

Implications of Questionnaire Redesign and Challenges to the Continuation of Trend in the 2011 Police-Public Contact Survey

Glynis Ewing¹ and Marcus Berzofsky¹

¹RTI International, 3040 E Cornwallis Rd, RTP, NC 27709

Abstract

Each time an established periodic survey introduces a newly redesigned field instrument, there is a risk of introducing between-year measurement error that can interrupt the continuation of statistical trend. Regardless of improvements to the measures of interest, a redesigned questionnaire introduces potential instability to the validity of trend estimates for outcomes of interest between survey years. This paper will discuss sources of measurement error specific to redesigned survey instruments, describe methods for evaluating estimate comparability between survey years in the event of a redesign, and provide a solution for adjustment to correct for measurement error, illustrated by select examples from the 2011 Police-Public Contact Survey (PPCS).

Key Words: questionnaire design, redesign, periodic survey, trend estimation, ratio adjustment, measurement error

1. Introduction

Survey instruments used in periodic surveys are often changed over time for many reasons, with goals ranging from improving current measures to expanding the scope of an instrument to include emerging measurement objects of interest. For surveys that examine trends over time, it is critical that trend estimates for outcomes of interest are not interrupted by changes to the instrument. One of the difficulties in assessing the impact of an instrument redesign is determining the amount of an estimate change that is “real” (i.e., due to change in time) and the amount that is due to the change in the instrument. Therefore, any questionnaire instrument should be pretested or piloted prior to use in the field to ensure that the revised questionnaire is still measuring the outcomes of interest with valid and reliable comparability with past instruments for trend estimation. Regardless of the intent for improvement, changes to any instrument used in a periodic survey over time can introduce between-year measurement error. Once an instrument has been altered and used to collect data, it is critical to determine to what extent any observed changes in trends are due to a true change or measurement error introduced by the revised questionnaire. To do so, the revised questionnaire must be qualitatively and quantitatively evaluated in reference to the original questionnaire.

This paper will introduce some of the issues surrounding measurement error specific to redesigned instruments in periodic surveys and offer methods of assessment and correction for previous years’ measurement error in the case of a new and improved, redesigned instrument illustrated by select examples from the Police-Public Contact Survey (PPCS), a triennial supplement to the National Crime and Victimization Survey (NCVS).

2. Issues Related to Questionnaire Redesign

2.1 Understanding the problem

The main goal of any periodic survey administered over time is to understand how estimates of interest change over time. Even if not measured perfectly, as long as measured consistently, trend estimates can serve as valid indicators of change. However, once a redesign of an instrument occurs, any measurement differences between the two instruments need to be addressed in order to allow for continued trend analyses.

When there is a qualitative change to a survey instrument and there is evidence of a quantitative change in trend when comparing the resulting data with that collected by previous iterations of that instrument, it becomes difficult to determine if that change in trend is a valid change in outcome or if that change is due to a change in the measurement of the outcome. Therefore, to understand the impact of the instrument change, it is critical to utilize methods that help isolate the amount of change due to measurement as best possible.

2.2 Questionnaire Redesign & Measurement Error

Because survey instruments are ultimately designed to glean information from respondents, to create data with analytic utility, questionnaires are written, formatted, ordered, and administered in such a way to elicit a response appropriate for the outcome of interest (Bradburn et al., 2004). Typical measurement error in surveys is due to bias and naturally occurring variance that exists as a result of the interaction between interviewer, questionnaire, and respondent (Schaeffer, 1991). However, redesigning an instrument for use in an existing periodic survey places a different set of issues over the typical sources of measurement error encountered when designing an instrument for a new survey or making changes in a trend-breaking survey year.

When an instrument establishes a certain way of measuring particular outcomes of interest, any change to that instrument can result in a change in the way each set of respondents answer the instrument and thus potentially result in an estimate measuring something different than the corresponding estimate from a previous survey year. Due to the existence of an established trend measured by previous instrument versions, any changes to questionnaire design carry a likelihood of introducing a problematic level of change to the measurement of that estimate (e.g. measurement error). In the case of a redesign resulting in increased accuracy in the outcomes of interest, the improved survey instrument provides an opportunity to “calibrate the effect” of any changes to the questionnaire items (Clark et al., 2003). This section focuses specifically on potential sources of measurement error in redesigned survey instruments that can introduce instability to the validity of established trend estimates.

2.2.1 Importance of Respondent Domain

Respondent domain refers to the group of respondents presented with the opportunity to answer the given questionnaire item(s) of interest. Some questionnaire items may have all respondents included in the domain whereas some may have a very small percentage of respondents eligible to answer the question, or questionnaire items may expand on the existing domain. When initially designing a survey instrument, the respondent domains are generally well-defined and are not a major source of measurement error. When redesigning a questionnaire for use in an already-established survey, it is important to ensure that the

redesigned instrument allows for the re-creation of consistent respondent domains for questionnaire items across survey years. For this reason respondent domain is a critical consideration when assessing potential sources of measurement error due to a questionnaire redesign.

Without a consistent or reproducible respondent domain, a questionnaire item can exist in a revised instrument in the exact wording, formatting, or order as the original instrument and the resulting estimate will be impossible to compare to the original estimate from the previous survey year. Differences in respondent domain can be caused by various errors of questionnaire design, for example, *errors of commission* in which respondents receive questionnaire items they are ineligible or not supposed to answer and *errors of omission* occur when a respondent does not receive the questionnaire items he or she should have answered (Kreuter & Keusch, 2015). Errors of commission or omission can occur by moving a questionnaire item to a different section of the instrument, e.g. a different module, or can occur due to gate logic error where a respondent is routed to either irrelevant questionnaire items or out of the interview. However, the expansion of respondent domain can be desirable in the case of a redesign when previous instruments underestimated outcomes of interest.

2.2.2 Other Relevant Factors

Ideally, all potential questionnaire revisions would be based on results from cognitive interviews using the existing instrument and the resulting revised instrument would undergo extensive pretesting in order to ensure the revised measurements will “work” with existing trend estimates, regardless of the goals of the redesign (Clark et al., 2003). The U.S. Census Bureau implements a “minimal pretesting standard” in cases of new or updated questions added to any existing survey instrument which requires the questionnaire be pretested to validate and refine the questions before they are put into production (Clark et al., 2003). Pretesting is a general term encompassing “different methods or combinations of methods” such as pre-field techniques like cognitive interviews or respondent focus groups and field techniques like analysis of item nonresponse rates or split panel/split-sample tests (see Clark et al. 2003 for a detailed discussion).

The ideal pretesting methods can be time-consuming and do not always occur in practice prior to deploying a redesigned instrument in the field for an ongoing, periodic survey. Changes to question wording, formatting, or order within the instrument from year to year can introduce levels of measurement error because each set of respondents are being prompted differently. These types of changes to instrument items can range from small wording tweaks that introduce minimal between-year measurement error to alterations in item location within the instrument that result in errors of commission or omission that eliminate the possibility of any trend continuation. Trend estimates are more robust to changes like slightly re-worded questions and expanded question structures when the respondent domain remains the same.

3. Research Questions & Objectives

In light of potential trend-breaking differences between previous instruments and any redesigned instrument in a periodic survey, we propose two research questions:

1. How does an analyst assess the extent to which an estimate has changed due to measurement error rather than an actual change in trend?
2. Given a change in the estimates due to measurement error, how does an analyst adjust for that error in order to maintain trend estimates?

To answer these questions, the first objective is to establish which outcomes of interest are comparably measured between the original and revised survey instruments. Once those comparable estimates are established, the second objective is to quantitatively assess those estimates associated with the analogous questionnaire items from the previous instrument and adjust for any measurement error accordingly.

These objectives were applied in an evaluation of the 2011 Police-Public Contact Survey (PPCS), which will be referenced as an applied example of methods outlined in this paper.

4. Methods

The first objective described in *Section 3* is more qualitative in nature, and the second objective is quantitative. After a field instrument has been revised, the questionnaire items should maintain comparability with the previous questionnaire items despite changes in wording, formatting, or structure. In the case of an instrument revised with the goal of enhancing response to certain outcomes of interest, a split panel or split sample design is the most appropriate method to “calibrate” the changes made in the instrument (Clark et al., 2003).

4.1 Qualitative Measurement-based Comparability

We define two criteria that a revised questionnaire item must meet in order to be determined comparable with its previous instrument’s counterpart:

1. Measures same object or outcome of interest via similar or identical wording or expands upon the existing measurement;
2. Shares the same respondent domain

Data quality is heavily influenced by the questionnaire format (Sanchez, 1992). As such, the qualitative review of a revised instrument to determine if the two criteria are met should consist of the following to determine if there is the potential for measurement error between the two instruments:

- Question wording,
- Question placement, and
- Skip patterns

When reviewing question wording, one should be cognizant of different definitions or phrasing that may illicit a different interpretation by the respondent. If the question wording changes the entire meaning of the previous item then it may not be possible for an adjustment to account for the measurement error.

For question placement and skip patterns, one should determine what, if any, errors of commission or omission are possible. For example, if the new instrument alters the respondent universe for an item (i.e., the number of respondents eligible for the item changes) the extent to which the universe will change needs to be assessed. For outcomes that will produce rate or proportion estimates, a large change in the universe can greatly impact the estimate due to the change in the denominator of the estimate.

4.2 Quantitative Difference & Adjustment Eligibility

Given the structural and definitional differences between the original and redesigned instruments, it is necessary to quantify the magnitude that these differences have on estimates. Differences in estimates between across two survey periods are due to two attributes: (1) changes to the instruments, and (2) time.

In order to control for time, a split-sample design should be implemented. A split-sample design randomly divides the sample into groups and assigns different treatments to each group¹. A benefit of the split-sample design is that it may be “embedded” in a planned survey year data collection (Clark et al., 2003). In the case of an assessment of instrument changes, the sample can be split into two groups where Group 1 is administered the new instrument and Group 2 is administered the previous instrument. By administering both instruments to a random sample during the same calendar period, significant differences between the two groups are likely due to measurement error rather than random variation.

Another benefit of the split-sample design is that it allows for the quantification of the effect of changing questions (Clark et al., 2003). To compare the two samples, adjustments should be made for nonresponse and coverage errors. After these adjustments, the two samples should equally represent the target population at the same point in time. Bivariate comparisons can be made using t-tests. These tests can be computed across all respondents and within key subdomains – e.g., gender, age category, or race/ethnicity. A statistically significant difference for an outcome indicates that prior period estimates should be adjusted for any trend comparisons.

4.3 Method of Adjustment

When estimates from two survey instruments administered simultaneously to similar respondents are significantly different, there is evidence that valid trend estimation from past years cannot continue without adjusting the estimates to the revised instrument.

In 1992 the National Crime Survey (NCS) underwent a major instrument redesign. The resulting survey – renamed the National Crime Victimization Survey (NCVS) – had an altered instrument whose goal was to improve the cues used to determine if a respondent experienced a crime during the previous six months. In order to understand the impact of changes to the NCVS instrument on the estimates, Cantor and Lynch (2005) developed a ratio adjustment by which the NCS victimization rate (X) for crime type j is adjusted as follows:

$$X_{NCSj}^A = X_{NCSj} \times R_j$$

Where R_j is the ratio adjustment based on the 1992 split sample for crime type j which is defined as

$$R_j = \begin{cases} \frac{X_{NCVSj}^S}{X_{NCSj}^S} & \text{if } \frac{X_{NCVSj}^S}{X_{NCSj}^S} \text{ significantly different from 1 and } > \text{ than 1} \\ 1 & \text{Otherwise} \end{cases}$$

Where X_{ij}^S i =NCS or NCVS is the estimated victimization rate from the split sample for crime type j . Similarly, standard errors are adjusted assuming the ratio adjustment is a fixed constant

¹ The two groups need to be mutually exclusive, but do not need to be of the same size; however, splitting the sample in half maximizes the power to detect differences.

$$SE(X'_{NCS_j}) = SE(X_{NCS_j}) \times R_j$$

As noted in the definition of R_j , Cantor and Lynch only modified prior year estimates if two conditions were met: (1) R_j is significantly different from 1, and (2) $R_j > 1$. The second condition was added because the new NCVS instrument was aimed at eliciting more crime rather than less. Therefore, Cantor and Lynch did not think it was theoretically correct to increase the crime rate in prior years when comparing to the NCVS estimates.

5. Application to the Police-Public Contact Survey

The Police-Public Contact Survey (PPCS) is a triennial supplement to the NCVS that interviews a nationally representative sample of U.S. residents aged 16 and older about their experiences with the police during the year preceding the interview. Introduced in 1999, the PPCS collects data to describe the amount, nature, and perceptions of individuals that have formal contact with the police. Respondents are asked to describe the nature of their most recent police contact, report the outcomes of the encounter, and detail their perception of the officer's conduct (Langton & Durose, 2013).

5.1 Redesign of the 2011 PPCS

In 2011, the PPCS instrument was substantially revised to better capture police contacts and characteristics of police-public encounters. The PPCS was changed to enhance respondent recollection of interactions, to collect information about contacts that were not face-to-face, and to gain a more nuanced understanding of certain characteristics of police-citizen contacts such as use or threat of force.

Instruments used in the 2002, 2005, and 2008 survey years focused on collecting information about respondents' most recent face-to-face contact with the police in the 12 months preceding the interview date. Instead of focusing on a single most recent contact, the revised 2011 instrument introduced a screener module of questionnaire items to capture a range of contacts with police the respondent might have had in the year preceding the interview date. This screener module also expanded the definition of police-public contact to include non-face-to-face interactions (e.g. phone calls) and a wider range of contacts respondents voluntarily initiated with the police (e.g. participation in a block watch). Though the redesigned instrument added measures of contact that had never been measured in previous survey years, it also ensured to measure types of contact already existing in previous instruments. Beyond these goals of expanding the types of contact included in the questionnaire, the screener module also improved on the questionnaire design by allowing respondents to report multiple contacts.

In order to evaluate the impact of the 2011 instrument redesign of the PPCS, we followed the two methods detailed in Section 4. Namely, we (1) qualitatively reviewed the 2011 and 2008 instruments to make sure the underlying constructs to be estimated remain the same, and (2) through a split sample, conducted by the Census Bureau, we quantitatively assessed the impact of the instrument controlling for calendar time. Details of the specific methods for each approach are detailed in *Section 5.1.1* and *Section 5.1.2*, respectively.

5.1.1 Comparing face-to-face contact

An example of a redesigned measure from the revised 2011 PPCS instrument that meets the criteria to be comparable is face-to-face contact with the police in the past 12 months from the interview date. The two versions of the instrument differed in the format used to

measure face-to-face contact. In the previous instrument, respondents were asked a singular gate item with yes/no response choices to indicate whether or not they had a face-to-face contact with the police in the past 12 months; respondents answering yes were asked about types of contact, and respondents answering no were routed to the end of the instrument. Where the original instrument asked a yes/no gate question before asking about specific types of contact, the revised instrument eliminated this gate and asked respondents to indicate the types of contact they have had with police in the 12 months preceding the interview date using a vector of dichotomous questions. This revised method of measuring face-to-face contact with the police expanded upon the original instrument by not only assuming certain types of contact with the police would be face-to-face and expanding the measure, but the change in structure was intended to improve respondent recall by potentially reminding respondents of encounters they might have forgotten (Fowler, 1995). The revised measurement of face-to-face contact utilizing the vector of dichotomous items also shares the same respondent domain as the original questionnaire item, ultimately making the revised questionnaire item an improvement over the original questionnaire item even in the presence of changes to wording, order, and format. Because of this qualitative comparability, it can be inferred that any statistical difference detected between estimates from the split sample is due to measurement error.

5.1.2 Evaluating quantitative differences

In order to assess these differences one cannot simply compare the estimates from the 2011 PPCS survey to those of the 2008 PPCS survey. In order to control for time, when conducting the 2011 PPCS, the Census Bureau implemented a split sample design with random assignment allowing for comparisons between the original and revised PPCS instruments. Because BJS wanted to have reliable estimates to report based on the new instrument, the split-sample was allocated such that 84% of respondents (N=41,408) received the 2011 instrument while 16% of respondents (N=7,838) received the 2008 instrument. As the NCVS is a household survey, randomization was done at the household level such that all persons within a selected household received the same instrument. We can infer that, if the questionnaire item is similar between instrument versions, the associated outcome from the revised item is comparable, so thus any statistical difference detected between the two iterations of that associated outcome estimate would signal to the presence of measurement error in the revised instrument or highlight measurement error existing in previous estimates.

Table 1 shows results weighted to the U.S. population for use or threat of force during face-to-face contact with the police measured in the 2011 PPCS split sample with the corresponding significance testing results between instrument version overall estimates and the estimates by demographic subdomains. These estimates were qualitatively similar in the questionnaire text and format and were administered to the appropriate respondent domain, but the pairwise test shows the overall estimates are significantly different from each other at the 0.05 level. Though there were some significant differences between the estimates within subdomains, these groups either have a small sample size or coefficient of variation greater than 50% making tests invalid across subpopulation distributions.

Table 1: Quantitatively comparing of persons reporting use or threat of force during most recent contact with police by select demographic subdomains

	<i>Original Instrument[†]</i>			<i>Revised Instrument</i>		
	<i>Weighted Count</i>	<i>Weighted Percent</i>		<i>Weighted Count</i>	<i>Weighted Percent</i>	
Overall	1,176,000	2.5	%	2,277,300	4.1	%**
Sex						
Male	756,200	3.3	%	1,490,800	5.2	%**
Female	419,800	1.8		786,500	3.0	**
Race						
White/a	753,200	2.2	%	1,318,800	3.4	%*
Black/African American	233,800	4.3	!	424,000	6.5	
Hispanic/Latino	127,800	2.5	!	440,100	7.2	***
Other (non-Hispanic)	20,500	1.3	!	57,000	2.5	!
Two or more races	40,600	6.6	!	37,400	4.1	!
Age						
16-17	--	--	%!	74,200	5.2	%!
18-24	283,600	3.2	!	745,300	7.7	***
25-34	427,700	4.0		612,400	4.7	
35-44	171,000	2.2	!	327,900	2.9	
45-54	181,800	2.2	!	270,000	2.4	
55-64	111,900	1.8	!	140,700	1.7	
65 or older	--	--	!	106,900	1.7	

[†] Reference group

*Significant at the 0.10 level

**Significant at the 0.05 level

***Significant at the 0.01 level

! Interpret with caution; estimate based on 10 or fewer sample cases or coefficient of variation is greater than 50%

-- Number rounds to less than 0.5

5.2 Assessing Impact on Trend Estimation

As discussed in *Section 2*, redesigning survey instruments introduces potential measurement error regardless of improvement. Improved survey instruments can highlight existing measurement error in estimates created from previous instruments. The 2011 redesign of the PPCS survey instrument enhanced responses to the items associated with overall estimates for most types of contact and characteristics of contacts, showing where past instruments have suppressed or inflated certain estimates.

Considering the example of face-to-face contact with police, the change from a single questionnaire item in the 2008 instrument to a vector of questionnaire items to measure face-to-face contact with police in 2011 and the resulting statistically significant differences in estimates highlights the presence of measurement error in past instruments.

For comparability with the estimate from the revised instrument in the case of a significant difference, past year estimates need to be adjusted.

5.2.1 Adjustment in the PPCS

Where the Cantor and Lynch adjustment method introduced in **Section 4.3** made given crime type j the focus of the NCVS, the PPCS would need an adjustment method adapted for given police-public contact type (e.g., face-to-face contact, driver in a traffic stop). Since the 2011 redesign of the PPCS instrument was aimed at eliciting more responses about contact with the police, it would be appropriate to adapt the Cantor and Lynch adjustment method for application to past PPCS years. The resulting ratio adjustment used in the analysis of the 2011 PPCS split sample is made for each type or characteristic of contact, or for a given type of contact c , and it is defined as

$$R_c = \begin{cases} \frac{X_{2011c}^S}{X_{2008c}^S} & \text{if } \frac{X_{2011c}^S}{X_{2008c}^S} \text{ significantly different from 1 and } > 1 \\ 1 & \text{Otherwise} \end{cases}$$

And, the resulting adjusted estimate is calculated as

$$X_{YYYYc}^A = X_{YYYYc}^S \times R_c$$

Where YYYY represents the survey year prior to 2011 being estimated (i.e., 2002, 2005, or 2008).

Table 2 on the next page shows split sample estimates for some types of contact and characteristics of contacts with the corresponding t-test results and the resulting adjustment factor. Note that some measures comparable between instrument versions resulted in non-statistically different estimates, leading to an adjustment factor of 1.0.

Table 2: Overall weighted rates of outcomes of interest from the 2011 PPCS and associated adjustment factors

<i>Outcome of Interest</i>	<i>Original Instrument Estimate</i>		<i>Revised Instrument Estimate</i>		<i>Ratio</i>	<i>Adjustment Factor</i>
Face to face contact	19.10	%	22.80	%***	1.19	1.19
Use or threat of force during contact	2.50		4.10	**	1.64	1.64
Arrested during contact	3.63		5.94	***	1.64	1.64
Driver in a traffic stop as most recent contact	7.01		8.96	***	1.28	1.28
Passenger in a traffic stop as most recent contact	0.47		1.97	***	4.19	4.19
Traffic accident as most recent contact	0.36		0.15		0.42	1.00
Reported crime or problem to the police as most recent contact	4.42		6.96	***	1.57	1.57
Police behaved properly and respectfully during contact	88.60		90.00		1.02	1.00

*Significant at the 0.10 level

**Significant at the 0.05 level

***Significant at the 0.01 level

5.3 Results of Adjustment

The calculated adjustment factors were applied to the appropriate estimates from the 2002, 2005, and 2008 survey years to produce trend lines for each outcome of interest. To illustrate the practical application of the ratio adjustment, **Figure 1** on the next page presents the adjusted and unadjusted trend lines for the percentage of the population reporting a face-to-face contact with the police in the past year. An adjustment factor of 1.19 was applied to the overall estimates associated with face-to-face contact with the police for past survey years. After applying the adjustment factor, the results show that had the revised instrument been used in previous PPCS survey years, the percentage of the

population reporting a face-to-face contact with the police in the past year would have been higher overall. Even with the adjustment taken into account, the percentage of the population reporting a face-to-face contact with the police in the past year decreased in 2008.

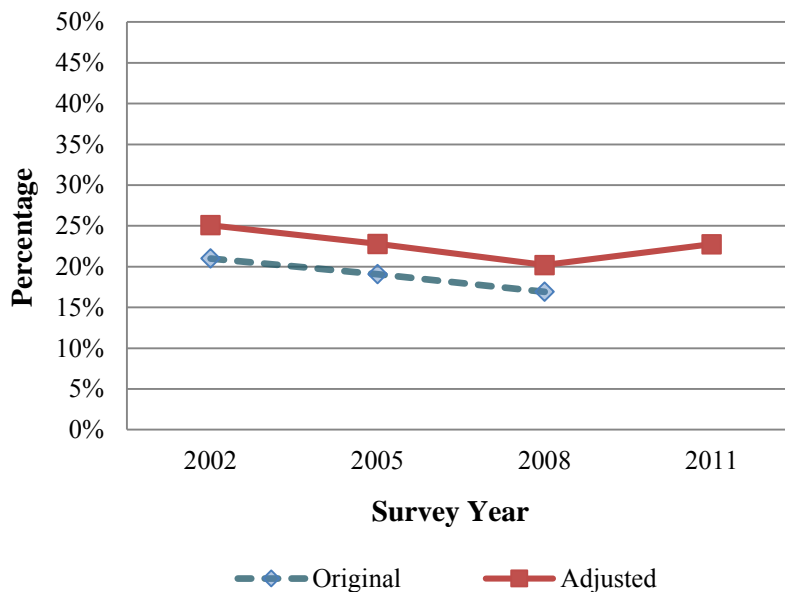


Figure 1: *Percentage of Respondents Reporting Contact with Police in the PPCS by Survey Year, Adjustment factor=1.19*

6. Conclusions

When introducing a redesigned field instrument to an already-established periodic survey, the continuity with established trends must be taken into account and differences between instruments should be understood on both a qualitative and quantitative level. Changes made to any survey questionnaire text have direct effects on the comparability of resulting estimates with past survey years. Revised items that share or expand upon the respondent domain of the analogous items found in the previous instrument allow for the possibility of trend continuation in the case of qualitative comparability between questionnaire item texts. Though continuation of trend is possible between two different survey instruments, the difference must be quantified and accounted for accordingly. In the case of the 2011 PPCS, a ratio adjustment was used to adjust estimates from past survey years to continue statistical trend in the event of a redesigned survey instrument.

Acknowledgements

The authors would like to thank the Bureau of Justice Statistics for sponsoring this research and comments provided by Mike Planty and Lynn Langton. However, we would like to note that the views expressed in this paper are those of the authors only and do not reflect on the views or position of the Bureau of Justice Statistics or the Department of Justice.

References

- Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking Questions: The Definitive Guide to Questionnaire Design – For Market Research, Political Polls, and Social and Health Questionnaires*, San Francisco, CA: Jossey-Bass.
- Cantor, D., & Lynch, J. P. (2005). Exploring the effects of changes in design on the analytical uses of the NCVS data. *Journal of Quantitative Criminology*, 21(3), 293-319.
- Clark, C. Z. F., Tinari, R., Singh, R., Tupek, A., Hogan, H., Killion, R. A., & Wright, T. (2003, July 25). *Census Bureau Standard: Pretesting Questionnaires and Related Materials for Surveys and Censuses*. Retrieved from <https://www.census.gov/srd/pretest-standards.html>
- Fowler, F. J. Jr. (1995). *Improving Survey Questions: Design and Evaluation*, Thousand Oaks, CA: SAGE Publications.
- Kreuter, F., & Keusch, F. (2015). *Week 9: Questionnaire from Start to Finish; Layout and Mode dependencies* [PowerPoint slides]. Retrieved from <http://jpsmonline.umd.edu/>.
- Langton, L. & Durose, M. (2013). Police behavior during traffic and street stops, 2011. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics (NCJ 242937).
- Police Public Contact Survey: A Supplement to the NCVS. (2011). Bureau of Justice Statistics.
- Police Public Contact Survey: A Supplement to the NCVS. (2008). Bureau of Justice Statistics.
- Sanchez, M. E. Effects of questionnaire design on the quality of survey data.” *Public Opinion Quarterly*, 1992, 56, 206-217.
- Schaeffer, N. C., & Presser, S. (2010). The science of asking questions. *Annual Review of Sociology*, 29, 65-88.
- Schaeffer, N. (1991). Conversation with a purpose or conversation? Interaction in the standardized interview, In Biemer P.P. et al. (Eds.), *Measurement Errors in Surveys*, New York, NY: Wiley, pp. 367-391.