3. Survey validity is thus of critical importance

B. Surveys with similar topics can have various outcomes as observed

1. Research validity can be obstructed by many factors (Betsch)
2. Case of Clever Hans the smart horse (Chegg)
3. Thesis statement: The expectation bias of researchers would subconsciously show on the narratives of questionnaire questions, and respondents would therefore be affected as well

C. Research Questions:

1. Does observer expectancy effect truly affect respondents?
2. Are the respondents aware of the effect?
3. Is the effect as effective when respondents have a firm stance?

III. Literature review

A. Experimenters do affect research validity

1. Study on narrative bias as supports the argument (Betsch)
2. Staff of the experiment are prone to have evaluation bias (Scriven)
3. Researchers tend to have personal preferences—the expectation bias (West)

B. Still other factors affect research validity

1. Review of cognitive bias modification and interpretation is another proof (Menne-Lothmann)
2. Another supportive point is the classification error (Burger)

C. Measurements to avoid bias and to correct survey validity

IV. Methodology

A. Questionnaire design: Part II—simulation questions

1. 7 situational questions in total
2. Using three point Likert scale from agree (3), no comments (2), to disagree (1).
3. Taking negative stance to be more stimulating for respondents (Betsch)

B. Distribution—through Facebook, using google form

C. Analysis

1. Using pie charts
2. Calculating counts of the simulation questions
3. Cross analyzing by comparing their stance with the counts

V. Results & Discussion

A. RQ 1: Does observer expectancy effect affect respondents?

B. RQ 2: Are the respondents aware of the effect?

C. RQ 3: Is it still effective when respondents have firm stances?

VI. Conclusion:

A. Implications: Half of the respondents are likely to be affected by observer expectancy effect even though they have firm stances and two-third of them are aware of it.

B. Limitations

1. Possible impact of social desirability bias (abortion)
2. Revelation of invalidity in certain conditions but not answer to all cases
    a. A more extreme scaling measure (3-point scale only)
    b. Overtly negative narratives

C. Suggestion: People should be more careful when facing surveys and results, also people should try to stay skeptical to the facts that seem granted

Are You Aware of the Alternatives Untold? — Impact of Expectation Bias

Research studies have a foothold in today's era of explosive information when people are accustomed to that so-called facts and knowledge are mostly based on statistics and research studies, and that these data go on to provide secondary sources for even more researches. These research studies form an important part of today's society, since these data are used in various fields that greatly influence people's everyday life, such as collecting the public's voice and making policies. What people are not aware of is that some secondary sources are mixed with noises due to error sources and are lack of confirmation to their validity (Burger et al.).

One of the most manifest aftermath of misused or invalid sources is the phenomenon that surveys of similar topics sometimes turn out having diverse outcomes, or even contradictory to one another. Such condition causes people to find themselves at a loss and not knowing which way to follow. There have been several researches that deal with the invalidity of research, and the factors are many. Some are problematic analytic process while some others have to do with bias (Scriven). In the analyzing process of survey data, there could be classification error and analytic error, for examples. The classification error lies in the failure to separate real statistics from noise, which means the "error." Burger makes an interesting argument that statistics can be constructed by administrative data alone while administrative data do not aim at serving statistical purposes. According to Bakker and Daas, the risks are "selective undercoverage, and administrative units and variables may not match statistical definitions" (qtd. in Burger et al. 490). These official statistics are facing an obstacle that they are often based on "a mix of source that typically do not involve probability sampling,"

(499) so that their reliability are called in doubts. Moreover, a series of investigation done in 2016 on the analytic errors in secondary analyses provides the answer to the failure of analyzing correctly, which are extended prior knowledge, ignoring respondents' sample units, and the replicate criteria, oversampling particular subgroups, misleading implications for descriptive estimates or standard errors that do not connect with sample design features (West). Among these factors, misleading implications for descriptive estimates is also relevant to the next discussion of the reasons for invalid survey—the cognitive bias of the biases.

Under the category of biases, there remains to be narrative bias, general cognitive bias, and expectation bias, to name only a few in this research. Starting from the narrative bias, studies show that narrative serves as an influencial stimulation when judging a risk or its probability even when a statistical base rate is provided (Betsch et al. 241). Even though statistics is enough for people to make a judgement, they are still inclined to be strongly affected by negative narratives since the possibility of risks increase people's awareness and enhance people's perception toward risks, which attribute to behavioral intentions (Betsch et al. 242, 253). This study clearly showcases the likelihood of survey being biased as respondents may be easily affected by negative narratives. The research study of Menne-Lothmann also backs up Betsch's findings. Menne-Lothmann and her crew use cognitive bias modification training as a means to boost the mood in the psychological field. Cognitive bias is the general thinking error that obstructs correct cognition. Using the modification training of cognitive bias in their experiment as a clinical tool, they reveal a positive relevance among the removal of cognitive bias, optimistic interpretation and decreasing bad mood which accord with the previous research result of

Betsch's. The expectation bias also exposes similar tendency to affect the survey. The expectation bias, also known as experimenter's expectation bias, is the observers' preference over the results they have hypothesized (W). All of the studies above reach a consensus to show that surveys can be influenced by experiment conductors' narratives and personal preference, that is to say, the researchers' stance do affect the survey and challenge its validity.

To follow up the argument that experimenters' attitude and personal point of view are affecting the survey, firstly it is critical to understand how are experimenters biased more precisely. The experiment staff are reported in Scriven's research to not be objective due to social and economic bonds (5). He also points out that the staff fails to be impartial as they succumb to "the cumulative effect of repeated acceptance (or rejection) of evaluative suggestions." (6) Despite the bias performed by the experiment team, Scriven also suggests that a bias can be a "statistically likely tendency to systematic errors" or "an actual increase in the frequency of errors" (19) while only the actual errors jeopardize research validity. In this research paper, whether there is a positive connection between tendency of error and actual amount of error frequency will be further discussed through the questionnaires of bias and their outcomes. Another point Scriven has made is that respondents' evaluation of issues should not be restricted by validity since it is not an inflexible question but more of a sense. However, putting validity aside, it remains a debatable question whether it put the survey result in vain if researchers have a hand in manipulating the audiences' perception on survey questions. To examine the exact effect of researchers' subconscious on the survey outcome, this research paper will compare and contrast the outcome under both experimenters' expectation bias and the subject-expectancy effect (the expectation bias of the audience) using questionnaire that showcases strong stance

in the wordings and attitudes on the questions, and finally ask respondents' personal stance at the end of the survey to see if the previous answers to strong stance differ from their original conception. Finally, with the data collected from questionnaire questions, this research paper will collate the result and manage to provide an explanation whether experimenters' attitude truly intervene the survey validity. At the end of the research paper, via analyzing the data and studying literary references, this study will also manage to provide notes for both researchers and respondents to avoid such bias when producing and answering questionnaires.

The methodology of this research paper is quantitative by collecting data with online questionnaire. The questionnaire is using a three-point Likert scale, scaling from "1" to "3", as "1" is disagreeing with the narration of the questions while "2" is having no opinion and "3" means consent. This design is due to a finding by Betsch (249), revealing that the 7-point scale is less sensitive and less capable of showing the difference between variables. The more options are provided for the respondents, the less drastic the outcomes will appear, which makes the survey outcomes subtle and difficult to analyze. Therefore, in this survey the scale point will be only three. In the questionnaire, there are three sections of questions, respectively; they are respondents' background information questions, biased stance simulation questions, and questions toward questionnaire-filling perceptions.

For starters, this survey does not aim at specific target audience since this issue could be universally critical for all walks of life, young or old, rich or poor; as a result, in the questionnaire, several conditions are provided for the respondents, trying to put them into different groups with an eye to having a clearer insight into which groups of people prone to be affected by the observer-expectancy effect to a

greater extent. The conditions provided to group the respondents include gender, age, level of education, and their personal experience of filling and designing questionnaires, that is, their familiarity and understanding of questionnaire.

In the second part, a set of biased stance questions are designed to simulate how researchers have their own perspectives affect the research. (Please refer to Appendix A). The subject matter adapted in this part is abortion, as abortion is a controversial issue and has no definite right or wrong. In such case, respondents, who are also message receivers while reading narratives of the questionnaire, can be more open for possibilities and opposing opinions under diverse circumstances. According to Betsch, readers are more sensitive about negative narrative comparing to a positive one since the former is like a warning of possible threats (242). To make the questionnaire more perceptible and to have a rather explicit survey outcome, the simulation questions use negative narrations that deny abortion as the researcher's stance. Also, at the very beginning of the second part of the questionnaire, respondents are asked about their personal stance toward abortion. Their stance toward abortion will later be used as the criteria for reference to compare with the survey outcome to see to what extent are respondents affected by the negative narration.

The third part of the questionnaire is perception questions, that is, to inquire respondents of their feelings after answering questions in section two (Also see Appendix A). The point of this part is to examine whether respondents are aware of the researcher's desire to manipulate their answers. Aside from checking whether respondents are under the effect of expectancy bias with their original stance and their degree of consent to the negative narratives, the third part goes an extra mile to inquire the respondents directly on whether they themselves consider the biased stance influential to them after filling out questions above, and whether they

consider the negative narrations as effective when they themselves have a firm stance toward the topic. By asking respondents per se about the effect of the questionnaire, not only can the result be double-checked, but testing whether respondents are aware of the effect on themselves as well.

To collate the data, the author compares respondents' stances answered in the first question of the second part to the counts calculated from the simulation questions. The result of such cross examination shows whether respondents were moved from their original stance, therefore answer to the first research question—Are the respondents truly affected? Afterwards, put together the result of the first research question with the respondents' recognition of whether they think the narratives are influencing them. Of those people who shift their stances in the first question, the ones who think the narratives are not effective are deemed as unaware of the effect. As for the third question, counts and stance are again pull together for discussion on the respondents' shifting condition when they have firm stances. The primary tools to analyze respondents' reaction is pie charts and counts calculated from the simulation questions.

In order to answer the first research question— "Does observer expectancy effect affect respondents?" —respondent's stances answered by the first question of the second part and their counts calculated from the seven simulation questions must be taken into consideration. In the stance question— "Are you for, against, or neutral about abortion?" —46 respondents (18.6%) answered "against." (See Pie Chart 2 in Appendix B) However, when going on to examine the points in the seven simulation questions, 149 respondents (60.3%) have points above average (14). What's more, 30% of them even have rather high points above 18. Since there are 7 questions and each one is scales from 1 to 3, the average counts is 14, and as three counts is a range, total counts above 17 (14+3) is considered rather extreme.

Also, as the narratives of the questionnaire is negative, the higher count represents higher level of disapproval of abortion. The result is quite manifest as the respondents who show disapproval of abortion increase from 18.6% to 60. (See Chart 1 in Appendix B) This survey result also resonates with West's study that misleading implications for descriptive narrations do affect respondents and thus, have certain impact on survey validity.

The second research question— "Are the respondents aware of the effect?" —is answered by the third part of the questionnaire, which is the perception questions. According to the survey outcome (For this discussion, please turn to Pie Chart 3, 4, and 5 in Appendix B), 36.8% of respondents are either not sure or not thinking the narratives have a stance while the narratives are deliberately designed to be negative. This result shows that the 36.8% of respondents are not sensitive to researcher's attitude. To double-check with the result, respondents were also asked what the stance of the questionnaire is as they perceive. Interestingly, 31.5% of them think that the questionnaire is either neutral or even holding positive attitude while the case is totally the opposite. This result again showcases the unawareness and recklessness about narratives. Then the survey also raises inquiry about whether the respondents think that questionnaire narratives can affect the answers; however, 35.2% of the respondents answered "no" while in effect, according to the first research question, respondents are indeed influenced. Concluding from the results above, it seems that roughly one-third respondents are not aware of the effect. Last but not least, another question in the perception part asking "have you ever filled out any other questionnaires that makes you feel they have a stance" also attributes to the conclusion as 106 respondents (42.9%) states that they do not remember whether they have encountered surveys with bias before, which shows the respondents' insensitivity.

Then it comes to the third research question— “Is the effect as effective when respondents have firm stance?” The answer is analyzed from both the results of the first research question and the responses from the third part of the questionnaire. (See Table 6 in Appendix B) By inquiring the respondents whether they think the questionnaire exert same influence when participants have firm stance already, the author gets negative feedback from 144 respondents (58.8%). However, the outcome after comparing respondents’ original stance and their reaction in simulation questions seems very different from what the respondents think. In the research, respondents’ stances are tracked and compared with the total counts of simulation questions. Among the 32 people who claim to support abortion, 2 shift to the opposite stance though simulation questions and both of them have high counts (both 19) that imply strong disapproval of abortion. Their counts of 19 are even higher than many respondents who hold neutral stance. Of the 166 participants who claim to take a neutral stance, 38 of them shift to be against abortion and 2 of them even get the full points of 21. It appears for the third research question that people with firm stance have relatively little chance of changing their viewpoints but not impossible. Those who are more conservative and hold neutral stance are especially prone to shift stances.

In conclusion, this survey implies that under certain conditions like extreme narratives in a negative style, more complicated or even controversial topic, and highly sensitive scaling measure (3-point only), respondents indeed are affected by the observer expectancy effect. The result of the research indicates that more than a half of the respondents but not all of them are affected while only about one-third of them are aware of the researcher’s intention or subconscious stance. Moreover, even if the participants have strong and firm stance, they are still possibly to be affected but with much lesser chance comparing to those who hold neutral stance.

There is no direct literature review that explains how psychological effects have impact on respondents, but this survey, to some degree, helps to bridge the gap between researchers and respondents, proving that observer-expectancy effect truly affects the respondents more or less.

Undeniably, there are some limitations to the survey. For example, as the topic is on a rather controversial issue— abortion, the respondents may also be affected by social desirability bias, which is the effect that people tend to follow the moral code of the society. That is to say, respondents may think that abortion is still somehow immoral to our society and thus make decisions to oppose it. However, the author has managed to cut down on the effect by distributing the questionnaire online anonymously. Other factors that challenged this survey include the certain conditions under which the survey is being done as mentioned in the conclusion. The results may differ should the issue of simulation questions is different. As abortion is a complicated issue, the reaction can be overtly different under different situations. The survey result here can only reveal that the survey outcomes can be misleading if only concerning the statistics, but the actual perception or what the respondents actually think about abortion cannot be observed from this survey. The negative wordings also have to do with the overt influence; if the observer is holding a more positive stance then the survey outcome may not be as affected. The use of three-point Likert scale is also a factor that leads to such research results. The usage of a more sensitive scaling is to exemplify the possible effect, but the condition designed in the questionnaire is not the normal case for research studies. The author is trying to simulate an extremely biased condition in terms of questionnaire study with an eye to manipulating the respondents. However, a more comprehensive survey would probably avoid these biases by various measures such as using a more objective Likert scale of seven,

raising questions from not only negative aspects but positive ones as well and so on.

Finally, there are some suggestions that may help avoid misleading research. Always be skeptical when doing questionnaires and as a media receiver, do not take all the data accessed for granted. The research results of a single study cannot represent the comprehensive condition. What's more, as survey designer, if you want to be more objective, use a less sensitive but fairer 7-point Likert scale rather than a 3-point one, since the more options the respondents are granted, the more objective they would be, and last but not least, always ask questions from the both sides to double check with the respondents.

Appendix A

Questionnaire Title:

期望偏差對問卷研究之影響 Impact of Expectation Bias on Questionnaire Research

Face-sheet:

您好

我是輔大英文系三年級的學生，目前正在進行一項關於問卷設計者和受試者因為不同立場所造成的研究偏差。

本問卷資料僅供學術用途，所有答案都會幫您保密，請放心按照真實想法填答。

以下問題需要您五分鐘的注意力，請在讀過敘述後回答問題。

非常感恩您的配合！祝您一切順心。

學生：輔仁大學英國語文學系　唐欣

指導教授：施佑芝　教授

For those who may concern,

I am a junior in the English Language and Literature Department of Fu-Jen Catholic University. This survey is to help me collect data for my study on the relation between respondents' perceptions on questionnaires and the survey designers' personal outlook of the subject matter. This survey is only for academic purposes and will be kept confidential. Feel free to answer the questions with your real thoughts. Please answer the following questions intuitively after reading the narration. It takes about five minutes to answer all the questions.

Thanks a lot for your kind corporation!

Instructor: Prof. Doris Shih

Student: Carol Tang

Questions:

I. 基本資料 Background Questions:
    1. 請問您的性別? What is your gender?
        a. 男 Male /女 Female
    2. 請問您幾歲? How old are you?
        a. ~18/ 18~22/ 23~30/ 30~40/ 40~50/ 50~60/ 60~
    3. 請問您的教育程度? What is your level of education?
        a. 國中或以下 Junior High School/ 高中 Senior High School/

大學 College/ 碩士 Master's Degree/ 博士 PhD

4. 請問您填答過多少問卷? How many questionnaires have you filled out?
    a. 這是第一份 This is the first one/ 少於五份 Less than five/ 五到十份 Between five and ten/ 十到二十份 Between ten and twenty/ 多於二十份 More than twenty
5. 請問您曾經製作過問卷嗎? Have you ever designed a questionnaire?
    a. 有 Yes/ 沒有 No

II. 模擬問題 Simulation Questions

在這個區段將會以「墮胎」作為討論主題。請在詳讀以下敘述後，選擇 1~3 來表達您對題目論述的認同程度。3:認同，2:不一定或是沒有意見，1:不同意

In this part, "abortion" will be used as the subject matter.
Please choose from 1 to 3 after reading the narratives below.
3 = agree, 2 = not sure/no comments, 1 = disagree

1. 我對墮胎的立場是? Are you for, against, or having no comments about abortion?
    a. 同意 Agree/ 沒有意見 No comments/ 反對 Disagree/ 不一定 It depends
2. 我覺得墮胎是不負責任的行為 I think abortion is irresponsible
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree
3. 我覺得墮胎這件事並不尊重生命 I think abortion disrespect life
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree
4. 我認為現在的年輕人並沒有認真看待墮胎的嚴重性 I think adolescents nowadays do not regard abortion serious enough
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree
5. 如果我聽說同學去墮胎，我會覺得她私生活很亂 If I hear that my classmate had an abortion, I would have negative association on her personal life
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree
6. 我想大多數的人並沒有意識到墮胎可能會有的後果(比方說血崩,不孕...) I think most people are not conscious about the possible aftermaths of an abortion (metrorrhagia, infertility, etc.)
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree
7. 墮胎讓年輕人對於性變得更隨便 I think the possibility of abortion is making youngsters reckless about sex
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree

8. 我覺得墮胎對社會的傷害比幫助還多　I think abortion do more bad than good to the society
    a. 1 Disagree/ 2 Not sure or no comments/ 3 Agree

III. 受試者感受問題 Perception Questions

以下將探討主觀的提問是否帶給受試者偏頗的感覺，請依照剛才回答前一個區段問題的感受來回答下列問題。In this part, we will discuss whether biased question is perceived by respondents. Please answer questions below with your experience and feelings while you were answering questions in the previous section.

1. 您是否感覺上面的問題帶有主觀立場? Do you feel that the survey itself has a stance?
    a. 是 Yes/ 否 No/ 我不確定 I'm not sure
2. 您覺得上面問題對墮胎的立場為? What do you think is the stance of the questions above?
    a. 同意墮胎 For abortion /反對墮胎 Against abortion /中立 Neutral
3. 您是否同意此問卷的看法? Do you agree with the viewpoint of the questions above?
    a. 是 Yes / 否 No/ 不一定 It depends
4. 您覺得問卷的立場是否會影響受試者? Do you think the questionnaire attitude affect the respondents?
    a. 是 Yes/ 否 No
5. 承上題，您認為當受試者本身對議題很有想法時會有一樣的效果嗎? Do you think the questionnaire exert the same influence when respondents have a stance already?
    a. Yes 是/ No 否
6. 填完問卷後，您是否對墮胎改變想法? After filling out the questions, do you change your thoughts about abortion?
    a. 是 Yes/ 否 No
7. 您是否曾經做過其他問題很主觀的問卷? Have you ever filled out any other questionnaires that makes you feel they have a stance?
    a. Yes 是 /No 否 /I can't remember 我不記得了

作答完成 Questionnaire Complete!

非常感謝您的配合！如果對問卷有任何疑問，歡迎與我聯繫。

Thank you again for your patience and cooperation. Please contact me if there is any problem regarding to the questionnaire.

caroline1217tang@icloud.com

Appendix B

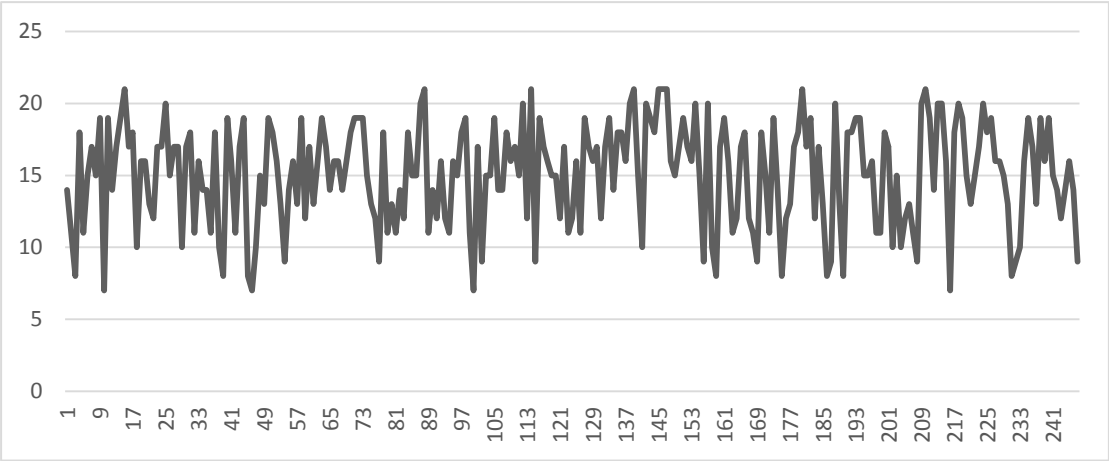
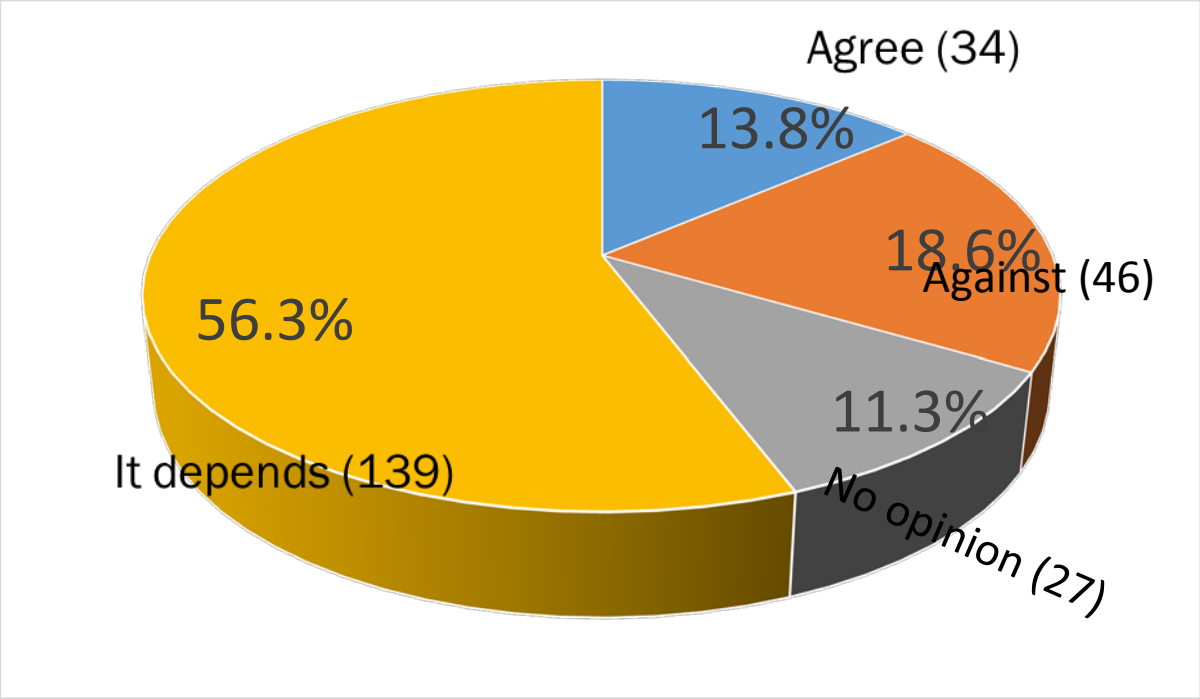
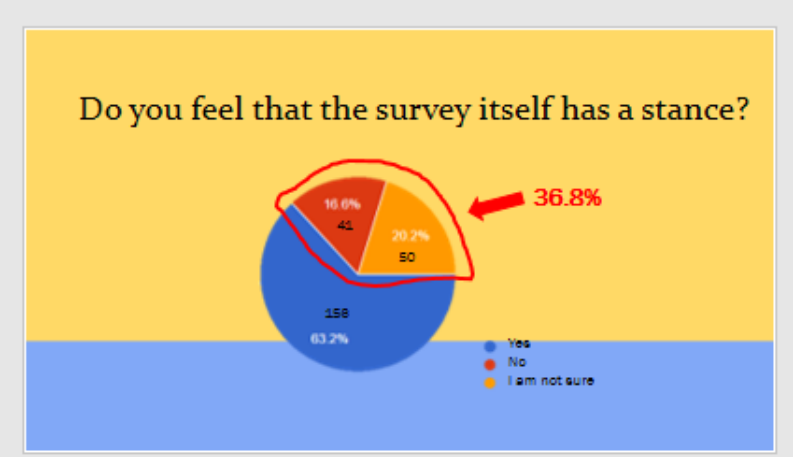


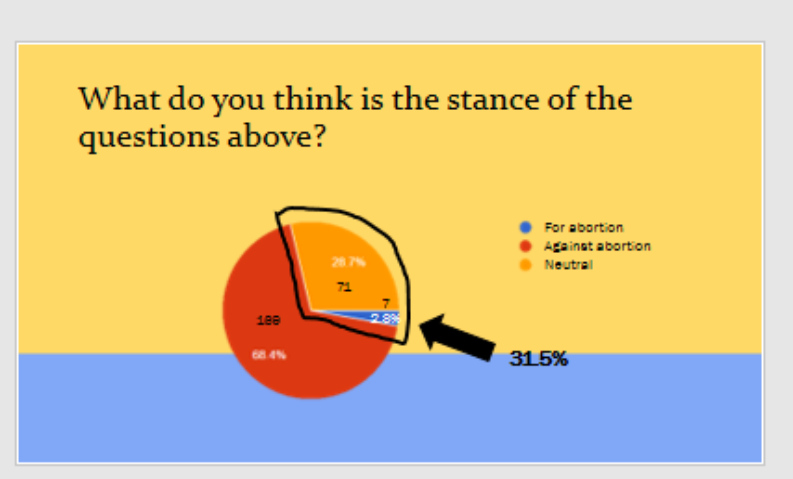
Chart 1: This is the calculated counts of the 247 respondents. 60.3% of them have points above 14; 30% of them have even higher points than 17. (149>average counts; 79=average counts; 17<average counts)



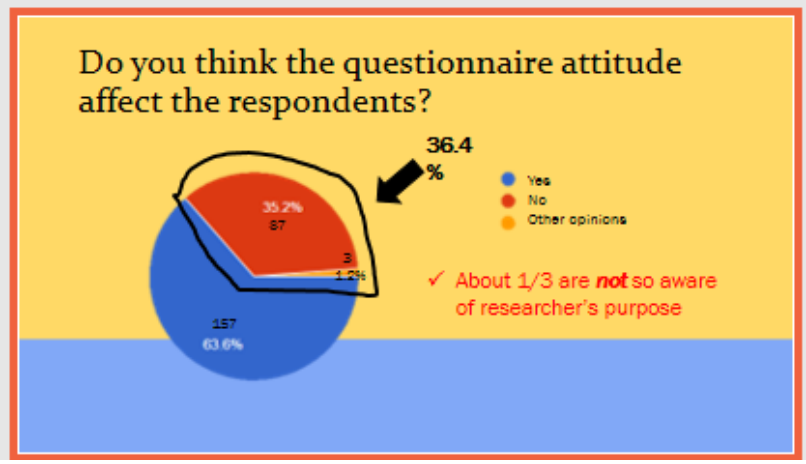
Pie Chart 2. This chart shows the responses of the question, “Are you for, against, or having neutral about abortion?”



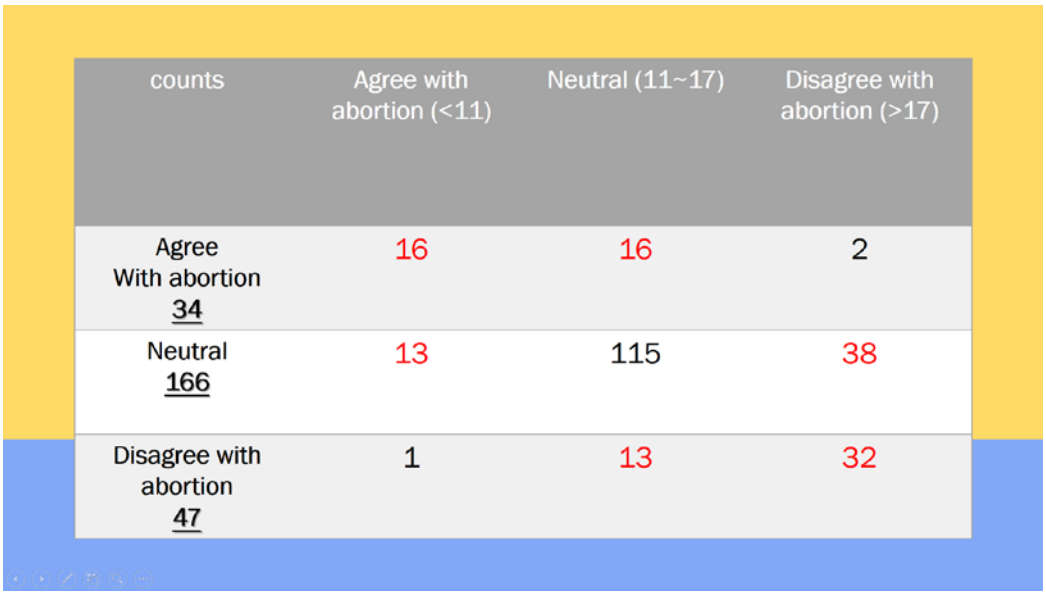
Pie Chart 3. It answers to the question, “Do you feel that the survey itself has a stance?”



Pie Chart 4. It answers to the question, “What do you think is the stance of the questions above?”



Pie Chart 5. This chart answers to the question, “Do you think the questionnaire attitude affect the respondents?”



counts	Agree with abortion (<11)	Neutral (11~17)	Disagree with abortion (>17)
Agree With abortion 34	16	16	2
Neutral 166	13	115	38
Disagree with abortion 47	1	13	32

Table 6. This is the results of how many respondents were affected by the observer-expectancy effect compared and tracked from the calculation counts and respondents' original stands.

Works Cited

Bakker, Bart F. M. "Estimating the Validity of Administrative Variables." *Statistica Neerlandica* 66.1 (2012): 8-17. Print.

Bakker, B.F.M. and P.J.H. Daas. “Methodological Challenges of Register-Based Research.” *Statistica Neerlandica* 66: (2012) 2-7. Web. http://dx.doi.org/10.1111/j.1467-9574.2011.00505.x.

Betsch, Cornelia, et al. “The Narrative Bias Revisited: What Drives the Biasing Influence of Narrative Information on Risk Perceptions?” *Society for Judgment & Decision Making* 10.3 (2015): 241-264. Print.

Burger, Joep, Arnout van Delden, and Sander Scholtus. “Sensitivity of Mixed-Source Statistics to Classification Errors.” *Journal of Official Statistics* 31.3 (2015):489-506. Print.

Carroll, Robert. “Experimenter Effect.” *The Skeptic’s Dictionary*. 1994. Web. http://skepdic.com/experimentereffect.html

Chegg. “*Observer-expectancy Effect.*” http://www.chegg.com/homework-help/definitions/observer-expectancy-effect-13

Menne-Lothmann, Claudia, et al. “How to Boost Positive Interpretations? A Meta-Analysis of the Effectiveness of Cognitive Bias Modification for Interpretation.” *PLoS ONE* 9.6 (2014): 1-26. Print.

Sriven, Michael. “Evaluation Bias and Its Control.” *U of California, Berkeley* June 1975. Web.

Toews, Ingrid, et al. "Extent, Awareness and Perception of Dissemination Bias in Qualitative Research: An Explorative Survey." *PLoS ONE* 11.8 (2016): 1-16. Print.

W, Justin. “Cognitive Bias Codex.” *DailyNous* 2016. web.

West, Brady T., Joseph W. Sakshaug, and Guy Alain S. Aurelien. "How Big of a Problem Is Analytic Error in Secondary Analyses of Survey Data?" *PLoS ONE* 11.6 (2016): 1-29. Print.