Testing the Validity of the Crosswise Model: A Study on Attitudes Towards Muslims

David Johann, German Centre for Higher Education Research and Science Studies, Berlin
Kathrin Thomas, City, University of London

12.05.2017

How to cite this article: Johann D. & Thomas K. (2017). Testing the Validity of the Crosswise Model: A Study on Attitudes Towards Muslims

Abstract

This paper investigates the concurrent validity of the Crosswise Model when “high incidence behaviour” is concerned by looking at respondents’ self-reported attitudes towards Muslims. We analyse the concurrent validity by comparing the performance of the Crosswise Model to a Direct Question format. The Crosswise Model was designed to ensure anonymity and confidentiality in order to reduce Social Desirability Bias induced by the tendency of survey respondents to present themselves in a favourable light. The article suggests that measures obtained using either question format are fairly similar. However, when estimating models and comparing the impact of common predictors of negative attitudes towards Muslims, some puzzling results are revealed raising concerns about the validity of the Crosswise Model.

Keywords

attitudes towards Muslims, crosswise model, validity

Acknowledgement

This work was supported by the Austrian National Election Study (AUTNES), a National Research Network (NFN) sponsored by the Austrian Science Fund (FWF) [S10902-G11]. We thank Ben Jann for his kind support and the anonymous reviewers for their constructive feedback.

Copyright

© the authors 2017. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)
Testing the Validity of the Crosswise Model: A Study on Attitud...

Introduction

Social Desirability Bias (SDB) induced by other-deception – i.e., respondents purposely misrepresenting the truth as a form of impression management motivated by the desire to avoid evaluation by a third party, such as survey interviewers (Goffman, 1959; Nederhof, 1985) – poses a threat to the quality of survey responses. Early research on socially desirable response patterns has indicated that, even when factual questions are concerned, researchers need to be cautious about the validity of respondents’ answers. It is suggested that the range of invalidity of responses to more sensitive items is “sufficient to cause worry, and indicates a great need for further research on the truthfulness of respondents’ statements” (Parry, & Crossley, 1950; see also Bradburn, & Sudman, 1979). In addition to this, Cannell, Oksenberg, and Converse (1977) and Cannell, Oksenberg (1981) further distinguished the effects of positive and negative distortion in survey responses, linking positive bias to socially desirable and negative bias to socially undesirable behaviour. As such, there has been demand for the improvement of old and development of new techniques to enable survey respondents’ to more truthfully answer survey questions and, thus, to enhance the overall quality of survey data.

So-called Randomized Response Techniques (RRTs) have been developed to improve the way in which respondents’ anonymity and confidentiality in reporting sensitive attitudes or behaviour in surveys is preserved and to reduce bias due to socially desirable response behaviour (Warner, 1965; Horvitz et al., 1968; Greenberg et al., 1969; Boruch, 1971; Bradburn, & Sudman, 1979; Locander, Sudman, & Bradburd, 1976; Kuk, 1990; Daniel, 1993; see also Lensvelt-Mulders, Hox, van der Heijden, & Maas, 2005; Krumpal, Jann, Autispurg, & von Hermann, 2015; Fox, 2015; Chaudhuri, 2016). Typically, RRTs rely on randomization devices, which obscure the meaning of an answer communicated to the interviewer and then allow estimating the prevalence of the sensitive characteristic using elementary probability theory (Krumpal et al., 2015). However, while successfully reducing SDB, RRTs may induce bias by design, for example, by providing complex or unusual instructions which immediately alert respondents (Ulrich, Schröter, Streegl, & Simon, 2012) and may lead to instant self-protective ‘no’-answers (Krumpal et al., 2015).

The Crosswise Model (CM) was especially designed to overcome these issues (Yu, Tian, & Tang, 2008). It relies on a simple design that only requires respondents to provide one simple yes/no-answer to a set of two different questions without burdening them with complex instructions or activities (Ulrich et al., 2012, Höglinger, Jann, & Diekmann, 2014, Kornrörfer, Krumpal, & Schmukle, 2014). While one question directly asks about the sensitive attitude, the other one directly enquires about an unrelated non-sensitive topic. However, respondents are instructed to provide only one answer to both questions. Response A, if the answers to both questions are the same (either both ‘yes’ or both ‘no’), Response B, if the answers to the two questions differ (either ‘yes’ and ‘no’ or ‘no’ and ‘yes’).

Simple probability methods then allow estimating the prevalence of the sensitive item, if (1) both behaviours are captured as binary responses; (2) the non-sensitive behaviour is unrelated to the sensitive one; and (3) the non-sensitive behaviour has a known probability distribution p (Krumpal et al., 2015). For example, assuming a uniform distribution of birthdays, p = 0.25 if the unobtrusive item asks whether a person was born in October, November, or December.

The prevalence of the sensitive item $\lambda = P(\text{yes in sensitive item}) = P(\text{yes in non-sensitive item}) - P(\text{yes in non-sensitive item}) \times P(\text{yes in sensitive item})$ can then be estimated (Krumpal et al., 2015; see also Yu et al., 2008; Warner, 1965) by

$$r_{CM} = \frac{\hat{p} + 1}{2p - 1}, \hat{p} \neq 0.5,$$

where p is the known population prevalence of the non-sensitive item – e.g., p = 0.25 in the birthday example – and $\lambda = P(\text{yes in sensitive item})$ is proportion of respondents giving the same answer to both questions in the CM.

The sampling variance is then obtained as follows:

$$\text{Var}(r_{CM}) = \frac{\hat{p}(1 - \hat{p})}{n(2p - 1)^2} \text{Var}(1 - r_{CM})$$

Individual respondent’s responses to the sensitive item are covered up by the design of the question, so that interviewers and researchers are unable to identify whether or not the respondent provided a yes-answer to the sensitive item. By design, the question should thus encourage respondents to more honest self-reports. In addition to this, the CM is also an attractive way of asking about sensitive behaviour across different survey modes (Yu et al., 2008; Krumpal, 2015), given that sufficient statistical power is provided by the survey design to ensure similar precision of the estimates as direct question formats (e.g., Jann, Jerke, and Krumpal, 2012, Ulrich et al., 2012, Kornrörfer, Krumpal, & Schmukle, 2014).

Previous studies have indicated that the CM successfully reduces under-reporting of socially undesirable behaviour, such as plagiarism (Coulls, Jann, Krumpal, & Näher, 2011, Jann, Jerke, & Krumpal, 2012, Höglinger, Jann, & Diekmann, 2014), cheating in games (Hoffmann, Diedenhofen, Verschure, & Musch, 2015; Höglinger, & Jann, 2016), substance abuse (Nakhaee, Pakravan, & Nakhaee, 2014, Chaudhuri, 2016), donation of organs (Höglinger, & Diekmann, 2017), sexual behaviour (Vakilian, Vakilian, & Keramat, 2014; Vakilian, Vakilian, Keramat, & Keramat, 2014), tax evasion (Korndörfer, Krumpal, & Schmukle, 2014; Kundt, 2014; Höglinger, & Jann, 2016), undeclared employment (Schnell, Thomas, & Noack, 2017), non-voting (Höglinger, & Jann, 2016), and Xenophobia and Islamophobia (Hoffmann, & Musch, 2016).

However, little research has attempted to more critically evaluate the CM’s underlying mechanisms and their consequences. For instance, only two studies so far indicate that the CM is prone to producing false positive results that should not be ignored (Höglinger, & Diekmann, 2017; Höglinger, & Jann, 2016). It remains open how the technique performs when (1) high incidence behaviour, i.e. behaviour that occurs more frequently (e.g., Wolter, & Laier, 2014), (2) representative samples (with the exception of Schnell, Thomas, & Noack, 2017), or (3) real-world settings (Krumpal et al., 2015) are concerned. Hence, validation studies exploring how well and in which situations the CM works are required (Jann et al., 2012).

Our core research aim is to investigate the CM’s validity when high incidence behaviour is concerned looking at prejudice against Muslims in the Austrian context. More precisely, we test the concurrent validity of the CM focusing on self-reported attitudes towards Muslims. We proceed as follows: we begin by discussing our approach in further detail. Next, we present the results of our research. We close with a discussion of our findings and their implications for future research.

Testing the Concurrent Validity of the CM

Our research relies on data collected by the Austrian National Election Study (AUTNES) (Kritzinger, Johann, Thomas, Glantschnigg, Achtholzer, Ginztler, Gründl, Oberläugauer, & Wagner, 2016a; Kritzinger, Johann, Thomas, Glantschnigg, Achtholzer, Ginztler, Gründl, Oberläugauer, & Wagner, 2016b). Respondents were sampled from an online access panel provided by Lightspeed GMI/TNS Option Brussels. The non-probability panel consisted of approximately 15,000 panelists recruited on an opt-in basis using a combination of indirect approximation methods (Vehovar,
Testing the Validity of the Crosswise Model: A Study on Attitudinal Influences of Respondents' Left-Right Self-Placement on Negative Attitudes Towards Muslims in Both Models: More Rightist Citizens Appear to Hold Substantively Implies That Well Educated Respondents Are Less Likely to Hold Negative Attitudes Towards Muslims. We Do Not Find a Similar Effect Table 1 Summarizes Our Models Predicting Negative Attitudes Towards Muslims. Looking at Models 1a and 1b, It Is Noteworthy That We Find a Statistically Significant Positive Influence of Respondents' High Educational Attainment on Negative Attitudes Towards Muslims in the DQ, Which Substantially Implies That Well Educated Respondents Are Less Likely to Hold Negative Attitudes Towards Muslims. We Do Not Find a Similar Effect Looking at the CM. A Wald Test Indicates That the Coefficients of Both Models Significantly Differ (p < 0.05). Moreover, Our Results Indicate Positive Influences of Respondents' Left-Right Self-Placement on Negative Attitudes Towards Muslims in Both Models: More Rightist Citizens Appear to Hold More Negative Attitudes Towards Muslims in Their Neighbourhood.

Results Table 1 Summarizes Our Models Predicting Negative Attitudes Towards Muslims. Looking at Models 1a and 1b, It Is Noteworthy That We Find a Statistically Significant Positive Influence of Respondents' High Educational Attainment on Negative Attitudes Towards Muslims in the DQ, Which Substantially Implies That Well Educated Respondents Are Less Likely to Hold Negative Attitudes Towards Muslims. We Do Not Find a Similar Effect Looking at the CM. A Wald Test Indicates That the Coefficients of Both Models Significantly Differ (p < 0.05). Moreover, Our Results Indicate Positive Influences of Respondents' Left-Right Self-Placement on Negative Attitudes Towards Muslims in Both Models: More Rightist Citizens Appear to Hold More Negative Attitudes Towards Muslims in their Neighbourhood.

In addition to the socio-demographic controls, Models 2a and 2b include self-reported attitudes towards Muslims. At the first glance, the results reveal relatively consistent results regarding reports about the influence of negative views about Muslims on respondents' preferences to live in a neighbourhood without Muslims across the CM and DQ model: the coefficients of four critical attitudes point in the same positive direction in the DQ and CM condition indicating that negative attitudes towards Muslims and foreigners correspond with a lower likelihood of wanting to live in a...
neighbourhood with Muslims. However, the liberal statement suggesting that the European and Muslim lifestyles are compatible stands out: the results reveal a negative influence in the DQ format suggesting that more liberal views correspond with a lower likelihood to live in a neighbourhood with Muslims (which we would expect). However, the coefficient indicates a positive direction for the CM question (which we did not expect). A Wald Test indicates that the coefficients differ significantly (p < 0.001).

Looking at implicit preferences for Occidentals over Orientals in Models 3a and 3b, the results imply that those respondents also hold more negative attitudes towards Muslims, as suggested by the positive direction of the AMP coefficient.

For completeness, we also present full results in Models 4a and 4b, which reinforce some of the inconsistencies regarding explicitly stated attitudes towards Muslims and which additionally indicate a statistically different coefficient for the statement “Muslims are largely to blame for occasional tensions between Muslims and Non-Muslims in Austria” (p < 0.05).

### Discussion and Conclusion

The Crosswise Model has repeatedly been found to reliably reduce Social Desirability Bias when estimating low incidence behaviour. However, little is known about its validity when high incidence behaviour or a real-world setting is concerned. By investigating the concurrent validity of the Crosswise Model on reported negative attitudes towards Muslims in Austria, we contribute to a study that may help us to further evaluate the properties of the Crosswise Model.

We would have expected to find that our predictors of negative attitudes towards Muslims point in the same direction across the Direct Question format and the Crosswise Model. However, our results raise further questions about the validity of the Crosswise Model with regard to the concurrent validity in predicting negative attitudes towards Muslims. While we identify some factors that appear to predict the estimates obtained on the basis of the Direct Question format and Crosswise Model well (i.e., regression coefficients point in the same direction and reach conventional levels of statistical significance), other results are rather puzzling.

Traditional indicators of liberalism, such as respondents’ self-reported placement on the left-right continuum, as well as other attitudes towards Muslims seem to produce fairly consistent results indicating that both question formats are able to estimate respondents’ rejection of Muslims as neighbours quite well. However, we are particularly puzzled by some insignificant coefficients of self-reported attitudes towards Muslims in the Crosswise Model, which even point into the opposite direction than in the DQ model. As we rely on strong measures that should indicate an influence in the same direction, such as other negative attitudes towards Muslim and implicit associations, we believe that our results point in the direction that the Crosswise Model may come with problems that have not been fully uncovered yet. As Schnell, Thomas, and Noack (2017) indicate the success of the Crosswise Model in capturing sensitive behaviour or attitudes may be related to respondents cognitive abilities, especially their educational background: well educated respondents may be better able to understand the core principle of the Crosswise Model protecting their anonymity and confidentiality, while less well educated respondents are more concerned about the unusual question format. Our rather peculiar finding regarding respondents education in this study, may be interpreted as support for their findings pointing in a similar direction. Furthermore, we may observe that more sophisticated respondents are not genuinely less negative about Muslims as studies based on Direct Question formats suggests, but that they are simply more aware or more prone to give socially acceptable answers per se, which some research on this matter proposes (e.g., Jackman, 1973, Wagner, & Zick, 1995, Ostapczuk, Musch, & Moshagen, 2009).

To conclude, although the Crosswise Model has previously been commended to be a promising method to reduce Social Desirability Bias, our research suggests that we yet do not fully understand under what circumstances the method is applicable. Without completely questioning the validity of the Crosswise Model, we encourage future research to more rigorously test questions using the question format in larger population studies, ideally based on probability samples; to investigate high and low incidence behaviour or attitudes; and to study the performance of the method in different countries. In order to fully understand under what conditions the Crosswise Model works and to further develop the technique, it indispensable to address these issues.

### Online Appendix

Testing the Validity of the Crosswise Model: A Study on Attitud...
Testing the Validity of the Crosswise Model: A Study on Attitudes Towards Muslims

We would like to report the prevalence of negative attitudes towards Muslims anyway. Both measures estimate almost identical proportions of negative attitudes towards Muslims (DQ: 59.8%, SE=1.4, n=1205; CM: 60.2%, SE=2.9, n=1205).

References


