



Data collection mode effect on feeling thermometer questions: A comparison of face-to-face and Web surveys



Mingnan Liu ^{a,*,1}, Yichen Wang ^{b,2}

^a SurveyMonkey, 101 Lytton Avenue, Palo Alto, CA 94301, USA

^b NERA Economic Consulting, 4 Embarcadero Center, Suite 400, San Francisco, CA 94111, USA

ARTICLE INFO

Article history:

Keywords:

Fact-to-face interview
Web survey
Feeling thermometer
Data collection mode effect

ABSTRACT

Feeling thermometer questions are widely used in political science research to estimate people's attitudes and feelings toward a political object, like a political figure or an organization. Given the popularity of the feeling thermometer question in population surveys, more work is needed to explore the measurement of this question type. This study examines the data collection mode effect on feeling thermometers. Using the 2012 American National Election Studies, we find that the measurement of feeling thermometers is not exactly comparable between face-to-face and Web surveys. Face-to-face respondents tend to provide warmer feelings, while Web respondents give relatively more reliable responses in comparison. In both survey modes, respondents are most likely to select the response options that are verbally labeled although the effect is more striking in face-to-face than Web survey. The item nonresponse between these two modes does not differ in a meaningful way. This study ends by discussing future research directions on feeling thermometer questions.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Feeling thermometer questions are frequently used in national surveys and political polls to measure the general population's attitudes and feelings toward a political party, political figure, or some groups and organizations (Nelson, 2008). It uses a 101-point rating scale where 0 indicates very cold and unfavorable feeling while 100 indicates very warm and favorable feeling. The popularity of this type of question is primarily due to the belief that one's feelings toward a certain object, such as a political party, can affect one's subsequent behaviors, like voting (Buell & Sigelman, 1985; Greene, 1999, 2004; Kaid, Leland, & Whitney, 1992; Lauderdale, 2010; McAdams & Johannes, 1988). In the literature, there are some empirical studies that attempt to examine the measurement of 101-point feeling thermometer questions. In the study that examined the feeling thermometer questions from the 1972 American National Election Studies (ANES), Wilcox, Sigelman, and Cook (1989) find that respondents provide systematically different

ratings to the feeling thermometers based not only on their actual evaluations of various social groups but also on other factors. For example, women, non-white, and less-educated respondents tend to systematically rate all social and political groups in the survey to be warmer than men, white, and more educated respondents. In another study, Smith (1987) shows that the choice of question wording could alter the responses to feeling thermometer questions. Specifically, he found that respondents rated "people on welfare" to be cooler or more unfavorable than "poor people." In an attempt to identify the cause for differential stability of symbolic and nonsymbolic political attitudes, Krosnick (1991b) finds that the different measurement reliability between two different question formats may have contributed to the lower stability of nonsymbolic than symbolic political attitudes. He reports that in the 1970 National Election Studies the 7-point fully verbally labeled rating scale produces higher test-retest reliability than the 101-point feeling thermometer questions. The study by Green (1988) finds that the nonrandom measurement error associated with feeling thermometer questions is higher than that of dichotomous like/dislike scale.

Given the popularity of the feeling thermometer question in population surveys, we believe that more work is needed to explore the measurement of this question type. One of the overlooked research areas is the mode effect on feeling thermometers. The purpose of this study is to understand whether and how

* Corresponding author. Tel.: +1 765 404 6621.

E-mail addresses: mingnanli@surveymonkey.com (M. Liu), yichen.wang@nera.com (Y. Wang).

¹ The paper was finished when he was a Ph.D. candidate in the Michigan Program in Survey Methodology, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA.

² Tel.: +1 734 604 9366.

people respond differently to the feeling thermometer questions in face-to-face versus Web probability surveys. As some flagship national surveys, such as ANES, are moving from face-to-face survey toward Web data collection, a better understanding of the mode effect of feeling thermometer questions is crucial for future academic research.

We are aware of only two studies that examine the mode effects on feeling thermometer questions. One study compares the reliability of two feeling thermometer questions (feelings toward Bush and Gore) among three data collection modes, namely telephone survey, probability Web survey and non-probability Web survey (Chang & Krosnick, 2009). The authors conclude that telephone produce the least reliable estimates between two waves of data collection while non-probability Web survey produces the most reliable estimates. The reliability of probability Web survey falls in between. In another study, Malhotra and Krosnick (2007) compare the means of three feeling thermometers (feelings toward Clinton, Gore and Bush) between a probability face-to-face survey and a nonprobability Web survey and find that the Web survey elicits warmer feelings for first two questions while face-to-face elicits warmer feelings for the last question. The drawback of this study, as the authors acknowledge, is that the observed difference can be contributed to mode effect, sampling difference, or a combination of both.

To date, there have been some research efforts spent on examining and comparing the mode effect between face-to-face and Web surveys (for example, see Liu & Wang, 2014). Our review of the literature shows that previous studies have researched three general conclusions. First, face-to-face surveys tend to have higher response rates than Web surveys (Christensen, Ekholm, Glümer, & Juel, 2014; Heerwegh & Loosveldt, 2008). This is largely due to the higher level of interviewer contact in the former than the latter mode. Interviewers can introduce the survey, explain its importance, and address respondents' concerns, all of which are not possible in the self-administered Web survey. Second, face-to-face surveys usually have a higher level of social desirability bias than Web surveys (Heerwegh, 2009). When facing a human interviewer in a face-to-face interview, respondents have a higher tendency to misrepresent their true opinions well by under-reporting undesirable attitudes and over-reporting those that seem more acceptable by the interviewers in order to avoid potential judgments and tensions. Third, data quality, which is measured by satisficing indicators such as item nonresponse rate and non-differentiation, is higher in face-to-face surveys than Web surveys (Goldenbeld & de Craen, 2013; Heerwegh, 2009; Heerwegh & Loosveldt, 2008). This is once again due to guidance and motivation given to the respondents from the interviewers in face-to-face surveys.

What do we expect the mode effects on feeling thermometer questions to be? First, providing a favorable rating is usually seen as a more desirable response than a negative rating. Through a warmer rating, face-to-face respondents can avoid creating a potential negative self-image resulting from negative responses (Couch & Keniston, 1960; Leary & Kowalski, 1990). In contrast, in the self-administered Web survey, respondents are less pressured to present themselves positively by providing more favorable responses due to the absence of interviewers.

Hypothesis 1. The face-to-face survey elicits more favorable or warmer feelings than a Web survey.

Second, respondents in face-to-face and Web surveys are likely to show different patterns of rounding (or response heaping) of their answers to the integer of 5 or 10. Rounding is usually the result of estimation (Burton & Blair, 1991). It eases the task of answering survey questions as respondents only need to come

up with an approximate number rather than a precise one (Schaeffer & Presser, 2003; Tourangeau, Rips, & Rasinski, 2000). If we treat rounding as an indicator of satisficing behavior, then we would expect more such behaviors in Web surveys as respondents are more likely to take cognitive shortcuts in a self-administered mode (Heerwegh & Loosveldt, 2008). The non-verbal communication between interviewer and respondent in face-to-face surveys motivates the respondents to give a more accurate and proper response to the question (Holbrook, Green, & Krosnick, 2003).

Hypothesis 2a. Web respondents are more likely to give rounded answers than face-to-face respondents.

On the other hand, the higher time pressure in face-to-face surveys may force the respondents to go through an incomplete anchoring and adjustment process, in which the heaping points serve as the anchoring points. If the face-to-face respondents do not have enough opportunity to make a full adjustment, there will be more rounded answers in the face-to-face than Web survey.

Hypothesis 2b. Face-to-face respondents are more likely to give rounded answers than Web respondents.

Third, we assess the item nonresponse rate in the two survey modes. The satisficing theory (Krosnick, 1991a) predicts more item nonresponse in the lack of motivation condition, such as in self-administered surveys. The presence of an interviewer in the face-to-face interview can motivate the respondents to give a more thoughtful and substantive answer and less item nonresponse in comparison to Web respondents.

Hypothesis 3. Web survey elicits more item nonresponse than face-to-face survey.

Fourth, the reliability of the responses to the feeling thermometer questions can differ by modes. As will be introduced in the next section, the data we use come from a two-wave panel study. Four of the thermometer questions were asked in both the first interview and the re-interview, which allows us to compare the reliability of the responses across two waves between these two modes. We are not aware of any study that compares the reliability of responses in face-to-face and Web surveys. However, the study by Chang and Krosnick (2009) suggests that the interviewer-administered telephone interview produce less reliable responses than self-administered Web survey. The authors do not offer a clear explanation on why the amount of random measurement error is higher in telephone survey than Web survey. We think that the face-to-face survey resemble the telephone survey on several aspects, including interviewer involvement and channel of communication, and locus of control, all of which may affect the level of reliability of responses (Couper, 2011). Therefore, similar to the telephone versus Web survey contrast on response reliability, we think the reliability of face-to-face survey is also likely to be poorer than the Web survey.

Hypothesis 4. The response reliability for the feeling thermometer questions is higher in Web survey than face-to-face survey.

2. Data and measures

This study utilizes the 2012 ANES survey data. The ANES conducted the surveys in two modes, namely computer-assisted interviewer-administered face-to-face and Web, using two independent national representative samples and one identical questionnaire. The target population for the survey is U.S. citizens aged 18 or old-

Download English Version:

<https://daneshyari.com/en/article/6838329>

Download Persian Version:

<https://daneshyari.com/article/6838329>

[Daneshyari.com](https://daneshyari.com)