

Foreigner Effects in Field Research: An Experimental Study

Kate Baldwin, Constantine Manda and Liz McGuire

Abstract

Does the presence of foreigners change individuals' responses in and their engagement with research projects with which the foreigner is affiliated? The effects of foreigners on individuals' responses to survey and interview questions is important to understand both for the design of field research, but also for understanding the effects of and best practices in implementing development assessments. This project adopts an experimental approach to examine the effect of foreigner presence as part of a research team on how citizens relate to the research project and how they respond to questions about themselves and their communities. Specifically, it randomly varies whether a foreigner is included in the research team that travels to study villages and conducts components of a multi-method field research project, and it randomly varies whether household survey respondents are exposed to the foreigner. By comparing survey responses and behavioral measures depending on exposure to foreigners as part of the research team, we will be able to measure the effect of foreign researchers.

Introduction

Do foreigners change individuals' portrayal of themselves and their engagement with research projects? The effects of foreigners on how individuals represent themselves and engage with the projects with which foreigners are associated have multiple implications. First, these effects have implications for the design of field research (Kapiszewski, MacLean and Read 2015; Paluck 2010; Udry 2003). If individuals present themselves in a different fashion in the presence of foreigners, it is important that research teams take this into account when designing projects. Second, these effects have implications for the design and implementation of international development projects. Development organizations often make decisions about the role and visibility of expatriate staff in their projects, including during baseline assessments. Yet, although there are reasons to think that these decisions influence how local communities present themselves and interact with the projects to which expatriate staff are assigned. This project will provide experimental evidence that sheds light on these questions.

This project seeks to contribute to the literature on best practices in data gathering and assessment in international locations (Kapiszewski, MacLean and Read 2015; Paluck 2010; Udry 2003). At the moment, much of the advice imparted in this literature is impressionistic rather than evidence-based. Apart from a subset of this literature that looks at the effects of interviewer identity, gender and proximity on survey responses (Adida et al. 2016; Blaydes and Gillum 2013; Campbell 1981), there has been little rigorous research on many aspects of field research design. Most relevant for our project, Weinreb (2006) finds that locally known interviewers achieve higher response rates and collect more reliable data than unknown (stranger) interviewers.

Our research focuses specifically on the effects of foreigner presence on the research process. Foreign principal investigators are often conflicted about how much time to spend in the field, due to two competing concerns. On the one hand, they would like to provide high levels of oversight of their projects, but on the other hand, they do not want to distort responses through their presence. Recent research by Cilliers, Dube and Siddiqi (2015) finds that the presence of a white foreigner does change how respondents play laboratory games. We are interested more broadly in how foreigner presence as part of a field team changes the responses provided in both qualitative and quantitative research.

The proposed experiment extends existing research in this area in several ways. First, it builds explicitly on political scientists' common best practices in quantitative and mixed methods field research in international locations. Second, it considers the effects of foreigner presence as part of a research team on both the *levels* and *accuracy* of reporting in interviews, focus groups and surveys.

Hypotheses

This project adopts an experimental approach to identify the effects of foreigners. Our interest in this question is mostly applied. However, we draw on existing political economy and social psychology literature to develop hypotheses about the types of responses and respondents that will be most affected by the presence of a foreigner as part of the research team and the direction

of the effects. By engaging in hypothesis testing about the types of variables likely to be affected by foreigner presence, we hope to provide a theoretical structure that can help other researchers decide whether the content and location of their studies are more or less vulnerable to misreporting based on research team composition.

We will test the following hypotheses regarding the direction of effects foreigners will have on measurement levels and measurement accuracy.

Measurement Levels:

H1. Exposure to a foreign vs. a local researcher will cause respondents to provide responses more aligned with foreign interests and values on *variables in which foreigners' positions are clear and salient*. (Social desirability hypothesis)

H2. Exposure to a foreign vs. a local researcher will cause respondents to provide more positive responses on *variables capturing individual or community status*. (Status hypothesis)

H3. Exposure to a foreign vs. a local researcher will cause respondents to provide more negative responses on *variables associated with neediness*. (Neediness hypothesis)

Measurement Accuracy:

H4. Exposure to a foreign vs. a local researcher will cause less accurate reporting on *variables capturing foreign alignment and neediness*. (Pandering hypothesis)

H5. Exposure to a foreign vs. a local researcher will cause *more accurate reporting on variables that are politically sensitive*. (Confidentiality hypothesis)

H6. Exposure to a foreign vs. a local researcher will cause more accurate reporting on *variables that require effort* in the presence of foreigners. (Stakes hypothesis)

H7. Exposure to a foreign researcher will cause respondents to provide less accurate responses on *variables that can be verified with insider knowledge*. (Detectability hypothesis).¹

Heterogenous Effects:

H8: The effects of foreigner presence will be larger in communities with more exposure to foreigners.

¹ Our design can only test this hypothesis under the assumption that there are increasing returns to more local researchers, as we never have a data gathering exercise with a research team with no local researchers.

Other Design Effects:

Although our primary purpose in this project is to measure the effect of foreign researchers, our design also permits us to analyze the effects of conducting field research in advance of surveys. As a result, we will also test the following hypotheses:

H9: The effects of field research in advance of surveys will cause greater social desirability effects and neediness effects insofar as it gives communities a chance to rehearse their responses.

H10: The effects of field research in advance of surveys will cause less accurate responses on variables subject to pandering but more accurate responses on variables subject to confidentiality concerns, detectability and recall effort insofar as it makes the research purpose more obvious and makes the researchers appear more informed/engaged.

Table 1. Predicted Effects of Foreigners on Variables and Contexts

<i>Variables</i>	<i>Context</i>	
	Low foreign exposure	High foreign exposure
Alignment with foreigner interests/values	Higher and less accurate outcomes	Higher magnitude effects
Status	Higher outcomes	
Neediness	Lower and less accurate outcomes	
Verifiable	Less accurate outcomes	
Politically Sensitive	More accurate outcomes	
Effort	More accurate outcomes	

Research Design

This project employs an experimental methodology in order to assess the effects of including foreigners in research teams. The intervention will randomly expose villages to one of three experimental arms, as illustrated in table 2. In the **control group**, there will be no pre-survey research conducted in the village (aside from any short logistical visits necessary to secure permissions). In the **qualitative research group**, there will be one day of focus groups and elite interviews conducted by an exclusively Tanzanian field team approximately one week before the household survey. In the **foreigner plus qualitative research group**, there will be one day of focus groups and elite interviews conducted by a field team that includes one foreigner approximately one week before the household survey. Communities will be randomized into the treatments, stratifying by region and previous exposure to foreigners.

Table 2. Experimental Conditions

<p>(1) Control: Survey only</p> <p>40 villages</p> <p>40 village chairperson interviews²</p> <p>800 household surveys</p>	<p>(2) Qualitative Research: Qualitative research by Tanzanian field team + survey</p> <p>40 villages</p> <p>40 focus groups (FGs) 40 village chairperson interviews</p> <p>800 household surveys (400 w/ FG invitees, 400 w/ others)</p>	<p>(3) Foreigner + Qualitative research: Qualitative research by field team with one foreigner + survey</p> <p>40 villages</p> <p>40 focus groups (FGs) with foreigner 40 village chairperson interviews with foreigner</p> <p>800 household surveys (400 w/ FG invitees/Foreign exposure, 400 w/ others)</p>
---	--	--

One week prior to the survey, mobilizing teams containing no foreigners will mobilize 20 people to participate in the household survey. In arms (2) & (3) of the study, 10 people will be selected at random and invited to participate in a focus group one or two days later, and 10 people will be selected at random to participate in a HH survey one week later. In arm (1) of the study, 20 people will be selected at random to participate in a HH survey one week later.

One or two days later, the focus groups will be conducted with the invited participants in arm (2) and (3) of the study. The focus groups will be on development issues, community needs, and local governance. In the foreigner plus qualitative research group, the foreigner will be present at the focus group. The elite interviews will be with the village chairman. In the foreigner plus qualitative research group, the foreigner will be present at this interview.

About one week later, all-Tanzanian research teams will return to the same villages to conduct a household survey (The foreigner will never return, even in condition 3).

The household surveys will ask about household and community needs, individual identity and status, individual values and cultural norms, attitudes toward the research project, attitudes toward foreigners, and willingness and reported ability to engage in local development projects.

The study villages have been purposefully selected for their plausibly exogenous variation in exposure to foreigners. In particular, we have tracked the presence of foreign missionaries across Tanzania from 1900, and our sample is made up of villages with no missionary presence, sporadic missionary presence and consistent missionary presence as a result of an exogenous shock that caused the removal of missionaries of certain denominations and nationalities from Tanzania post-WWI. Randomization will be blocked by region and previous exposure to foreigners.

² In this arm, the village chairperson interviews will be conducted at the same time as the household surveys.

All of the research activities will be implemented by IPA-Tanzania. One of the authors – Liz McGuire – will serve as the foreigner in the focus groups and elite interviews. The focus groups will be conducted in teams of three, and the elite interviews will be conducted in teams of two. All other team members will be Tanzanian. All team members will rotate across different roles and will work with different team members on different days; this will allow us to isolate the effect of Liz from the effect of other individuals.

We are also randomizing whether Liz presents in either a more locally embedded or a more foreign manner. In 21 communities, she will present in a more business-like fashion, wearing pants/business casual wear and speaking English. In 19 communities, she will present in a more locally embedded manner, wearing a skirt and speaking Kiswahili. Based on a survey of academics who conduct mixed methods field work in April 2018, both modes of presentation are common. This additional randomization will allow a tentative assessment of the extent to which mode of presentation matters.

Measures of Interest

We are interested in how the presence of a foreigner on the research team affects the way respondents respond to a variety of questions commonly asked on public opinion and household surveys. In particular, we will test the effect of foreign researcher exposure on responses to standard questions related to wealth, food security, income, literacy, gender equality, quality of local services, partisan alignment, and voter turnout. We will test the effect of foreign research presence on a standard behavioral measure of generosity (through a donation activity). We will also test the effect of foreign researcher exposure on three novel questions that help us to isolate the effects of different variable characteristics on reporting levels and accuracy – a question about extreme weather events, a question that asks for directions and a behavioral measure of engagement with the survey. All of these measures are included on the HH survey, and some of them are included in the village chairperson survey too (at which the foreigner will be directly present).

We predict how much foreign researcher exposure should affect the level and accuracy of responses to various types of questions based on the whether the variables signal alignment with foreigner interest/values, status or neediness; whether an individuals' responses are plausibly verifiable by a knowledgeable insider; whether the responses are politically sensitive; and whether the responses recall effort to recall. Our predictions about the effects of exposure foreign researchers and exposure to research advance teams on these different outcomes are indicated in table 3.

Table 3. Variables, Measurement and Expected Effects of Foreigners and Advance Teams

Variable	Variable Characteristics						Measurement	Predictions	
	Alignment with foreign interests/values	Status	Neediness	Verifiable with local knowledge	Politically sensitive	High effort required		Foreigner Effects	Advance Team Effects
Wealth			Yes	Yes			VC & HH (& verified)	Depressed levels; less accurate	Depressed levels; less accurate
Food Security			Yes				HH	Depressed levels; less accurate	Depressed levels; less accurate
Income			Yes			Yes	VC & HH	Depressed levels; unclear effects on accuracy	Depressed levels; unclear effects on accuracy
Literacy		Yes		Yes			HH (& verified)	Augmented levels; less accurate	No predicted effects on levels; more accurate
Foreign Involvement	Yes		Yes	Verifiable with foreign knowledge			VC ³	Unclear prediction on reported presence of other foreigners; augmented	

³ We ask similar questions on the household survey but these responses could also be affected by changes in the set of foreigners who have previously visited the village.

								benefit levels	
Gender Equality	Yes						VC & HH (& verified)	Augmented levels; less accurate	Augmented levels; less accurate
Quality of Services			Yes	Yes			VC & HH (& verified)	Depressed levels; less accurate	Depressed levels; less accurate
Pol. Alignment					Yes		VC & HH (& verified)	Depressed support for incumbent; more accurate	Depressed support for incumbent; more accurate
Turnout					Yes	Yes	HH (& verified?)	Depressed levels; more accurate	Depressed levels; more accurate
Weather Events			Yes	Yes		Yes	VC & HH (& verified)	Augmented levels; unclear effects on accuracy	Augmented levels; more accurate
Directions						Yes	VC & HH (& verified)	Augmented length; more accurate	Augmented length; more accurate
Engagement with Survey						Yes	VC & HH	Augmented levels	Augmented levels
Generosity		Yes	Yes				HH	Unclear prediction on levels	Unclear prediction on levels

Importantly, we are interested both in measuring the effects of foreigners on how respondents' answer questions in these areas, but also – where possible – in assessing whether the presence of foreigners leads to more or less accurate reporting. Below, we discuss our measurement of the variables above.

A. Wealth:

Measurement: We will measure wealth using a shortened version of the battery of questions asked on the DHS surveys, as developed by equitytool.org (Chakraborty et al. 2016). This battery of questions includes an asset list (electricity, television, radio, iron, bank account), floor material, wall material, roof material, type of fuel used for cooking and type of energy used for lighting.

Verification: Our enumerators will independently observe and record the floor, roof and wall materials.

B. Food Security

Measurement: We will measure food security using a battery of questions, validated as part of FANTA-2 (Food and Nutrition Technical Assistance Program 2) (Deitchler et al. 2010). These questions ask about the occurrence and, if yes, the frequency of three types of hunger in the past month (Was there ever no food at all in your household because there were not resources to get more? Did you or any household member go to sleep at night hungry because there was not enough food? Did you or any household member go a whole day or night without eating anything because there was not enough food?)

C. Income

Measurement: We will measure income through one question: “Approximately how much income does your household make per year?”

D. Literacy

Measurement: We will measure self-reported literacy by asking respondents “Can you read in Kiswahili?” and then, if yes, “Which of the following statements best captures how well you can read in Kiswahili? Cannot read any words, can read syllabi but not words, can read words fluently but not sentences, can read sentences fluently but not paragraphs, or can read paragraphs fluently.”

Verification: Later in the survey, we will verify self-reported literacy by asking respondents to read a passage from the 2012 Uwezo test.⁴ The enumerators will be trained, as per the Uwezo test guidelines, to indicate whether the respondent could not read any words, could read syllabi

⁴ The exact paragraph used is from page 4 of the 2012 Tanzania Test Booklet's Kiswahili Set 1. The full test booklet can be downloaded at http://www.uwezo.net/wp-content/uploads/2012/08/TZ_2012_Tests.pdf.

but not words, could read words fluently but not sentences, could read sentences fluently but not paragraphs or could read paragraphs fluently. This test has previously been administered by Twaweza, a Tanzanian NGO, to children 6 to 16 in Tanzania, Kenya and Uganda. The level of the test is based on a 2nd Grade level of competency.⁵

E. Foreign Involvement

Measurement: We will measure reported foreign involvement in the community using a question about how frequently white foreigners have been in the village in the past five years. We will measure the perceived benefits of their involvement through a question asking “Do people in the community talk about their impact on the community being very negative, somewhat negative, mixed, somewhat positive or very positive?”

F. Gender Equality

Measurement: We will measure gender equality in interhousehold decision-making, using a series of three questions frequently used in gender empowerment indices (Donald et al. 2017): In this household, who usually makes decisions about making major household purchase, like buying a goat or a cookstove? Who usually makes decisions about making purchases for daily household needs, like soap, sugar and kerosene? Who usually makes decisions about visits to your family or relatives?

We will also measure gender equality in occupational and educational aspirations for the youngest generation of one’s family (Beaman et al. 2012). Respondents will be cued to think either about the youngest girl or the youngest boy in their family, and then asked “What year of education do you expect him/her to achieve?” and “What occupation would you hope for him/her to do when he/she grows up?”

Verification: We will partly verify answers to the first set of questions by having the enumerator record whether there were any questions where the respondent said they would have to ask their spouse to get a good answer.

F. Quality of Services

Measurement: We will measure the quality of services through questions that ask the respondent to “rate the quality of the nearest primary school?” and to “rate the quality of the nearest health facility?”

Verification: We will verify the quality of services by having the senior field officers on the team fill out a survey about the infrastructure and number of workers at the nearest primary school and nearest health clinic. The questions used to measure the quality of schools and clinic are based on Twaweza’s Wananchi Survey (Twaweza 2012a; Twaweza 2012b).

⁵ In the 2012 Uwezo survey, 57 percent of children aged 10 to 16 could read the Kiswahili passage fluently.

G. Political Alignment

Measurement: We will measure political alignment by asking respondents, “Over the course of your life, have you ever belonged to or felt close to a political party? If so, which ones?” This is slightly different from standard questions about political parties, developed after discussion with numerous academics from and working in Tanzania about the sensitivities surrounding *current* partisanship in Tanzania.

Verification: We will partly verify this by having the enumerator independently make note of any partisan paraphernalia observed during the interview.

H. Turnout

Measurement: We will measure self-reported turnout by asking respondents “Did you vote in the last national election?”

Verification: We will partly verify this by comparing village turnout statistics to turnout statistics at the ward level (provided we can obtain this statistic from administrative sources).

I. Weather Events

Measurement: We will ask the respondent whether “In the last five years, have you experienced any natural disasters, such as drought or too much rain, in this ward?”

Verification: We will verify these self-reported data against geospatial data on droughts and rainfall at the ward level.

J. Directions

Measurement: We will ask for the respondent to provide us with directions from the village chairman’s office to the nearest bus stop.” We will measure the amount of detail they provide qualitatively and the amount of time they take to respond quantitatively.

Verification: We will verify the accuracy of the directions by comparing them with directions worked out by the senior field officers and their mobilizing teams.

K. Generosity

Measurement: We will measure generosity through a donation exercise (Frey and Meier, 2004). Respondents will be given 2000 TSh for completing the survey. Then we will tell them “Separately, we are collecting money for a Tanzanian orphanage and we would like to give you the opportunity to donate some portion of your payment to this fund. The decision to donate is up to you, and you do not need to make any donation.” The amount they are willing to donate is our measure of generosity. We will not allow them to donate more than 2000 TSh.

L. Engagement with Survey

Measurement: After the survey is complete, we will measure engagement with/support for the survey by seeing if the respondent is willing to answer a few more questions, as follows: “You have completed our main survey now. Thank you for your time. However, if you are willing, we have just a few more questions we’d like to ask. Would you be willing to respond to a few more questions?”

Analysis

We will analyze the outcome data at the individual level using ordinary least squares regression with strata/block fixed effects and standard errors clustered at the village level. The main coefficient of interest is the effect of interactions between household-level treatment assignments and village-level treatment assignments (i.e. the effect of foreign exposure is measured through the interaction between the effect of being invited to a focus group in a community where a foreign researcher was present at the focus group).

Level Effects:

We will measure level effects as follows:

$$y_{ij} = \beta_1 FG_{invitee_i} + \beta_2 ForeignerFG_j + \beta_3 FG_{invitee_i} * ForeignerFG_j + \mathbf{B}_k + \epsilon_{ij}$$

where β_3 measures the effect of exposure to foreign researchers absent within-community spillovers.

Accuracy Effects:

We will measure accuracy effects using the same formula but setting $y_{ij} = y_{ij}^s - y_{ij}^v$ where y_{ij}^s is the self-reported measure of the variable and y_{ij}^v is the validated/actual measure. (For a similar approach to measurement error, see Bound and Krueger 1991; Bound et al. 2001; Blattman et al. 2016).

In addition, we will measure accuracy effects as the correlation $r(y_{ij}^s, y_{ij}^v)$ among individuals exposed to different arms of the experiment, testing for significant differences in $r(y_{ij}^s, y_{ij}^v)$ by converting correlations to z-scores using the Fisher z-transformation.

Heterogenous Effects:

We will measure heterogenous effects by historic exposure to foreigners by interacting all variables with a measure of historic exposure to foreigners as follows:

$$y_{ij} = \beta_1 FG_{invitee_i} + \beta_2 ForeignerFG_j + \beta_3 FG_{invitee_i} * ForeignerFG_j + \beta_4 Highexposure_j + \beta_5 Highexposure_j * ForeignerFG_j + \beta_6 Highexposure_j * FG_{invitee_i} + \beta_7 Highexposure_j * FG_{invitee_i} * ForeignerFG_j + \mathbf{B}_k + \epsilon_{ij}$$

where β_7 measures the differential effect of exposure to foreign researchers in high-foreign exposure communities. We will also test whether the effect of foreigner varies by individual-level exposure to foreigners, replacing $Highexposure_j$ with $Highexposure_i$.

Advance Research Team Effects:

We will measure the effect of the advance research team on responses as follows:

$$y_{ij} = \beta_1 FG_{invitee_i} + \beta_2 FG_{community_j} + \mathbf{B}_k + \epsilon_{ij}$$

where β_1 measures the effect of exposure to an advance research team absent within-community spillovers. We will measure the effect of advanced team exposure on both the levels of responses and the accuracy of responses, as described above.

Bibliography

- Adida, Claire, Karen Ferree, Daniel Posner and Amanda Robinson. 2016. "Who's Asking? Interviewer Coethnicity Effects in African Survey Data." *Comparative Political Studies* 49 (12): 1630-1660.
- Beaman, Lori, Esther Duflo, Rohini Pande and Petia Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science* 335 (6068): 582-586.
- Blattman, Christopher, Julian Jamison, Tricia Koroknay-Palicz, Katherine Rodrigues and Margaret Sheridan. 2016. "Measuring the Measurement Error: A Method to Qualitatively Validate Survey Data." *Journal of Development Economics* 120: 99-112.
- Blaydes, Lisa and Rachel Gillum. 2013. "Religiosity-of-interviewer effects: Assessing the impact of veiled enumerators on survey response in Egypt." *Politics & Religion* 6 (3): 459-485.
- Bound, John and Alan Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1): 1-24.
- Bound, John, Charles Brown and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." *Handbook of Economics* 2: 3705-3843.
- Campbell, Bruce. 1981. "Race-of-interviewer effects among Southern adolescents." *Public Opinion Quarterly* 45 (2): 231-244.

- Chakraborty, Nirali, Kenzo Fry, Rasika Behl and Kim Longfield. 2016. "Simplified Asset Indices to Measure Wealth and Equity in Health Programs: A Reliability and Validity Analysis Using Survey Data from 16 Countries." *Global Health: Science and Practice* 4 (1): 141-154.
- Cilliers, Jacobus, Oeindrila Dube and Bilal Siddiqi. 2015. "The White-Man Effect: How Foreigner Presence Affects Behavior in Experiments." *Journal of Economic Behavior and Organization* 118c: 397-414.
- Deitchler, M, T. Ballard, A. Swindale, and J. Coates 2010. "Validation of a Measure of Household Hunger for Cross-Cultural Use." Academy for Educational Development [AED], Food and Nutrition Technical Assistance Project 2 [FANTA-2]. Washington, D.C.
- Donald, Aletheia, Gayatri Koolwal, Jeannie Annan, Kathryn Falb and Markus Goldstein. 2017. "Measuring Women's Agency." *Policy Research Working Paper 8148*.
- Frey, Bruno and Stephan Meier. 2004. "Pro-social Behavior in a Natural Setting." *Journal of Economic Behavior & Organization*. 54 (1): 65-88.
- Kapiszewski, Diana, Lauren MacLean and Benjamin Read. 2015. *Field Research in Political Science: Practices and Principles*. New York: Cambridge University Press.
- Paluck, Elizabeth Levy. 2010. "The Promising Integration of Field Experimentation and Qualitative Methods." *Annals of the American Academy of Political and Social Science* 628: 9-71.
- Twaweza. 2012a. Wananchi Survey Tanzania 2012 School Questionnaire 04-10-2012. Available at: https://twaweza.org/uploads/files/SzW_qnr_sch_baseline_2012.pdf
- Twaweza. 2012b. Wananchi Survey Tanzania 2012 Health Facility Questionnaire 04-10-2012. Available at: https://twaweza.org/uploads/files/SzW_qnr_hf_baseline_2012.pdf
- Udry, Christopher. 2003. "Fieldwork, Economic Theory, and Research on Institutions in Developing Countries." *American Economic Review* 93 (2): 107-111.
- Uwezo. 2012. Uwezo Test 2012. Available at: http://www.uwezo.net/wp-content/uploads/2012/08/TZ_2012_Tests.pdf
- Weinreb, Alexander. 2006. "The Limitations of Stranger-Interviews in Rural Kenya." *American Sociological Review* 71 (6): 1014-1039.