Interviewer effects in studies of private demand for cholera vaccines:
evidence from India, Mozambique, and Bangladesh

by

Melissa E.B. McMahon

A Masters Paper submitted to the faculty of
The University of North Carolina at Chapel Hill
In partial fulfillment of the requirements
For the degree of Master of Regional Planning in the
Department of City and Regional Planning

Chapel Hill

2006

Approved by:

Advisor - Dale Whittington
I offer my sincere thanks to Professors Whittington and Eng, Joe Cook, and Brian Maskery for their advice and assistance with this project. It is unfortunate that most of what I’ve learned with the DOMI research team cannot be reflected in a research paper.

To Marc Jeuland, thank you for being such a superb research partner.

This research was made possible by the International Vaccine Institute (IVI) and the Bill and Melinda Gates Foundation.
Abstract

This paper investigated the impact of interviewers on respondent self-reported willingness to pay for cholera vaccines in three study sites: Kolkata, India; Matlab, Bangladesh; and Beira, Mozambique. Specifically examined were the relationships between interviewer gender and respondent gender, age, and education level. The analysis also measured the overall variance in responses caused by individual interviewers. It was found that only individual interviewer effects are present in the Kolkata and Beira studies, while interviewer gender effects and gender-gender, -age, and –education interactions are present in the Matlab study. When controlling for individual interviewer effects in the Matlab analysis, only the gender-education interaction is still significant. Despite the high levels of significance of individual interviewer effects in each study, the magnitude of these effects on response variance as indicated by their coefficients is still fairly small. These results suggest that these CVM studies may be more resistant to interviewer characteristic effects than other studies that a) focus on gender-sensitive issues, b) use smaller numbers of interviewers or c) use smaller sample sizes. The results also suggest that while whole interviewer effects are present in all three studies, the size of the interviewer teams prevents the few aberrant interviewers from causing large amounts of variance across the whole sample, keeping the CVM results reliable and likely to be generalizable to their respective populations.
1. Introduction

The epidemiology of cholera

Cholera is a severe diarrheal disease that can result in acute dehydration and can lead to death. It is caused by the infection of a person’s intestine with the bacterium, *vibrio cholerae*, through drinking contaminated water or eating contaminated food. The source of the contamination is usually the feces of an infected person, which can then spread rapidly in areas with inadequate treatment of sewage and drinking water supplies. Hence, the incidence cholera is highest in urban areas with high population density and insufficient water and sanitation infrastructure. In 2003, over 110,000 cases of cholera were reported worldwide, which led to 1,894 deaths (World Health Organization [WHO], 2004). In 2004, Mozambique reported 20,080 cholera cases, which was more than any other country in the world (WHO, 2005). In Beira, the second-largest city in Mozambique, cholera is endemic and outbreaks are frequent, occurring most recently in the spring of 2004 and again in January of 2006. Certain environmental conditions have been attributed to Beira’s cholera problem for city residents. These conditions are: urban development in and around marshes that flood annually, human defecation in open areas, widespread use of non-tight pit latrines, and disposal of household wastewater in open embankments (Ansaruzzaman, et al., 2005).

The public health response to cholera

In the absence of infrastructure improvements which are prohibitively expensive, next-generation vaccines are a safe and effective approach to addressing a number of contagious and debilitating diseases in developing countries (Acosta, 2005). However, diseases differ in their relative impacts on a population, so it is important that researchers
try to measure the impact of the disease and compare this with the cost and effectiveness of the vaccine program in order to make convincing arguments for which vaccine programs should be implemented under tight budget constraints. Disease-related morbidity and mortality, vaccine effectiveness (numbers of death averted, disability-adjusted life years, DALYS, saved), people’s willingness to be vaccinated, and the cost of producing and distributing vaccines are all important considerations when deciding what vaccination programs to implement (Global Alliance for Vaccines and Immunization [GAVI], 2004; Pack, et al., 2006).

To justifying vaccination programs, a common approach is the use of a cost-effectiveness measure which compares certain benefits (deaths averted, DALYS saved, or fully-vaccinated persons) to total vaccination program cost. This approach generates a measure of efficiency: an amount of money spent on the program per death avoided, DALY saved, or fully vaccinated person. Donor organizations such as World Bank and Global Alliance for Vaccines and Immunization (GAVI) have used cost effectiveness measures of alternative programs to inform their decisions on which program to fund (Cavailler, et al., 2005; GAVI, 2004; Naficy, et al., 1998). In addition, donor organizations have also suggested that governments conduct a more complete assessment of proposed programs using a cost-benefit analysis (CBA). In a CBA the total benefits to the population must be measured and assigned a financial value in order to determine if they outweigh the total costs. CBA would suggest that there are intangible benefits to vaccines that are particularly hard to measure, such as the values people would place on avoiding pain, suffering and death from cholera.
The contingent valuation method

The contingent valuation method (CVM) is one way to measure this total private value for vaccines. Even in imperfect markets (e.g. information is incomplete, there are barriers to entering the market, products are not homogeneous) policy makers can get at least a sense of the value people place on certain goods. In the absence of any market for the good in question, CVM presents a hypothetical scenario to respondents in which a vaccine would be offered for a given price and then poses a series of closed-ended questions to measure respondents’ willingness to pay for the vaccine. Typically CVM generates estimates of benefits that are greater than the estimates of benefits obtained through measuring public and private cost of illness avoided. CVM estimates of people’s willingness to pay for vaccines, therefore, can provide new information about the value they place on vaccines for themselves and their families, when taking into consideration their personal budget as well as the value they place on avoiding pain, suffering and death. The sum of many individual demand statements approximates a market demand curve for the good we are interested in, a tool that policy makers can use to inform future decisions about if and how to make that good available at a price that people can support. Because the data informing CV estimates of the value of goods comes directly from stated preferences, it is crucial that we refine our techniques for eliciting these preferences such that they come even closer to representing the choices that people will actually make in practice.

CVM studies are most commonly conducted through in-person household interviews. As with all data collected through in-person interviews, a potential bias can be introduced by the interaction between interviewers and respondents (Hyman &
Stember, 1950). These interviewer effects should be measured in order to determine if any systematic influences are present that confound the relationship between price and demand. If interviewer effects exist and can be measured, they can be controlled for, increasing the reliability and validity of the CVM study.

This paper examined the impact of interviewers on respondent self-reported willingness to pay for cholera vaccines in three study sites: Kolkata, India; Matlab, Bangladesh; and Beira, Mozambique. Specifically examined were the relationships between interviewer gender and respondent gender, age, and education level. The analysis also measured the overall variance in responses caused by individual interviewers. In addition to reporting CVM results from each site, I compared results from the three sites and discussed implications for the quality of data collected at each site. I conclude with methodological suggestions for future research.
2. Literature Review

Method bias

A common source of bias or variance in social research data is “method bias” which results from the measurement method rather than from the ideas or constructs being measured. In a review by Podsakoff, Mackenzie, Lee and Podsakoff (2003) of method biases in social and behavioral research, they concluded that systematic sources of measurement error, such as method bias, are more detrimental to research results than random ones because systematic error provides an alternative explanation for the observed relationships that is independent of the one hypothesized. Method bias errors can occur at many levels of a study, from the content of specific survey items and response formats to the broader context of response reporting biases like social desirability, acquiescence bias, or yea- and nay-saying. Furthermore, Podsakoff et al. (2003) connected types of method bias to different stages of the question-response process. For example, interviewer effects are method biases that occur during the “response reporting” stage of the question-response process.

Interviewer effects studies

Early studies of the influence of interviewer effects on survey response quality (Ferber and Wales, 1952; Feldman, Hyman and Hart, 1952) found contradictory results. While Ferber and Wales (1952) reported significant differences in response reporting between interviewers, Feldman et al. (1952) found that for most items in the survey responses did not differ significantly by interviewer. Both authors suggested that future studies of interviewer effects need to refine the research design and analytical techniques used.
A 1983 study by Singer, Frankel and Glassman approached the question of interviewer effects by examining certain interviewer characteristic effects individually, an attempt to replicate in telephone interviews the questions and findings in two previous studies that used personal interviews. This study examined the effects of interviewer characteristics and expectations on both overall cooperation rates and on item non-response rates and response quality. They found that older interviewers obtained better cooperation from respondents than younger interviewers, and that this effect was independent of interviewing experience. The effect of interviewer experience was positive to a point, but with longer experience screening rates leveled off and response rates actually declined. Attitudinal variables had no significant effect on screening rates, but interviewer expectations did have a strong and significant effect on the response rate, with interviewers with optimistic attitudes having significantly higher response rates than those with pessimistic attitudes. Most other findings in this study were insignificant. The study was only partially able to replicate effects found in previous studies. The authors attribute the weakness of results in this study to the fact that it was a phone survey sample (rather than in-person interviews), the response rate was lower than the in-person survey study, and there wasn’t as great a variation in interviewer expectations as there was in the in-person study.

A popular interpretation of direct interviewer effects such as those observed in earlier studies is called the Social Attribution Model (SAM) (Johnson and Parsons, 1994). This model postulates that interviewer characteristics alone are enough to influence respondents’ reporting of replies during an interview through the respondents’ observations about their interviewers and the subsequent editing or tailoring of their
replies. However, some common critiques of these early interviewer effects studies included: the samples of interviewers and respondents were small, only a select few measures that show significant outcomes were reported, and the influence of the individual interviewer was not controlled for in sample-wide analyses of interviewer actions, expectations or characteristics. Furthermore, early studies of interviewer effects did not include comparison studies. Consequently, journals editors requested explicit comparisons from authors of manuscripts about interviewer effects. In 1969 *The Public Opinion Quarterly* solicited a “reconciliation of findings” from three researchers, separate from their respective articles on interviewer effects, in an attempt to provide readers with a better sense of the generalizable conclusions that could be drawn from the studies. (Dohrenwend, Williams and Weiss, 1969).

*Individual characteristic and interaction effects*

According to Dijkstra (1983) the term “interviewer effect” refers to that portion of total response variance which can be attributed to a particular interviewer characteristic. He distinguishes this from “interviewer variance” which can be thought of as the cumulative effect of all interviewer characteristics, measured and unmeasured. Gender-of-interviewer effects describe the situations in which a portion of response variance can be attributed to the gender of the interviewer. Evidence of these effects is most convincing when it is apparent separately in each gender sub-sample of the respondent group, because this controls for the influence of respondent gender.

More recent studies have attempted to explain interviewer effects through statistical methods that measure the impact of specific interviewer characteristics such as race, gender or age, and interact these characteristics with characteristics of respondents
Huddy et al. (1997) also found similar evidence for gender-of-interviewer effects on responses related to women’s issues and controversial political topics related to feminism. They found mixed evidence for effects from gender-of-interviewer interacting with gender-of-respondent, and one of the two area surveys showed more pronounced gender-of-interviewer effects on responses from less well-educated and younger respondents as compared to more well educated or older respondents. Huddy et al. (1997) used a difference of mean scores on survey items across interviewer gender categories (t-tests), then regressed each dependent variable onto interviewer gender and dummies for all interviewers, followed by models with interaction variables designed to look for
respondent characteristics most susceptible to interviewer gender effects. The generalizability of these findings, however, is limited. Similar to the Kane and Macaulay (1993) study, Huddy et al.’s survey instrument was focused on gender issues.

Race-of-interviewer effects on the quality of survey responses have also been explored. One study (Davis, 1997) found that African American respondents interviewed by white interviewers were more likely to provide mutually contradictory opinions to questions about political parties and beliefs, a race-of-interviewer effect that was attributed to acquiescence. In this study the author used a t-statistic to evaluate sub-samples by race of interviewer, looking at the percent of respondents who were acquiescing to mutually contradictory political positions. He followed this with OLS to evaluate the impact of change in interviewer on respondent answers in a second interview. A later study by Davis and Silver (2003) on stereotype threat built on the acquiescence explanation. The 2003 study specified a more elaborate social-psychological explanation for race-of-interviewer effects to explain why interviewer effects could occur on factual questions as well as opinion and attitudinal ones. In this study, African American respondents interviewed by white interviewers got fewer answers correct to political knowledge questions than those interviewed by African American interviewers, controlling for educational background and gender of respondents. These studies provide further suggestive empirical evidence that characteristics of interviewers and respondents, in this case race, may influence the accuracy of survey results.

Some studies of interviewer effects have focused on reports of drug use behavior (Mensch and Kandel, 1988, Fendrich et al., 1999) and illustrate the challenges of eliciting
certain kinds of information through face-to-face interviewing. In the Mensch and Kandel (1988) study the authors explored the quality of drug data from a 1984 wave of the National Longitudinal Survey of Youth, focusing on the underreporting of illicit substance use. The primary statistical tool was f-tests within categories of independent variable (e.g. levels of education, marital status) by gender of respondent. In regards to interviewer effects, they found that familiarity with the interviewer depressed drug use reporting, and they explain this effect as being a product of the increased salience of normative standards in the context of past and expected future interactions with the same individual.

The Fendrich et al. (1999) study evaluated the importance of interviewer and subject effects on drug use disclosure in a sample of juvenile arrestees. This study found that interviewer effects were more salient in models predicting marijuana disclosure than in models predicting cocaine disclosure, subject race/ethnicity effects were salient for both drugs, and large interviewer cluster effects were present. Fendrich et al. (1999) adapted the Social Attribution Model (SAM) to what they call the “Conditional Social Attribution Model”(CSAM) which posits that, in addition to interviewer characteristics having direct effects on respondent answers (the SAM), judgments about interviewers are going to be influenced by respondent characteristics. This model suggests that survey results can be affected by interactions between interviewer and respondent characteristics. Findings from this study lack generalizability, though, because of the highly sensitive nature of the survey content. According to these authors, survey reports of drug use behavior are difficult because people have strong disincentives to provide
accurate information and because the behaviors being reported are associated with interviewer characteristics, opinions and beliefs in diverse ways.

A 2001 study by Heeb and Gmel on self-reported alcohol consumption built on the theoretical advances of Fendrich et al. (1999) by examining the effects of gender and age of interviewers and respondents simultaneously (per the CSAM) using a hierarchical linear model (HLM) to account for individual interviewer “clustering” effects. Clustering effects are one way to describe the impact that interviewers have on their respective samples that can differ from interviewer to interviewer. Clustering effects would suggest that some interviewers cause more characteristic effects than other interviewers. In addition to the HLM, the authors used other statistical tools: a two-way ANOVA, separately for gender and age; t-tests on fixed coefficients; and variances using chi-square. This study found a main effect for gender of respondents and for age of respondents, an interaction effect between interviewer and respondent ages, and a strong interviewer clustering “design” effect that, if not taken into account, could result in errors in significance tests and too-small confidence intervals.

In contrast, a 2003 study of interviewer and respondent race and sex similarity effects in highly structured interviews found no evidence for effects when contrasting ANOVA with the HLM approach (Sacco, Scheu, Ryan and Schmitt, 2003). The authors of this study proposed that the absence of effect might be due to sensitization to diversity issues, the highly structured nature of the interview, and the extensive interviewer training and experience. While modeling and testing for interviewer effects across disciplines has become more precise, the literature suggests the potential for biases that require further research in diverse contexts.
Interviewer bias in the context of contingent valuation studies

It has been fairly well accepted that the best approach to collecting preference information in CVM studies is to perform in-person surveys, at the respondent’s household if possible (Mitchell and Carson, 1989; Arrow, et al., 1993). However, considering the primacy of the in-person household survey approach, and extensive literature supporting interviewer effects in public health and political survey contexts, there is a need for economists to develop a better understanding of the influences of interviewers on response-editing in CV surveys.

Some studies have examined mode effects in contingent valuation surveys (Leggett, Kleckner, Boyle, Duffield and Mitchell, 2003; Davis, 2004). Leggett et al. (2003) provide empirical evidence to support a social desirability bias theory, finding that in-person surveys obtain willingness-to-pay measures that were 23 percent higher than measures obtained through self-administered surveys. They also find higher WTP with face-to-face surveys than with self-administered surveys. The authors used a non-parametric, two-way contingency table as well as a multivariate logit function to test for statistical differences between the two types of survey administration. However, their study did not go beyond general mode effects to explore the nature and magnitude of influence that interviewer and respondent characteristics may have had on stated preferences.

A recent study by Loureiro and Lotade in 2005 analyzed how the presence of two different interviewers affected people’s willingness-to-pay (WTP) for fair-trade, shaded and organic eco-labeled coffee. They found a relationship between the interviewer and WTP responses for the fair trade and shaded coffees, with the African interviewer
eliciting higher WTP than the American interviewer. The authors attributed this difference to the consumer’s association of the interviewer with the coffee’s country of origin. While these results supported the sensitivity of WTP values to interviewers generally in face-to-face surveys, the authors could not determine what characteristics of the interviewers produced the difference in WTP. In addition, as Feldman (1952) noted in regard to early studies, “too much may hinge on the accidental employment of one aberrant interviewer.” (p.735)

In sum, a literature review across the fields of public health, economics, and opinion survey research suggests that while interviewer effects have been observed and discussed in the past, the relative magnitude and direction of these effects and important intervening variables that contribute to the process are still largely unexplored. Further, many survey studies today still neglect to measure or even mention the interviewer as a source of error. Similar vaccine demand studies have been conducted to measure demand for hypothetical vaccines against diseases such as malaria (Cropper, Haile, Lampietti, Poulos and Whittington, 2004), HIV/AIDS (Whittington, Matsui, Frieberger, Houven and Pattanayak, 2002) and typhoid fever (Canh, et al., 2005). But only a few papers were found that examine interviewer effects on CVM studies (a few more mention the issue briefly but do not examine it). To this author’s knowledge there are no studies that systematically examine interviewer effects on responses in CVM studies conducted 1) on vaccines, 2) outside the U.S. or 3) across related CVM studies conducted in different sites.

This paper contributes to the literature on interviewer effects in general and contingent valuation studies in particular by examining the interaction effects of
interviewer and respondent characteristics on findings from three CVM studies of WTP for cholera vaccines conducted in India, Bangladesh and Mozambique. All three studies addressed the same basic research questions, used similar interviewer training protocols, administered similar questionnaires, and were guided by the same team of international experts for evaluation and feedback. The analysis focused on: 1) interviewer gender, 2) respondent characteristics of gender, age and education level, and 3) interviewer influence on respondents through unobserved characteristics, preferences or actions, a factor generally called interviewer variance or “whole interviewer effects”.
3. CVM Study Sites and Data Description

The three studies that are the subject of this comparative analysis are part of a larger initiative of the International Vaccine Institute’s Diseases of the Most Impoverished (DOMI) program, funded by the Bill and Melinda Gates Foundation. The DOMI program, initiated in 2000, was designed to support the development and introduction of new-generation vaccines in developing countries; the program has epidemiological, socio-behavioral, technological and economic research components. The sub-series of economic studies attempts to measure private demand for vaccines in developing countries through the use of the contingent valuation method. Participating countries include Pakistan, China, Vietnam, Indonesia, India, Bangladesh, and Mozambique (Whittington, et al., 2005; Cook, Whittington, Canh, Johnson, and Nyamete, 2005). This analysis will focus on the India, Bangladesh, and Mozambique CVM studies conducted in the summers of 2004 and 2005. The brief descriptions of the study sites are adapted from unpublished materials with the permission of each research team.

**Kolkata, India**

The India study was conducted in two neighborhoods in Kolkata, the third largest city in India (a city-wide population of about 13 million). Tiljala is a densely-populated, mostly low-income slum. Beliaghata has more diverse living conditions and incomes. Tiljala is predominately Muslim while Beliaghata is mostly Hindu. Beliaghata contains many middle-class families living in apartment buildings that are in relatively good condition, but has several small slums with living conditions similar to those in Tiljala.
Interviewer Selection Process

The local partner organization, National Institute of Cholera and Enteric Disease (NICED) handled recruiting interviewers in Kolkata. There was no known formal announcement or advertisement for positions, but rather word-of-mouth was used. Many of these interviewers were college students studying health policy and management and had heard about the position through one of their professors. The research team in Kolkata was generally not in a position to fire people unless there were serious problems. Nevertheless, the project managers kept an eye on interviewers during training and gave them “tests” in order to assess how hard they were working and how well they knew the questionnaire. All the interviewers did fine but one. Despite warning her several times about serious errors, the issue was brought to NICED and it was requested that she be moved off the project.

Training began with 15 enumerators. Data collection began with 13 after one was fired and another quit at the end of training, and two additional interviewers left during the fieldwork itself. The study in Kolkata also had a team of 15 “field supervisors” to find households and set up interviews, so as interviewers were lost to attrition, the best of these supervisors were trained to take over as enumerators.

When asked to subjectively evaluate the quality of the interviewer team, the lead investigator at the Kolkata site stated he believed that the interviewers were an intelligent group of students. As they finished pre-testing and began data collection, he felt they all knew the questionnaire very well and could conduct a smooth, pleasant interview. The main challenge was keeping them motivated and working hard.
**Matlab, Bangladesh**

The Bangladesh study centered at Matlab, a rural town 55 kilometers southwest of Dhaka. Matlab has a population of approximately 224,000 people. The International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) operates a hospital in Matlab that provides free treatment to anyone with diarrhea. The use of hand pumps is common but problematic because of extensive arsenic contamination in the Matlab area. Due to arsenic contamination, villagers in the Matlab area have been advised not to boil water, putting them at greater risk for cholera transmission and exposure.

*IInterviewer Selection Process*

At the Bangladesh site interviewers were screened and hired by ICDDR,B based on their past experience doing similar work on health surveys. It was requested that more female than male interviewers be hired for this research, based on a sense from the literature that women make better interviewers than men. Of 60 interviewers interested in the job, 20 were selected. Training did not play a role in the interviewer selection process; the study kept all of the interviewers that were trained. There were 20 interviewers in total, all of whom remained for the duration of the study.

In his subjective evaluation, the lead investigator at this study site stated he felt he had a pretty good team of interviewers. The interviewers all had previous experience working on health-based questionnaires. They also wanted to work with ICDDR,B in the future. The lead investigator reported that no cheating was detected during field supervision. The main challenge with survey delivery was making sure that interviewers stuck to the script.
**Beira, Mozambique**

Beira is in the Sofala province of Mozambique. It is the second largest city in the country, with twenty-two *bairros* (neighborhoods) and a population of approximately 450,000. The official language in Mozambique is Portuguese but many people in Beira speak one or more unwritten local languages, such as Ndau and Sena. The sanitation conditions observed among households foster conditions for cholera: half of respondents have simple pit latrines, and about fifteen percent defecate in the open.

*Interviewer Selection Process*

In Beira interviewers were recruited by the partner agency, Centro de Higiene Ambiental e Exames Médicos (CHAEM), through a collection of resumes of individuals used for an earlier data collection process. Some of these interviewers, therefore, had previous experience. This group of interviewers was primarily college-age, some who had not gone to college and others who were enrolled students on break. Some of them were the students of the local economist working on study.

The training program was the primary way interviewers were chosen for the project. Training began with 26 candidates, of which 21 were retained for the project. Of those retained for the project, three were chosen to be interview set-up personnel rather than interviewers. Part way into the study a fourth interview set-up person was hired to balance their work load and allow for better coordination across field locations.

The largest challenge with data collection in the field was ensuring the close connection between the household sampling done by interview set-up personnel and the actual interviews that were completed. Unlike other study sites, in Beira the lack of census data or voting registration records made it impossible to perform a random sample
at the household or individual level. The randomized sampling process extended through the *bairro* and *unidade* levels (larger and smaller neighborhood units, respectively), after which set-up personnel were trained to sample every fifth household. This technique adds both difficulty for sampling and the challenge that interviewers in the field were tempted to pressure setup personnel to “choose” households that were closer, or had people at home. In Beira, unlike Matlab and Kolkata, interviews were conducted in pairs for safety reasons, though in each case there was a primary interviewer responsible for asking the questions and recording the responses. This interviewer is the one attributed to the interview in the coding of the dataset.

Overall, the investigators for the Beira study feel their interviewing team was capable, though perhaps not as consistent and precise in their technique as interviewing teams on other research projects. Some interviewers were more professional and seemed to take their jobs more seriously than others.
### Table 1. Summary of field sites and survey protocol in three countries

#### Overview of Comparison Study Sites and CV Methods

<table>
<thead>
<tr>
<th>Setting</th>
<th>Kolkata, India</th>
<th>Matlab, Bangladesh</th>
<th>Beira, Mozambique</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban, two neighborhoods, one very poor and one middle class</td>
<td>Rural</td>
<td>Mostly urban, some semi-urban</td>
</tr>
<tr>
<td>Respondents</td>
<td>835 respondents; Mother or father of children in hh under 18</td>
<td>591 respondents; Mother or father of children in hh under 18</td>
<td>996 respondents; Mother or father of children in hh under 18</td>
</tr>
<tr>
<td>Sampling protocol</td>
<td>Beliaghata: Voter list; Tiljala: NICED census; HHs with kids&lt;18</td>
<td>ICDDRBR Health and Demographic Surveillance System database; HHs with kids &lt;18</td>
<td>Weighted stratified sample atbairro level, random sample of unidade s, &quot;every fifth house&quot; at household level; HHs with kids &lt; 18, understands Portuguese (a few in Sena or Ndau)</td>
</tr>
<tr>
<td>Attributes of the vaccine offered in the scenario</td>
<td>Cholera 50%, 3 yr, no variation of attributes</td>
<td>Cholera 50%, 3 yr, no variation of attributes</td>
<td>Cholera 3 years, generic effectiveness description (excellent effectiveness in the first year, diminishing effectiveness in the second and third years)</td>
</tr>
<tr>
<td>CV design (unadjusted prices)</td>
<td>13/11 interviewers 4 prices: USD 0.22, 0.56, 1.11, 11.11</td>
<td>20 interviewers; 6 prices: USD 0.15, 0.37, 0.74, 1.12, 4.46, 8.93</td>
<td>18 interviewers; 5 prices: USD 0.20, 0.82, 1.63, 2.86, 4.08</td>
</tr>
</tbody>
</table>

All three studies conducted a contingent valuation survey designed to measure private willingness to pay for vaccines. The Mozambique and Bangladesh sites measured willingness to pay for cholera vaccines; the India site measured willingness to pay for both cholera and typhoid vaccines. This paper uses only the respondent willingness to pay results for cholera vaccines.

An important basis for comparison across these three study sites is that the interviews were conducted using the same basic questionnaire, though it was refined at each site according to cultural, social, economic and linguistic context. After translating the questionnaires from English into the preferred local language, the research teams pre-
tested the questionnaires and made adjustments to specific questions that confused people or were not relevant. For example, when interviewers asked respondents about their assets, they do so from a predetermined list that is designed to range from assets almost everyone should have to assets that only the most affluent households would have. This list was adjusted at each country site based on the most common assets for that community.

The contingent valuation section of the questionnaires remained fundamentally the same across all study sites, with minor variations. All sites used a household question that asked respondents how many vaccines they would purchase for members of their household, as well as whom specifically they were purchasing for. The India and Bangladesh sites also used a question that elicited the respondent’s willingness to pay for a vaccine for him or herself only. In Mozambique, pre-testing found this question to be problematic, so respondent demand was determined by looking at the table listing for whom vaccines were purchased to see if the respondent was included in the list. Since there is nothing in the literature or the author’s experience in this research to suggest that household demand and respondent demand would show different interviewer effects, this analysis uses as the dependent variable only respondent demand results in the form of bivariate (yes/no) responses.
4. Methodology

*Research questions*

The analyses at each site were designed to answer the following specific questions about the relationship between interviewer characteristics and respondent willingness to pay for cholera vaccines:

**Direct Effects/Social Attribution Models:**

1. *Do any specific interviewers generate systematically different results for respondent willingness to pay as compared to the other interviewers at the site?* – This question is examined by including individual interviewer dummy variables in the basic regression model. For each study site the investigator was instructed to drop the dummy variable for that interviewer he thought most “average” in performance (based on other statistics like mean interview time, price given, interview workload, etc), so the measurements of variance the rest of the interviewers show are variance from this average interviewer.

2. *Are respondents more or less likely to say they are willing to pay for a cholera vaccine for themselves if the interviewer is male instead of female?* – This question is examined through the use of a basic dummy variable for interviewer gender (male interviewer = 1).

**Indirect Effects/Conditional Social Attribution Models:**

3. *Is there a difference in female respondent willingness to pay when interviewed by male interviewers as compared to female interviewers?*
a. Likewise, is there a difference in male respondent willingness to pay when interviewed by male interviewers as compared to female interviewers?

– These questions are examined through the use of the four way interaction variable for respondent and interviewer gender.

4. Are older respondents more or less likely to say they are willing to pay for a cholera vaccine for themselves when interviewed by male interviewers as compared to female interviewers?

a. And, are younger respondents more or less likely to say they are willing to pay for a cholera vaccine for themselves when interviewed by male interviewers as compared to female interviewers?

– These questions are examined through the use of the four way interaction variable for respondent age and interviewer gender.

5. Are respondents with less education more or less likely to say they are willing to pay for a cholera vaccine for themselves when interviewed by male interviewers as compared to female interviewers?

a. Are respondents with more education more or less likely to say they are willing to pay for a cholera vaccine for themselves when interviewed by male interviewers as compared to female interviewers?

– These questions are examined through the use of the four way interaction variable for respondent education level and interviewer gender.
Modeling approach

Modeling of respondent demand in this paper is based on the models from the original CVM analyses (working papers, or submitted for publication). In those studies the household decision maker chooses whether or not to purchase a two-dose vaccine in order to maximize his/her utility function, which is a function of consumption and health and is subject to the household budget constraint. This yields a demand function for immunization $A$. In the CVM analyses demand depends on vectors of household members’ health ($\mathbf{H}$) and other characteristics ($\mathbf{Z}$) (e.g., sex, birth order, education), as well as total income ($y$) and the prices for preventive ($p_v$) and mitigating ($p_m$) health goods:

$$A = g(y, p_v, p_m, \mathbf{Z}, \mathbf{H})$$  \hspace{1cm} (1)

For the purposes of this paper respondent demand for a cholera vaccine is modeled as a function of these socio-demographic characteristics of the respondent and respondent’s household, with the addition of interviewer characteristics ($\mathbf{I}$) and interviewer and respondent characteristic interactions ($\mathbf{C}$), yielding an adjusted model that includes these variables:

$$A = g(y, p_v, p_m, \mathbf{Z}, \mathbf{H}, \mathbf{I}, \mathbf{C})$$  \hspace{1cm} (2)

Interviewer characteristics used include gender and a “whole interviewer effect” modeled by including individual interviewer dummy variables in the regression. The whole interviewer effect refers to the amount of variation in responses that is explained by individual interviewers after measuring the variation in response that is explained by the measurable characteristics of those interviewers, in this case gender.
The education and age categories for the interactions denote, respectively, higher and lower levels of education, and higher and lower ages. In the interaction models we created 4-way dummy variables representing each of four possible interactions between interviewer gender and the respondent characteristic of interest.

Table 2. Four-way dummy interaction variables

<table>
<thead>
<tr>
<th>Respondent characteristics</th>
<th>Male Interviewer</th>
<th>Female Interviewer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent Male</td>
<td>Interviewer Male</td>
<td>Respondent Male</td>
</tr>
<tr>
<td>Female</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent Female</td>
<td>Interviewer Female Respondent Male</td>
<td></td>
</tr>
<tr>
<td><strong>2. Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older (Age&gt;=35)</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent Older</td>
<td>Interviewer Male</td>
<td>Respondent Older</td>
</tr>
<tr>
<td>Younger (Age&lt;=35)</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent Younger</td>
<td>Interviewer Male</td>
<td>Respondent Younger</td>
</tr>
<tr>
<td><strong>3. Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent High Edu</td>
<td>Interviewer Male</td>
<td>Respondent High Edu</td>
</tr>
<tr>
<td>Low</td>
<td>Interviewer Male</td>
<td>Interviewer Female</td>
</tr>
<tr>
<td>Respondent Low Edu</td>
<td>Interviewer Male</td>
<td>Respondent Low Edu</td>
</tr>
</tbody>
</table>

In order to test the three 4-way interaction models, the regressions are run to isolate the impact of interviewer gender on one respondent variable value at a time (e.g. interviewer gender on respondents with high education, then in the next regression, interviewer gender on respondents with low education), resulting in six different regressions in total that address interaction effects. When running the models we dropped one of the dummies as the baseline for comparison, so when results are reported, significant differences are reported as they differ from the specific baseline case that was dropped from the regression.

Due to the relatively large number of interviewers, the procedures for allocating respondents to interviewers at each site, and the lack of highly gendered subject matter in the questionnaires, the author expects that interviewer effects will be smaller than those described in previous studies of interviewer effects in survey research. The author
expects that some cases of observed direct effects of interviewer gender might be partially explained by the presence of indirect effects of interviewer gender interacting with specific respondent characteristics. In general, the lead investigators at each site expect to see some individual interviewer effects, but limited if any interviewer-respondent characteristic interaction effects. The lead investigator of the Bangladesh site proposed that we might see a bias toward yes responses by male respondents when being interviewed by female interviewers.

Other General Methods

The lead investigators of each site used the computer software STATA 9 to estimate probit regressions on their datasets. A short set of questions was posed to each of the principal investigators before running the analysis. These questions were designed to gather basic impressions of characteristics of the study site and the interviewers, brief descriptions of the interviewer training process at each site, and preliminary opinions about what the interviewer effects analysis would show for each site. A follow-up dialog occurred that allowed all the investigators to exchange ideas about how to best model these relationships, and what the results mean for their data. This dialogue feeds into the closing discussion of site-specific results, as well as how the sites compare with one another in terms of more general study implementation choices and challenges.
<table>
<thead>
<tr>
<th>Variable Description</th>
<th>Variable</th>
<th>Beliaghata</th>
<th>Tiljala</th>
<th>Matlab, Bangladesh</th>
<th>Beira, Mozambique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviewer Gender, male = 1</td>
<td>intgen</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Length of interview</td>
<td></td>
<td>43</td>
<td>48</td>
<td>39.4</td>
<td>39.4</td>
</tr>
<tr>
<td>If given time to think, = 1</td>
<td>ttt</td>
<td>0.523</td>
<td>0.467</td>
<td>0.479</td>
<td>0.47</td>
</tr>
<tr>
<td>Gender of respondent, male = 1</td>
<td>male</td>
<td>0.52</td>
<td>0.511</td>
<td>0.718</td>
<td>0.125</td>
</tr>
<tr>
<td>Age of respondent (yrs), continuous</td>
<td>age</td>
<td>34.8</td>
<td>35</td>
<td>37.1</td>
<td>32.1</td>
</tr>
<tr>
<td>Number of people in household</td>
<td>hhcount</td>
<td>5.04</td>
<td>5.29</td>
<td>5.07</td>
<td>5.78</td>
</tr>
<tr>
<td>Respondents with high education</td>
<td>education</td>
<td>0.434</td>
<td>0.5</td>
<td>0.178</td>
<td>0.125</td>
</tr>
<tr>
<td>Household income</td>
<td>hhinc2*</td>
<td>0.206</td>
<td>0.17</td>
<td>0.239</td>
<td>0.259</td>
</tr>
<tr>
<td>Household income</td>
<td>hhinc3*</td>
<td>0.288</td>
<td>0.296</td>
<td>0.178</td>
<td>0.161</td>
</tr>
<tr>
<td>Household income</td>
<td>hhinc4*</td>
<td>0.292</td>
<td>0.33</td>
<td>0.135</td>
<td>0.0536</td>
</tr>
<tr>
<td>Lives in ICDDR,B service area</td>
<td>icddrb</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>0.225</td>
</tr>
<tr>
<td>Has soap in the house today = 1</td>
<td>soap</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>Na</td>
</tr>
<tr>
<td>Has assets in high value group = 1</td>
<td>asset4</td>
<td>0.0842</td>
<td>0.066</td>
<td>0.0842</td>
<td>0.066</td>
</tr>
<tr>
<td>Neighborhood, Tiljala = 1</td>
<td>tiljala</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

*The numbers for household income for Beira were predicted because of underreporting.

*Not adjusted for purchasing power parity.
5. Results

Kolkata, India

In the Kolkata study each interviewer was responsible for between 2.78 and 8.47 percent of all interviews. This wide range is due to the fact that two interviewers quit before the data collection was complete, and they were replaced with new interviewers. Three interviewers were more likely to have respondents say they would buy a vaccine for themselves than the average interviewer, at a significance level of 5 percent. These interviewers had comparable workloads to the rest of their team. Two other interviewers show an individual effect at the 10 percent level. When the influence of interviewer gender on respondent willingness to pay was tested with gender and interaction models, interviewer gender was not found to be significant for respondents in general nor for respondents of specific genders, ages or education levels.

Beira, Mozambique

In the Mozambique study each interviewer was responsible for between 5.32 and 7.33 percent of all interviews. The interviewer who only completed 5.32 percent was taken out of the field toward the end of the study to help with translations. One male interviewer was more likely to have respondents say they would buy a vaccine for themselves than the average interviewer, at a significance level of 5 percent. Three other interviewers show an individual effect at the 10 percent level. When the influence of interviewer gender on respondent willingness to pay was tested with gender and interaction models, interviewer gender was not found to be significant for respondents in general nor for respondents of specific genders, ages or education levels.
Matlab, Bangladesh

Interviewer workload

In the Matlab study each interviewer was responsible for between 4.74 and 5.25 percent of all interviews. One interviewer had a large impact on the presence of interaction effects, which is a different result than the whole interviewer effects found in the other two study sites. Included below are the results with the whole sample, and results with that interviewer’s sample removed from the analysis.

Basic gender effect

Using the full sample, respondents in the Matlab sample are 39 percent less likely to say they will buy a cholera vaccine for themselves when interviewed by a male rather than a female interviewer. This finding is statistically significant at the 1 percent level.

Gender-gender interactions

Examining the gender effect more closely, there is a strong statistical difference between situations in which men are interviewed by women as compared to men. Specifically, male respondents appear to be about 45 percent less likely to say they will buy a vaccine for themselves when interviewed by a male rather than a female interviewer (5 percent level). There is less statistical significance (10 percent level) in the difference between cases in which women are interviewed by women as compared to men, but female respondents are also about 36 percent less likely to say they will buy a vaccine when interviewed by male rather than female interviewers.

Gender-age interactions
Using similar 4-way dummy interaction variables to examine the combination of interviewer gender and respondent age, the Bangladesh data show that older respondents are 61 percent more likely to say they will buy a vaccine for themselves when interviewed by female interviewers as compared to male interviewers (p=.001). However, in the same model interviewer gender does not appear statistically significant for younger respondents (p = .296).

Gender-education interactions

In interactions between interviewer gender and high or low levels of respondent education, we see again that male interviewers elicit lower rates of willingness to pay for cholera vaccines by respondents regardless of respondent education level. In these models the coefficients show the increase in the likelihood respondents buy vaccines when the interviewer is a female as compared to when the interviewer is a male. The coefficient is larger (.597 v .305) and has greater significance (1 percent v 5 percent) for the higher education group than for the lower, suggesting that the differential impact of interviewer gender on respondent willingness to buy vaccines increases with education.

Whole interviewer effects

When interviewers are included as separate variables in the regression, one interviewer had significantly lower rates of respondent willingness to pay for cholera vaccines as compared to the average interviewer, at a significance level of 1 percent. When the respondent sample from this interviewer is excluded from the models, the only interaction that remains significant at the 5 percent level is interviewer gender and high respondent education, for which female interviewers have a positive effect on respondent willingness to pay as compared to male interviewers.
Table 4. Summary of significant variables

**Significant Results by Study Site**

<table>
<thead>
<tr>
<th>Site</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Matlab, Bangladesh</strong></td>
<td>Interviewer gender</td>
<td>-.387***</td>
</tr>
<tr>
<td></td>
<td>IMRM</td>
<td>-.453**</td>
</tr>
<tr>
<td></td>
<td>IMRF</td>
<td>-.356*</td>
</tr>
<tr>
<td></td>
<td>IFRold</td>
<td>.610***</td>
</tr>
<tr>
<td></td>
<td>IFRloed</td>
<td>.305**</td>
</tr>
<tr>
<td></td>
<td>IFRhied</td>
<td>.597***</td>
</tr>
<tr>
<td></td>
<td>Interviewer 16</td>
<td>-1.28***</td>
</tr>
<tr>
<td><strong>Kolkata, India whole sample</strong></td>
<td>Interviewer 7</td>
<td>.655*</td>
</tr>
<tr>
<td></td>
<td>Interviewer 8</td>
<td>.799*</td>
</tr>
<tr>
<td></td>
<td>Interviewer 16</td>
<td>.849**</td>
</tr>
<tr>
<td></td>
<td>Interviewer 17</td>
<td>1.51***</td>
</tr>
<tr>
<td></td>
<td>Interviewer 18</td>
<td>.880**</td>
</tr>
<tr>
<td></td>
<td>Interviewer 3</td>
<td>.706*</td>
</tr>
<tr>
<td></td>
<td>Interviewer 8</td>
<td>.759*</td>
</tr>
<tr>
<td><strong>Beliaghata only</strong></td>
<td>Interviewer 3</td>
<td>.487*</td>
</tr>
<tr>
<td></td>
<td>Interviewer 6</td>
<td>.593**</td>
</tr>
<tr>
<td></td>
<td>Interviewer 13</td>
<td>.478*</td>
</tr>
<tr>
<td></td>
<td>Interviewer 16</td>
<td>.467*</td>
</tr>
</tbody>
</table>

* Indicates significance at the 10% level

** at the 5% level

*** at the 1% level
6. Discussion

A focus on Bangladesh

In the Bangladesh full sample there are significant interviewer effects that could not have been controlled for without first observing them using these tests. In the full sample interviewer gender is important to all respondents, but the results of the interaction tests suggest gender is more important to male respondents than to female respondents, with male respondents giving significantly fewer yes responses to male interviewer than to female interviewers. Older respondents were significantly more likely to give a yes response to a female interviewer than a male interviewer. Interviewer gender is also important for respondents regardless of education level, but in this sample respondents with more education were more likely to give yes responses to female interviewers than respondents with lower education.

These interviewer effects seem quite problematic until one conducts a test of significance for whole interviewer effects. This test showed one interviewer out of twenty obtained significantly fewer yes responses than the “average” interviewer on that team. In the sample that excludes this one interviewer’s questionable results (Interviewer 16), the statistical tests show that only interviewer gender interacted with respondent high education remains significant.

It is not obvious why in the Bangladesh sample one interviewer was able to cause such specific and strong interaction effects to appear. In the other studies whole interviewer effects did not have this influence. On the other hand, in those studies the effects were distributed over more than one interviewer, and individually, each interviewer’s effect was not as significant as the one in Bangladesh.
Table 5. Matlab Interviewer Workload Statistics

<table>
<thead>
<tr>
<th>Interviewer</th>
<th>Number of Interviews</th>
<th>Percent of Total Workload</th>
<th>Average Price(Tk) Offered</th>
<th>Percent Yes Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28</td>
<td>4.74</td>
<td>125</td>
<td>0.536</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>4.91</td>
<td>161</td>
<td>0.464</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>5.08</td>
<td>185</td>
<td>0.500</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>5.08</td>
<td>181</td>
<td>0.567</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>5.08</td>
<td>159</td>
<td>0.333</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>5.25</td>
<td>175</td>
<td>0.323</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>5.08</td>
<td>170</td>
<td>0.533</td>
</tr>
<tr>
<td>8</td>
<td>30</td>
<td>5.08</td>
<td>104</td>
<td>0.700</td>
</tr>
<tr>
<td>9</td>
<td>29</td>
<td>4.91</td>
<td>213</td>
<td>0.345</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>5.08</td>
<td>104</td>
<td>0.733</td>
</tr>
<tr>
<td>11</td>
<td>31</td>
<td>5.25</td>
<td>139</td>
<td>0.419</td>
</tr>
<tr>
<td>12</td>
<td>29</td>
<td>4.91</td>
<td>168</td>
<td>0.517</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>5.08</td>
<td>193</td>
<td>0.467</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
<td>5.08</td>
<td>163</td>
<td>0.552</td>
</tr>
<tr>
<td>15</td>
<td>28</td>
<td>4.74</td>
<td>125</td>
<td>0.464</td>
</tr>
<tr>
<td><strong>16</strong></td>
<td><strong>28</strong></td>
<td><strong>4.74</strong></td>
<td><strong>210</strong></td>
<td><strong>0.125</strong></td>
</tr>
<tr>
<td>17</td>
<td>31</td>
<td>5.25</td>
<td>121</td>
<td>0.323</td>
</tr>
<tr>
<td>18</td>
<td>29</td>
<td>4.91</td>
<td>165</td>
<td>0.571</td>
</tr>
<tr>
<td>19</td>
<td>29</td>
<td>4.91</td>
<td>100</td>
<td>0.586</td>
</tr>
<tr>
<td>20</td>
<td>29</td>
<td>4.91</td>
<td>174</td>
<td>0.310</td>
</tr>
</tbody>
</table>

From some descriptive statistics we see the interviewer in Bangladesh did not complete a larger portion of the interviews than any other interviewer at the site, but he did have a much lower rate of obtaining positive WTP from respondents than other interviewers. This interviewer also did not receive a larger than average number of high prices to offer (through the random allocation of prices to questionnaires) which would have been able to explain a higher than average rate of no responses.

The fact that one interviewer explains so many interviewer gender effects in the Bangladesh sample suggests that rather than whole interviewer effects being the product of some unknown or unquantifiable aspect of that interviewer’s personality or presence, the unique impact of that interviewer could sometimes be closely tied to predictable and measurable interviewer or respondent characteristics, in this case gender.
7. Conclusions from this study

The results of this analysis indicate that interviewer gender effects can be significant for CV studies, and they are sometimes explained by the influence of specific interviewers in particular. This influence can be measured and controlled for statistically. Significant interviewer gender effects were only seen in one of the three study sites. At the Kolkata and Beira study sites whole interviewer effects were evident, but they did not result in interviewer gender effects specifically, as they did in Bangladesh.

These results lead me to some tentative conclusions:

- It is likely that employing larger numbers of interviewers will dissipate the influence of one or a few biased interviewers, and it will make it easier to remove their data from the analysis without jeopardizing the sample size and statistical significance of the results.

- Another method for controlling interviewer effects used in the Beira study was to pair interviewers and mix up the pairs periodically – this may have mitigated both whole interviewer direct effects and gender effects in the cases of those pairs with both a male and a female interviewer.

- It is possible that similar and systematic training procedures are not going to prevent interviewer effects in all cases. This could be because it is hard to know what the effect of an individual will be on survey results, and it could also be because interviewer training in the data collection process is not necessarily going to change the social importance of certain interviewer or respondent characteristics. Good training may be able to mitigate the influence of open biases the interviewers may have, but in the absence of conscious biases that can take the
form of response expectations and social desirability biases, gender or other characteristic effects may remain.

- When cultural information guides us to implement surveys a certain way, for instance using men to speak with men and women to speak with women, it is important that the researcher understand how these choices can systematically influence the results of the survey.

- Multivariate analyses can be done to check for interviewer effects and control for them after the data is collected.

In the larger context of social science research for public health and welfare, in-person surveys are a very popular tool to allow the public to inform decision-making. They are systematic, consistent, and fairly simple to administer and analyze. This is as compared to other community research techniques like community meetings, focus groups, and participatory research that tend to require highly specialized skill sets and a longer-term commitment to data collection, reflection, and revision of the procedures. No matter which technique is used, however, it is critical that the voice of the public is accurately portrayed. Analyses like this help researchers better understand the significance of collection methods in the conclusions that data will suggest.
8. The survey research context: Other considerations

Barriers to cholera vaccination in developing countries

This paper focused on one way that researchers can improve their data analyses such that their results best reflect the population’s preferences concerning cholera vaccines. That way is through statistical control of known variables that introduce bias and variance in responses to willingness to pay questions. This focus, however, is many steps removed from the logistical challenge of actually vaccinating people against cholera in developing nations. Before the research task is even posed to investigators, other considerations are at play.

For individuals, a possible barrier to acceptability of a cholera vaccination program involves the disease cycle. While cholera is endemic in some regions, there are epidemic outbreaks that occur periodically resulting in exponential increases in incidence rates. One might expect that cholera vaccination would seem more urgent and important to individuals and families when there is an outbreak and more people are sick than in between outbreaks when few people suffer from the disease.

However, there is less evidence in the data than expected that demand for cholera vaccines is affected in one way or the other by prior household experience with the disease or death from the disease. In the case of Beira, demand was also not correlated with an area’s incidence rates of diarrheal disease in general. In Matlab respondents who reported believing they would contract cholera and believing it was a serious disease were more likely to say they would purchase the vaccine, but respondents who actually knew someone in the family who had cholera were no more likely to purchase a vaccine than the average person.
What else influences people’s willingness to pay for cholera vaccines? In Beira it seemed to be the distance to vaccination locations, the price of the vaccine, respondent education and wealth, and the time respondents were given to think about their answers to the question. In Kolkata price, respondent income, respondent education, risk aversive behavior and perceived severity and prevalence all influenced demand for cholera vaccines. In Bangladesh price, time to think, income, access to credit, age, perceived severity and perceived susceptibility were found to influence demand for cholera vaccines. Each of these variables can be considered a barrier or facilitator depending on the value the variable takes and the vaccine distribution context. For instance, one could consider a poor education system a barrier to acceptability of cholera vaccines in so much as lower education reduces individual’s willingness to buy vaccines for themselves or their families.

Another individual level barrier to acceptability of cholera vaccines could be perceptions of legitimacy and safety. Some respondents in the Beira sample expressed the perception that the vaccine trial in 2004 and 2005 was an experiment on the population, that the vaccine itself was “experimental” or not safe. For these respondents this perception seemed to carry over into their interviews about willingness to pay for a hypothetical cholera vaccine, and these respondents often did not wish to pay anything to receive the vaccine.

Population level or policy level barriers to access

From a community wide or policy level perspective, there are slightly different barriers to access to cholera vaccines. Physical accessibility can be considered a community level barrier in the sense that decisions about the distribution of vaccines will
likely impact which populations receive the vaccine and which do not. The timing of vaccination campaigns also produces barriers to access, particularly if, as in the case of cholera vaccine used in these CV surveys, respondents have to make two separate trips to receive both doses of the vaccine.

There are also other health priorities that demand attention of the ministries or offices of health in developing nations. The most well-known of those priorities is HIV and AIDS which clearly exacts a larger toll on the population of these nations than cholera does. Therefore, while outbreaks of cholera can directly impact daily life for many people, finite health resources must be brought to bear on a seemingly endless list of health concerns for these communities. Where cholera falls in terms of importance will probably vary from place to place and from year to year.

Another policy level barrier to access to cholera vaccines is the conflicting responsibilities governments have toward their constituencies. While a government may hold as a fundamental goal reducing disease related morbidity and mortality in their country, they also must respond to public expectations that basic healthcare and protection from disease is a right and should be provided to all citizens at no charge. Though “no charge” is a simplistic portrayal of public financing, this is a principle held by many governments across the world. Political and cultural norms thus curb experimentation with alternative provision mechanisms and policies that might be shown effective or feasible by research like that presented here (e.g. fee for service).

Almost all literature concerning the design and distribution of vaccines in developing countries assumes that vaccines should be provided for free, usually through the financial support of international donors. The EPI vaccine bundle for children under
one was implemented in that way, and became very successful in some areas. Research is now exploring how to provide next-generation vaccines that are not included in that bundle, particularly vaccines for adults, older children, or vaccines that must be taken more than once in a lifetime. The research, however, still rests on the assumption that all vaccines should be provided for free and (therefore) that developing nations will need to partner with donor agencies in order to afford to run the programs. The international financial support is not intended to be a long-term solution, but these programs are rarely self-sustaining when the international money runs out.

An economic argument suggests that where there is demand for vaccines, there may also be willingness and ability to pay for vaccines. If this is so, governments may find they do not need to pay all of the expense of a mass vaccination campaign. Rather, the expense could be shared in some fashion between the government and the people. There are immediate and clear equity issues involved with putting a price on vaccines; only those who can afford them will buy them and gain direct benefits. Some research, though, suggests that even when many people in a community do not get vaccinated, if enough people do, a “herd immunity” effect will confer benefits on many others (Ali, et al., 2005). To increase political and community support for fee-for-service cholera vaccination programs, it will be critical for CVM studies to use data collection and analysis methods that account for, rather than ignore, potential bias from culturally-based interactions between interviewers and their respondents. While studies such as these conducted in India, Bangladesh, and Mozambique provide information that could inform a fee-for-service approach to vaccination in developing nations, there is currently little political support for charging fees for vaccinations.
9. Suggestions for future research

This study’s close examination of the data collected by these three studies revealed trends in interviewer-respondent effects that were only in a few instances statistically significant, but nonetheless point to important implications for future CVM studies to consider.

The findings suggest that more detailed information about interviewer characteristics should be collected in order to control for their effects on respondents. In these three studies, interviewer age was not recorded, nor was interviewer education level or race. The interviewers were, in nearly all cases, of the same race or ethnicity as the respondents, and therefore, omitting race from the records might not have been a problem. Nonetheless, in multicultural nations and communities, there are subtleties that are glossed over by using the term race. For example, skin tone can represent social class in some countries, and could be used as a proxy measure for social class to determine if interviewer-respondent differences or similarities influenced responses. In all three sites the teams of interviewers were more homogeneous in age and education than the populations of respondents, but exact data on this difference was not recorded. A better research design to test for these effects would include measuring race, age, and education levels of interviewers and, if possible, allocating a variety of interviewers to each site based on these measurements.

CVM training protocols have been refined over time but tend to focus on getting the CVM scenario correct and eliminating diversity of interviewer styles and personalities. It would be useful to see if similar groups of interviewers trained under
protocols with different emphases would collect statistically different survey results from similar randomly selected respondent pools.

An analysis of interviewer effects in survey studies usually assumes interviewer effects are undesirable. In contrast, the CVM study design is intended to inform respondents about how cholera works as a disease and how to prevent it, how vaccines work against cholera, and the effectiveness of this particular cholera vaccine. Rather than arguing that interviewers in CVM surveys do not have an impact on respondent answers, one might argue the contrary – that interviewers intentionally improve respondents’ pool of decision information regarding the choice under consideration (in this case a vaccine). Given the potential influence of interviewers on respondent answers, might interviewers and the in-person household interview process be used more directly for health promotion, disease prevention, or for other general civic, lifestyle, or educational goals?
References


