

The Effects of Questionnaire Frames on Indicators of Data Quality

Jihyang Choi

The School of Journalism

Indiana University

Choi20@indiana.edu

Abstract

This study aims to empirically test how different kinds of frames of survey questions –“informative” and “accepting consequences” frames– affect the quality of the survey data collected. The “don’t know” response rate and the level of opinion diversification were adopted as criteria to measure data quality.

The study reveals “informative” frames, in which background information about particular issues or policies is provided in the form of a preamble, did not increase the data quality significantly. “Accepting consequences” frames yielded contradictory results depending upon the level of the consequences involved. When the questions contain frames that facilitate the respondents making an evaluation about the personal-level consequences, the data quality significantly improved. However, the quality did not increase in the frames of the national-level consequences.

Introduction

Public opinion researchers have long struggled to find methods of minimizing measurement errors and, thus, maximizing data quality. However, suggested remedies to these problems have yielded conflicting consequences. This is the case because the quality of data itself is an extremely ambiguous, contradictory concept which is difficult to test empirically (Price & Neijens, 1997).

Studies have suggested a variety of criteria with which to evaluate the quality of survey outcomes: the stability or consistency of opinions (Converse, 1964; Nie, Versa, & Petrocik, 1976); the amount of information supporting respondents' opinions (Bishop, Oldendick, Tuchfarber, & Bennett, 1980; Lane & Sears, 1964); the level of conviction (Schuman & Presser, 1981); or, the plausibility of implementing the survey results (Key, 1961).

Researchers also have focused on discovering variables in the opinion forming process that affect the quality of the collected data. Two of the most frequently mentioned variables are the amount of information provided to respondents (Fishkin, 1991, 1995), and the respondents' capacity to understand the consequences of the issues involved (Yankelovich, 1991). Another group of studies has explored a more technical tool to increase the data quality by testing the effectiveness of different questionnaire designs: questionnaire length (Regula & Bachman, 1981; Yammarino, Skinner, & Childers, 1991); layout formats (Manfreda, Batagelj, & Vehovar, 2002; Norman, Friedman, Norman, & Stevenson, 2001; Sanches, 1992); or questionnaire structure (Hess et al., 2001; Schaefer & Dillman, 1998).

However, empirical studies of the effects of maximizing the amount of information provided and the level of emphasis on the consequences of the issues concerned in a questionnaire setting have been scarce. With that concern in mind, this study aims to add

knowledge to the existing studies of survey data quality by testing the prevailing assumptions, within the questionnaire frames, that more information and a greater understanding of the consequences of the issues tends to lead to data of higher quality. This study attempts to arrive at empirical evidence of how different kinds of frames –informative frames and “accepting consequences” frames– of questions affect the quality of the data gathered. We also hope to add more detailed observations concerning the effects of the respondents’ acceptance of consequences by separately comparing the outcomes of different levels of frames –specifically, the personal-level and national-level consequences. Moreover, the study also aims to offer a practical guide to how survey research can improve the quality of collected data by refining the questionnaire frames used.

Literature Review

Questionnaire design and data quality

Previous studies have shown that measurement errors can arise from questionnaire design.

Babbie (1979) asserts that, particularly in self-administered surveys, the questionnaire design needs to be less complex and burdensome than that used in face-to-face surveys.

Studies have revealed that different layout formats and other graphic cues yield detectable differences in data quality (Sanchez, 1992). Questionnaire structure and wording has also been investigated. The study of the effects of personal-level versus house-level questionnaire design has suggested that the use of a personal-level questionnaire slightly increases the data quality in terms of the completeness of the survey (Hess et al., 2001). The balance of the questions posed has also attracted researchers’ attention. However, studies of this issue have yielded contradictory results. Some studies found that fully balanced questions that state both competing points of view tend to yield a more balanced distribution of answers

(Narayan & Krosnick, 1996); however, Shaeffer, Krosnick, Langer, and Merkle (2005) assert that greater efficiency can be obtained through a minimally balanced questionnaire –that is, using questions that state only one viewpoint entirely and only briefly acknowledge the alternative viewpoint. Some other studies have pointed out the negative effects of question length (Crawford, Couper, & Lamias, 2001; Galesic & Bosnjak, 2009), or the different effects of open-ended and closed-ended questions on people’s response pattern (Kwak & Radler 2002; Schaefer & Dillman 1998).

Recently, questionnaire design has become a keen focus of interest, since web surveys are increasingly popularized, and anyone with access to the Internet can create a questionnaire and perform a survey (Couper, Traugitee, & Lamias, 2000). However, those survey questionnaires designed by people with no professional or academic training might possibly result in data of lower quality and, consequently, would not represent people’s opinions in a valid way.

Researchers find the reason for the larger errors which are found in web surveys in their respondents’ “satisficing” behavior. According to “satisficing theory,” some people refuse to conduct the considerable amount of cognitive work that is required in order to interpret the meaning of the survey questions asked, and also refuse to arrive at the best possible responses, thereby providing irrelevant answers (Krosnick, 1991). In web surveys, respondents are even more likely to demonstrate this sort of “satisficing” behavior, since most people in the Internet environment tend to be engaged in a variety of activities at the same time (e.g., *multitasking*; Holbrook, Green, & Krosnick, 2003). Empirical study of the difference in data quality between web surveys and face-to-face surveys reveals that web survey respondents are more likely to produce data of lower quality (Heerwegh & Loosveldt, 2008).

Public opinion quality

Previous studies suggest that quality data can be obtained only when people hold quality opinions. One common assertion found in the existing studies about what goes into the creation of opinions of good quality is that they can be produced by means of rational deliberation by informed citizens. Thus, supporters of this argument have emphasized the importance of providing the responding public with more information, appropriate education, or more opportunity for discussion of the issues involved in the questionnaire.

Therefore, in order to facilitate better-informed public opinion, a variety of methods have been created: deliberative polls through discussions (Fishkin 1991, 1995; Luskin, Fishkin, & Jowell, 2002); and educational polls which inform respondents about the background of complex public issues in an attempt to collect more knowledgeable opinions (Neijens 1987; Van Knippenberg & Daamen, 1996).

On the other hand, some researchers argue that being well-informed does not always guarantee a higher quality of survey opinions because mere information by itself does not lead people to arrive at solutions. Yankelovich (1991), as an alternative to this information trap, proposed public engagement or judgment as a criterion for quality. He suggested that quality public opinion can be acquired when the public accepts “responsibility for the consequences of its views” (p. 24). This idea is related to the argument which Fishbein put forward in his “attitude formation model,” which assumes that the more a person believes that a particular attitude will result in good consequences, the more favorable that person’s attitude will be (Fishbein, 1963).

Several survey projects have attempted to increase data quality by maximizing those two factors –information and acceptance of consequences– in survey questions. A series of

educational surveys by the Americans Talk Issues Foundation (ATIF) tested small variations in the wording of questions. For example, the survey provided a series of limiting conditions such as “how would you rate this proposal...if you know that...” or stated the possible consequences of their choice in questionnaires (Kay, Henderson, Steeper, & Lake, 1994). The Choice Questionnaire (Neijens, 1987; Neijens, de Ridder, & Saris, 1992) also attempted to improve the quality of public opinion by providing more background information on issues and policies, and also by encouraging respondents to carefully think about the possible consequences of their choices.

Operationalization of data quality

Previous studies of data quality have indicated that the operationalization of data quality is a very complex process, due to the diversity of views about the criteria used to assess data quality. Researchers have adopted a mixture of a variety of measurements in order to test the level of data quality: the “don’t know” rate and the nondifferentiation of response (Heerwegh & Loosveldt, 2008); missing data including the “don’t know” refusal (Couper, 1997); the level of consistency between attitudes within individuals’ belief systems (Sturgis, Roberts, & Allum, 2005); opinion consistency (Hess et al. 2001; Shaeffer et al., 2005); the response times, response rates and length of open answers (Galesic & Bosnjak, 2009); and the item nonresponse rate and the prevalence of longer open-ended responses (Kwak & Radler, 2002).

The item nonresponse rate is also frequently employed: when the average number of unanswered questions is small, the collected data is regarded to be of higher quality. Another operationalization is the response length. For example, Schaefer and Dillman (1998) regarded longer responses to open-ended questions to be an indicator of effective survey measurements.

However, among a variety of criteria that claim their own legitimacy, one of the most

frequently employed criteria for judging comparative survey quality is the “don’t know” rate. Previous studies suggest that the presence of a lot of "don't know" responses tends to interfere with researchers' reading of opinions accurately, since a "don't know" response is a very ambiguous choice which may occur for a number of different reasons. Some researchers argue that respondents should be allowed to provide "don't know" answers (especially in attitude questions) since they sometimes really have no opinion or have not thought about the issue (Converse, 1974). However, another group of researchers argue that a large number of "don't know" responses is problematic and may decrease the overall data quality since some respondents tend to provide a "don't know" answer as an easy way out even when they really do have their own opinion about the issue. Smith (1984) also noted that some people in the "don't know" response group may have a certain opinion, despite the fact that their position may not be firm or fully articulated. Gilljam and Granberg (1993), based on a survey on nuclear power in Sweden, probed the real attitudes of people who offered "don't know" responses and found empirical evidence of a "pseudo-nonattitudes" group who refrained from expressing themselves, even though they do have an opinion. Studies explain the reason for people's propensity to provide a "don't know" response with the lack of respondents' attention to questions or their desire to escape from having to do the cognitive work required in order to respond to the survey (Krosnick & Fabrigar, 2005).

Therefore, survey researchers have tried to find a means of minimizing the number of "don't know" responses and, consequently, increase the data quality. One common method is not to provide a "don't know" option unless this option has significant implications for the study. Another method is to train interviewers to probe all "don't know" responses in an attempt to induce substantive responses (Schuman & Presser, 1981).

In evaluating the quality of aggregate public opinion, researchers have suggested two

opposing criteria: stability and diversity. Some have argued that the level of consistency of public opinion is a reflection of the quality of the data obtained. Since the time of Converse's work (1964, 1970), the stability of opinion –which refers to the degree to which people's opinions remain consistent over time– has been regarded as an important criterion of data quality. However, studies of the evolving process of public opinion imply that stability may not be applied as an indicator of data quality in every phase of opinion development. This is the case since, as numerous studies have demonstrated, an individual's opinion is prone to change until it becomes final, carefully considered opinion as one is exposed to opposing ideas, information and arguments via media and interpersonal contacts.

Price and Neijens (1997) suggested that different criteria for judging quality might be employed, depending upon the phases of opinion formation. These researchers have conceived of the opinion forming processes in a serial fashion: first, properly eliciting the collective goals or values to the maximum level; then, developing a full range of options; then, examining the consequences; then, evaluating options; and, finally, making a sound choice. Similarly, the Public Agenda Foundation (2010), a public opinion research organization, sees the process of opinion evolution as a seven stage sequence: awareness; a sense of urgency; a search for solutions; reacting and resisting; wrestling with alternative choices; intellectual assent or resolution; and full resolution. Yankelovich (1991) also noted that the quality of public opinion would be formed by its passing through three different stages: consciousness raising; working through; and, finally, a resolution stage.

Based on these researchers' observations that public opinion evolves according to certain stages or phases, it can be inferred that the criteria used for evaluating data quality would vary according to the stage in the opinion forming process. As Rohrmann (1986) asserted, judging the quality of decision outcomes with a single, fixed criterion is problematic.

Specifically, when a certain issue is under vigorous and contested debate throughout the nation and, thus, the public is going through, for example, the process of the “evaluating options” or “reacting and resisting” stages, the stability or consistency of opinion would not be an appropriate criterion by which to measure data quality. At these stages, members of society may be encountered who have conflicts and/or disagreements over the policy solutions, since the heightened public discourse is not one-sided but rather two-sided or even multisided (Zaller, 1992). Therefore, it is natural for public opinion to fluctuate accordingly, and, rather than there being a consistency of public opinion, how opinions are diversified might better reflect how attentively people are engaged in the opinion forming process.

Previous studies also indicate that opinion diversification could be another important criterion of data collected. It has been argued that the existence of diverse opinions on one issue leads people in a society to have a more accurate perception of public opinion (O’Gorman, 1975; 1979). Moreover, Huckfeldt, Johnson, and Sprague (2004) also assert that the presence of disagreement and political heterogeneity within a society has a positive effect on the capacity of citizens to deal constructively with an issue.

Hypotheses

As mentioned above, the present study attempts to test whether more information and more acceptance of consequences do, indeed, lead to higher quality data. As a starting point, based on previous studies, the present research assumes that the amount of information and the emphasis on accepting consequences within a questionnaire setting would have certain effects on data quality. It is also assumed that the direction of the effects would be positive when the respondents are provided with more information or more cues that cause them to think about the possible consequences in survey questions. Thus, the following hypotheses are proposed:

H1: Data quality is significantly higher when survey questions adopt “informative” or “accepting consequences” frames than when questions adopt general frames.

H2: Data quality is significantly higher when survey questions adopt “informative frames” than when questions adopt general frames.

Given that the personal-level and household-level questionnaire design yield different data quality (Hess et al., 2001), the study also assumes that variations in the consequence level would also affect the survey quality, leading to the following hypotheses:

H3: Data quality is significantly higher when survey questions adopt “accepting personal consequences” frames than when questions adopt general frames.

H4: Data quality is significantly higher when survey questions adopt “accepting national consequences” frames than when questions adopt general frames.

Methods

In this study, the independent variable is the kinds of frames used in survey questions. The dependent variable is the data quality, as measured by two criteria: the “don’t know” response rates and the level of opinion diversification. Thus, three types of data were gathered for the study: frames of survey questions; the “don’t know” rates of respondents; and the level of opinion diversification.

Sampling and materials

The survey questions and their corresponding responses were obtained from the Gallup

Organization with the help of *Gallup Brain*, Gallup's survey database. Within this abundant survey data, this study chose the survey results from three topics (economy, healthcare reform, and the war on terrorism) over a three year period (October, 2007 to September, 2010). These data allow for the exploration of differences in the quality of the data collected. These three topics were chosen because they were the most actively debated issues among the public during that time period, and the three were consistently listed among the top ten problems mentioned in Gallup's "What is the most important problem facing this country today?" question from 2007 to 2010.¹

The search criteria were designed to select relevant questions and corresponding responses as precisely as possible. Sets of survey questions and responses were selected if they contained the following words: "economic situation," "financial situation," or "stimulus plan" for the economy topic; "healthcare" or "health policy" for healthcare reform; and "terror," "terrorism," or "Iraq" for the war on terrorism. Among the survey results, irrelevant results that were accidentally included were culled out. This process yielded a total of 308 sets of questions and responses.

Independent variables: the kinds of frames in survey questions

Each question was categorized into one of four categories of question frames: "informative" frames; "accepting personal-level consequences" frames; "accepting national-level consequences" frames; and general frames. The questions with "informative" frames refer to those that provide background information about issues in the form of a preamble in order to provide respondents with frames of reference.

For example, the following questions were categorized as adopting "informative" frames: "Next, as you may know, Congress and the president recently passed legislation to

provide up to \$700 billion of government assistance to address the problems being faced by U. S. financial institutions. All in all, do you think it is a good thing or a bad thing that the government passed this bill?” and “As you may know, Khalid Sheikh Mohammed has described himself as the mastermind of the September 11 terror attacks. He is scheduled to face trial in New York City. Just your opinion, do you think it would be better to have his trial in New York City or better to have it somewhere else?”

The “accepting personal-level consequences” frames refer to those that facilitate respondents’ making an evaluation about the possible impact of certain issues or policies on individuals or their families. For example, questions which fell under this category are those such as, “How do you think Obama’s economic stimulus plan would affect you and your family’s financial situation?”; or, “Now thinking more specifically about how the healthcare bill will affect you and your family; do you think the healthcare coverage you and your family receive will get better, will not change, or will get worse than it would if no healthcare bill passed?”

Likewise, frames of “accepting national-level consequences” refer to questions that ask respondents to evaluate the consequences that certain issues or policies would have on the U.S. Questions such as “...if the U.S. withdraws its troops from Iraq, would it be more likely to be attacked by terrorists...?”; or, “Taking everything into account, do you think in the long run that the new healthcare bill will make the U.S. healthcare system better, will not make much difference to it, or will make it worse?” were placed in this category. The "general frames" category applies to questions that did not adopt any of the three frames described above.

After the coding process, a total of 38 “informative” frames, 41 “accepting personal-level consequences” frames, and 37 “accepting national-level consequences” frames were all

grouped together. And, by default, the remaining 192 were categorized as "general frames." Among the 192 general frame questions, for the sake of comparison with the other groups, 41 questions were selected by using a systemic random sampling.

Coder reliability

Forty eight randomly selected questions, 12 from "informative" frames, 12 from "personal consequences acceptance" frames, 12 from "national consequences acceptance" frames, and 12 from the "general frames," were coded by an independent coder, in order to provide a coder reliability check. The coefficient as calculated using Scott's π , was 0.83.

Dependent variables

Data quality was measured by two different kinds of criteria: the "don't know" response rate and the level of opinion diversification. This study adopted the "don't know" response rate as one of two criteria for judging the quality of data for several reasons. First, it is one of the most conventionally used criteria of comparative survey quality. Krosnick and Fabrigar (1997) pointed out that some respondents may choose the "don't know" category as an easy way out. Second, given that the survey questions concerned in this study were regarding issues that were among those at the top of the national agenda during the time periods, and also given that all respondents might have been exposed to information about those issues from media and also from interpersonal communication, the study assumes that the "don't know" responses from participants are more likely to reflect the satisficing behavior of respondents rather than a true nonexistence of opinion.

Another criterion for judging the quality of survey data is the level of opinion diversification. Previous studies of the phases of opinion evolution imply that the public is

more likely to be engaged in the process of “evaluation of options” or “reacting and resisting” or “working through,” since the three issues concerned (economy, war, health care) were the prominent issues in the political and public arenas (Price & Neijens, 1997; Yankelovich, 1991). Moreover, studies also suggest that respondents tend to be exposed to conflicts and disagreements when a certain issue is under extensive debate. Thus, this study posits that a more diverse level of public opinion would more closely reflect how extensively, seriously and deliberately people are engaged in the discussion of those topics.

In sum, in this study, a lower “don’t know” rate and a higher level of opinion diversification indicate data of higher quality. The “don’t know” response rate simply refers to the percentage of respondents who chose the “don’t know” response category in surveys. The level of opinion diversity refers to how diverse the opinion is. To measure the level of diversity, this study adopted the Herfindahl Index as a main measurement. The Herfindahl Index, which is extensively used in Economics to indicate the level of market concentration, was used in order to determine how concentrated or diversified the respondents’ attitudes are. It is calculated simply by taking the sum of the squares of the percentile shares of each of the categories of responses.

For example, suppose the answer categories of a survey is divided into “a great deal,” “a fair amount,” “only a little,” and “not at all,” and that the percentile distribution of each category is:

- “a great deal” 25%
- “a fair amount” 10%
- “only a little” 5%
- “not at all” 60%

In this example, the Herfindahl Index would be 4350, or

$$\because 25^2 + 10^2 + 5^2 + 60^2 = 4,350$$

The Index figures range from 0 to 10,000; a high Index figure indicates that the market is excessively monopolized, while a low Index figure means that the market is diversified. Thus, by analogy, the study interprets a high Index figure to be an indication that public opinion is excessively uniform; while a low Index figure means that public opinion is diversified. In other words, the lowest possible Index figure (0) tells us that the respondents' evaluation or attitude about an issue is extremely scattered; the highest possible Index figure (10,000) indicates that public opinion has reached a state of absolute consensus, because the Index figure 10,000 means that a single response category has 100% of the share of responses. In sum, the lower the Index figure is, the higher the level of diversification.

The Herfindahl Index has rarely been used in the area of communications, though it has been used in some studies dealing with the media industry (Chan-Olmsted, 1991; Stuhlfaut, 2005), and in studies on the content diversity of news (Entman, 2006; Choi, 2009).

Research design

Data were analyzed using a *t*-test and an analysis of variance (ANOVA). The *t*-test was used to test differences in data quality between a group of “informative” or “accepting consequences” frames and the general frames, which is the concern of Hypothesis 1. The ANOVA was used to determine whether significant differences, in terms of data quality, would exist among the four groups. To answer Hypotheses 2 to 4, the Scheffe post hoc test was used. The study employed the Scheffe method since it always tends to be conservative as

the distribution becomes heavier in the tails, and, thus, to act as a control on the Type 1 error rate (Ringland, 1983).

Findings

Hypothesis 1 predicted that data quality would be significantly higher when survey questions adopt “informative” or “accepting consequences” frames than when questions adopt general frames.

The mean of the “don’t know” response rate in general frames ($M = 5.22$ $SD = 7.58$) was higher than that of the three other kinds of frames ($M = 2.70$ $SD = 3.81$); and the level of diversification was higher in the three other kinds of frames ($M = 3931.52$ $SD = 1192.01$) compared to the general frames ($M = 4528.50$ $SD = 1010.18$). *T*-tests revealed that differences in the “don’t know” rate ($p < 0.01$) and the level of opinion diversification ($p < 0.01$) are statistically significant, thereby supporting H1. The data quality was significantly higher when survey questions adopt “informative” or “accepting consequences” frames than when questions adopt general frames. The differences are presented in Table 1.

Table 1 about here

The ANOVA revealed that significant differences exist in the “don’t know” response rate ($F[3, 153] = 5.356$, $p < 0.01$) and also in the level of opinion diversification ($F[3, 153] = 5.356$, $p < 0.01$) depending on the use of different kinds of frames. The group means of data quality are presented in Table 2.

Table 2 about here

To answer H2 through H4, Scheffe post-hoc test was performed. Hypothesis 2 predicted that the data quality would be significantly higher when survey questions adopt “informative” frames than when they adopt general frames. The mean of the “don’t know” response rate of “informative” frames ($M = 4.53$, $SD = 5.99$) was lower than that of the general frame group ($M = 5.22$, $SD = 7.58$); and the level of diversification was higher in the “informative” frames ($M = 4210.27$, $SD = 1156.97$) than in the general frames ($M = 4528.50$, $SD = 1010.18$). However, contrary to the study’s assumption, the Scheffe test revealed that the differences are not statistically significant (the “don’t know” response, $p = 0.945$; the level of opinion diversification, $p = 0.671$), and, thus, H2 was rejected.

Hypothesis 3 proposed that data quality would be significantly higher when survey questions adopt accepting personal-level consequences frames than when the questions adopt general frames. The mean of the “don’t know” response rate of “accepting personal-level consequences” frames ($M = 1.42$, $SD = 1.19$) is much lower than that of general frames. The level of opinion diversification was also higher in “consequences acceptance” frames ($M = 3610.28$, $SD = 1058.94$) than that in the general frames. The Scheffe test also revealed that differences in the “don’t know” response rate ($p < 0.01$) and the level of diversification ($p < 0.01$) are statistically significant. Thus, H3 was supported: data quality was significantly higher when survey questions adopt “accepting personal consequences” frames than when questions adopt general frames.

The focus of Hypothesis 4 was in revealing the differences in data quality between “accepting national-level consequences” frames and general frames. The mean of the “don’t know” response rate in the “national” frames ($M = 2.24$, $SD = 1.52$) is lower than that of the general frames; and the level of diversification was higher in the “accepting national-level

consequences” frames ($M = 4001.19$, $SD = 1308.70$) than that of the general frames.

However, the differences in the “don’t know” response ($p = 0.075$) and the level of opinion diversification ($p = 0.245$) were not statistically significant, which means that H4 was rejected. Results from the testing of Hypotheses 2 to 4 are presented in Table 3.

Table 3 about here

Discussion and Conclusions

The objective of this study was to empirically test whether the information given and people’s acceptance of consequences indeed affect the quality of the data collected. The findings summarized in the previous section imply that, at least in the questionnaire setting, those two factors are not defining aspects that lead to higher quality data, contrary to the previous studies’ extensive emphasis on them as a prescription for data quality. Providing more background information about low salient issues in the form of a preamble in questions did not significantly increase the data quality. In terms of understanding consequences, the data quality was significantly enhanced only when the respondents were provided with cues to think about the consequences of an issue or policy on a personal-level; no significant increase in the data quality was found when the questions induced respondents to consider the consequences on a national-level.

Specifically, the findings about the minimal impact of information again imply that the role of information in the process of obtaining data quality is not straightforward, but is rather complex. Even though researchers have made efforts to inform or educate respondents of surveys through extensive programs, such as educational polls or deliberative polls in the hopes of gathering better opinion data, the findings of the present study suggest that an

injection of information is not a panacea for obtaining data of higher quality.

This result leads us to raise a question about the distinction between information that is useful or relevant in the decision making process and information that is not. Despite several ambitious survey projects which have tested the effects of information on data quality, the information so tested was not background information or objective knowledge, as it was in this study. Rather, in most of cases, the information was about the different social or political consequences of particular policies or issues. Thus, the information was about a series of subjective evaluations of an issue. For example, the Americans Talk Issue Foundation's educational survey (Kay et al., 1994) offered information to participants in the form of arguments which favor or oppose particular policies. In a similar way, the Choice Questionnaire (Neijens, 1987) tested the differences between situations in which respondents are provided with information about the consequences of certain issues and those in which no information is given. Thus, in those projects, it is not clear that the changes in the data quality derive from the information itself or from the persuasive power of the sources (Price & Neijens, 1998). Therefore, more detailed studies about how different kinds of information affect the quality of data, are required to clarify the role of information.

The present study also provides a valuable observation about the function of "accepting consequences" frames. The finding shows that the effects of those frames differ depending on the level of consequences; higher quality data was obtained only when the frames were dealing with the personal-level consequences. This result implies that participants tend to take the consequences of issues and policies seriously only when the questions are related to their own personal interests. Moreover, conversely, the findings also imply how complicated a task it is to make people more engaged in national issues; which, in turn, is related to Yankelovich's argument that people simply do not try to accept the

consequences of their opinions on some issues (1991). He further explains the possible reasons why people oftentimes do not evaluate the consequences of particular issues, while paying greater attention to the consequences of other issues: people may not confront the consequences of their opinions when they lack awareness; when they have no time; or when they have difficulty in resolving their own conflicting values and ambivalence (pp. 29-30).

This study is also meaningful in that it tested the applicability of a new method to measure the quality of survey outcomes: i.e., the level of opinion diversification. The fact that the evaluation of data quality using a new criterion –the level of opinion diversification– is consistent with the evaluation based on a rather conventional criterion –the “don’t know” response rate– supports the validity of the new criterion. Price and Neijens (1997, quoting Foote and Hart’s 1953 work), argued that, when the public is evaluating the options of a certain policy, the advocates of each competing policy seek to promote their causes using persuasive appeals, and that intense public discourse takes place (pp. 340). This is when people develop more diverse opinions on current issues. However, it is also worth mentioning that more systematic and in-depth studies are needed to clarify the relationship between opinion diversification and data quality.

Although this study contains valuable implications, it also has several limitations. First, it does not take into consideration broad kinds of topics. The targets of analysis were limited to the questions and corresponding responses to three topics over a three year period beginning in October of 2007. Despite the fact that the study attempted to investigate public opinion of issues that were under vigorous public discourse, the limited sample used hinders the study from arriving at a broader understanding about the relationship between questionnaire frames and data quality. Moreover, it is also worth noting that individual’s opinions on the three topics of concern in the study could possibly be influenced by their

ideological orientation, thus making the effects of the questionnaire's frames more subtle and complex. Therefore, further studies are needed to clarify the extent of influence that the questionnaire's frames have.

Another limitation regarding the criteria used to judge data quality is also worth mentioning. Based on the studies of the phases of the evolution of public opinion, the study proposed that the level of opinion diversification can be used as a criterion of data quality when the issues are under intense debate among the public. However, it is well to reiterate that future studies are needed in order to find out the fluctuation pattern of the level of opinion diversification depending on the phases of opinion evolution and, possibly, to refine the scope of the criterion's applicability.

Despite these limitations, the significance of the study lies in its contribution towards a better understanding of data quality in far more systemic ways. Based on the findings of the study that revealed the rather delicate role which information and accepting consequences have in the acquisition of survey data quality, this study can be applied in practical ways by suggesting that survey organizers pay more attention to the frames employed in their questionnaires. It also has valuable implications for researchers of public opinion. The results can serve as an opportunity for researchers to think about the ways to increase the data quality of surveys in a more systematic way.

Table 1. Comparison between groups

| | General Frames (N = 41) | “Informative” or “consequences” frames (N = 116) |
|---|--|---|
| Don’t know response | 5.22 (SD = 7.58) | 2.70 (SD = 3.81) |
| Level of opinion diversification | 4528.50 (SD = 1010.18) | 3931.52 (SD = 1192.01) |

Note: The mean differences are significant at the 0.01 level.

Table 2. Group means of data quality

| | “Informative” (N = 38) | “Personal -level consequences” (N = 41) | “National -level consequences” (N = 37) | General (N = 41) |
|---|-----------------------------------|--|--|-----------------------------|
| Don’t know response | 4.53 (SD = 5.99) | 1.42** (SD = 1.19) | 2.24 (SD = 1.52) | 5.22 (SD = 7.58) |
| Level of opinion diversification | 4210.27 (SD = 1156.97) | 3610.28** (SD = 1058.94) | 4001.19 (SD = 1308.70) | 4528.50 (SD = 1010.18) |

*Note: **. The mean difference is significant at the 0.01 level.*

Table 3. Mean differences from the general frames

| | “Informative” (N = 38) | “Personal -level Consequences” (N = 41) | “National -level consequences” (N = 37) |
|---|---------------------------|--|--|
| Don’t know Response | 0.69 | 3.80** | 2.98 |
| Level of opinion diversification | 318.23 | 918.22** | 527.31 |

Note: **. The mean difference is significant at the 0.01 level.

¹The Gallup Organization regularly carries our surveys asking people the question, “What do you think is the most important problem facing this country today?” Despite small fluctuations, the economy, healthcare, and the war on terrorism have consistently been listed high in the list of problems between 2007 and 2010. In the September, 2008 survey, the economy (58%), the war in Iraq (13%), and healthcare (9%) were the top three problems that the U.S. public considered to be the most important. In August of 2009, the economy (38%) and healthcare (25%) still remained as the top two problems, while people who responded that the war/situation in Iraq (5%) is the most important problem decreased.

References

- Babbies, E. R. (1979). *The Practice of Social Research* (2nd ed.). Belmont, CA: Wadsworth.
- Bishop, G. F., Oldendick R. W., Tuchfarber A. J., & Bennett, S.E. (1980). Pseudo-opinions on public affairs. *Public Opinion Quarterly*, 44, 198-209.
- Chan-Olmsted, S. (1991). A structural analysis of market competition in the US TV syndication industry, 1981-1990. *Journal of Media Economics*, 4(3), 9-28.
- Choi, J. (2009). Diversity in foreign news in US newspapers before and after the invasion of Iraq. *International Communication Gazette*, 71(6), 525-42.
- Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and Discontent* (pp. 206-261). New York: Free Press.
- Converse, P. E. (1970). Attitudes and non-attitudes: Continuation of a dialogue. In E. R. Tufte (ed.), *The Quantitative Analysis of Social Problems* (pp. 168-89). MA: Addison-Wesley.
- Converse, P. E. (1974). Nonattitudes and American public opinion: Comment. The status of nonattitudes. *American Political Science Review*, 68, 650-66.
- Couper, M. P. (1997). Survey instructions and data quality. *Public Opinion Quarterly*, 62(2), 317-38.

- Couper, M. P., Traugitee, M., & Lamias, M. (2000). Experiments on the design of web surveys. Paper presented at the fifth *International Conference on Social Science Methodology*, Cologne, Germany, October 3-6, 2000.
- Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web surveys: perceptions of burden. *Social Science Computer Review*, 19, 146-62.
- Entman, R. M. (2006). Punctuating the homogeneity of institutionalized news: Abusing prisoners at *Abu Ghraib* versus Killing Civilians at *Fallujah*. *Political Communication*, 23(2), 215-24.
- Fishbein, M. (1963). An investigation of the relationship between beliefs about an object and the attitude toward that object. *Human Relations*, 16, 223-40.
- Fishkin, J. S. (1991). *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven: Yale University Press.
- Fishkin, J. S. (1995). *The Voice of the People: Public Opinion and Democracy*. New Haven, CT: Yale University Press.
- Foot, E., & Hart, C. W. (1953). Public opinion and collective behavior. In M. Sherif and M.O. Wilson (Eds.), *Group Relations at the Crossroads* (pp. 308-331). New York: Harper & Bros.
- Galesic M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicator of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349-60.
- Gilljam, M., & Granberg, D. (1993). Should we take don't know for an answer? *Public Opinion Quarterly*, 57(3), 348-57.
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population. *Public Opinion Quarterly*, 72(5), 836-46.
- Hess, J., Moore, J., Pascale, J., Rothgeb, J., Keeley A., & U.S. Census Bureau (2001). The effects of personal-level versus house-level questionnaire design on survey estimates and data quality. *Public Opinion Quarterly*, 65, 574-84.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67, 79-125.
- Huckfeldt, R., Johnson, P. E., & Sprague, J. (2004). *Political Disagreement: The Survival of Diverse Opinions Within Communication Networks*. Cambridge: Cambridge University Press.
- Kay, A. F., Henderson, H., Steeper, F., & Lake, C. (1994). *Interviews with the Public Guide Us... On the Road to Consensus*. St. Augustine, Florida: Americans Talk Issues Foundation.
- Key, V. O., Jr. (1961). *Public Opinion and American Democracy*. New York: Knopf.
- Krosnick, J. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys, *Applied Cognitive Psychology*, 5, 213-36
- Krosnick, J., & Fabrigar, L. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. P. Biemer, M. Collins, E. de Leeuw, C. S. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality*. New York: Wiley.
- Krosnick, J., & Fabrigar, L. (2005). *Questionnaire Design for Attitude Measurement in Social and Psychological Research*. New York: Oxford University Press.
- Kwak, N., & Radler, B. (2002). A comparison between mail and web surveys: response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18(2),

257-273.

- Lane, R. E., & Sears, D.O. (1964). *Public Opinion*. Englewood Cliffs, NJ: Prentice-Hall.
- Luskin, R. Fishkin, J., & Jowell, R. (2002). Considered opinions: deliberative polling in Britain. *British Journal of Political Science*, 32(3), 455-87.
- Manfreda, K.L. Batagelj, Z., & Vehovar, V. (2002). Design of web survey questionnaires: Three basic experiments. *Journal of Computer-mediated Communication*, 7(3). Retrieved December 1, 2010, from <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2002.tb00149.x/full>
- Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly*, 60, 58-88.
- Neijens, P. (1987). *The Choice Questionnaire. Design and Evaluation of an Instrument for Collecting Informed Opinion of a Population*. Amsterdam: Free University Press.
- Neijens, P., de Ridder, J. A., & Saris, W. E. (1992). An instrument for collecting informed opinions. *Quality and Quantity*, 26, 245-58.
- Nie, N. H., Verba, S., & Petrocik, J. R. (1976). *The Changing American Voter*. Cambridge, MA: Harvard University Press.
- Norman, K. L., Friedman, Z., Norman K., & Stevenson, R. (2001). Navigational issues in the design of online self-administered questionnaires. *Behaviour and Information Technology*, 20(1), 37-45.
- O'Gorman, H. (1975). Pluralistic ignorance and white estimates of white support for racial segregation. *Public Opinion Quarterly*, 39(3), 313-30.
- O'Gorman, H. (1979). White and Black perceptions of racial values. *Public Opinion Quarterly*, 43, 48-59.
- Price, V., & Neijens, P. (1997). Opinion quality in public opinion Research. *International Journal of Public Opinion Research*, 9(4), 336-360
- Price, V., & Neijens, P. (1998). Deliberative polls: toward improved measures of "informed" public opinion? *International Journal of Public Opinion Research*, 10(2), 145-76.
- Public Agenda Foundation (2010, May). The Seven Stages of Public Opinion. Retrieved December 1, 2010, from <http://www.publicagenda.org/pages/seven-stages-public-opinion>
- Regula, H. A., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly*, 45, 549-59.
- Ringland, J. T. (1983). Robust multiple comparisons. *Journal of the American Statistical Association*, 78(381). 145-151.
- Rohrmann, B. (1986). Evaluating the usefulness of decision aids: A methodological perspective. In B. Brehmer, H. Jungermann, P. Lourens, & G. Sevón (Eds.), *New Directions in Research on Decision Making* (pp.363-81). Amsterdam: North-Holland.
- Sanches, M. E. (1992). Effects of questionnaire design on the quality of survey data. *Public Opinion Quarterly*, 56, 206-217.
- Schaefer, D. R., & Dillman, D. A. (1998). Development of a standard e-Mail methodology: Results from an experiment. *Public Opinion Quarterly*, 62, 378-397.
- Shaeffer, E. M., Krosnick, J. A., Langer, G. E., & Merkle, D. M. (2005). Comparing the quality of data obtained by minimally balanced and fully balanced attitude questions. *Public Opinion Quarterly*, 69(3), 417-428.
- Schuman, H., & Presser, S. (1981). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York, Academic Press.
- Smith, T. (1984). Nonattitudes: A review and evaluation. In C. Turner and E. Martin (Eds.)

- Surveying Subjective Phenomena* (pp. 215-55). New York: Russell Sage.
- Stuhlfaut, M. W. (2005). Economic concentration in agricultural magazine publishing: 1993-2002. *Journal of Media Economics*, 18(1), 21-33.
- Sturgis P., Roberts C., & Allum, N. (2005). A different take on the deliberative poll. *Public Opinion Quarterly*, 69(1), 30-65.
- Van Knippenberg, D., & Daamen, D. (1996). Providing information in public opinion surveys: Motivation and ability effects in the information-and-choice questionnaire. *International Journal of Public Opinion Research*, 8, 70-82.
- Yammarino, F. J., Skinner, S. J., & Childers T. L. (1991). Understanding mail survey response behavior: a meta analysis. *Public Opinion Quarterly*, 55, 613-39.
- Yankelovich, D. (1991) *Coming to Public Judgment: Making Democracy Work in Complex World*. Syracuse, NY: Syracuse University Press.
- Zaller, J. R. (1992). *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.