

ARTICLES

The role of interviewer ethnicity on survey responses: A case study in refugee camps in Bangladesh

Hannah Miles¹, Christian Els^{1 a}, Cholpon Ramizova¹

¹ Ground Truth Solutions

Keywords: Bangladesh, Rohingya, refugees, Cox's Bazar, humanitarian aid, ethnicity of interviewer effect

<https://doi.org/10.29115/SP-2022-0008>

Survey Practice

In this paper, we report results of a study conducted in one of the world's largest refugee camps, the Kutupalong–Balukhali expansion site in Cox's Bazar, Bangladesh, in early 2021. We compare quantitative results of a survey on Rohingya refugees' perceptions of humanitarian aid conducted by Bangladeshi and Rohingya interviewers. For 36 out of 40 perception questions, we found a significant difference between the interviewer groups. Rohingya respondents, when interviewed by a member of their own ethnicity, express much lower levels of satisfaction with aid services than those interviewed by Bangladeshi interviewers, are more likely to report rude or abusive behavior, and are less likely to admit to behaviors deemed socially undesirable. Existing literature on ethnicity-of-interviewer effects in public opinion research focuses mainly on the United States and ethnic minorities in Europe, and thus, this study provides new insight into interviewer effects in humanitarian settings. The findings from this study call into question the validity and comparability of surveys conducted in humanitarian settings around the world. For Bangladesh, our findings show the importance of recruiting interviewers from the same ethnic group as respondents in order to elicit honest opinions concerning issues that may be socially sensitive or seen as critical of the aid efforts undertaken on their behalf.

Background

In 2021, the number of people in need of humanitarian assistance worldwide reached almost a quarter of a billion, constituting more than 2.5% of the global population. Between August 2020 and July 2021, 80 million surveys were submitted to just one of the most popular data collection platforms in the humanitarian sector.¹ Obtaining accurate information from affected populations is crucial, as the results of these surveys directly impact how implementing agencies provide what kind of aid and to whom.

This research forms part of Ground Truth Solutions' continued engagement with crisis-affected communities in Bangladesh and took place in the Kutupalong–Balukhali expansion site in Cox's Bazar. Surveys among the Rohingya refugee population, including humanitarian needs assessments, are mainly conducted by Bangladeshi interviewers native to the area. Given

^a Corresponding author:

Christian Els
Ground Truth Solutions,
Wohllebengasse 16/5
1040 Vienna
Austria
christian@groundtruthsolutions.org

¹ Kobo Toolbox, hosting in the same period 100,000 user accounts on their humanitarian server.

considerable evidence that perception surveys of Rohingya refugees can exhibit significant bias (ACAPS / IOM 2021; Ground Truth Solutions 2020; REACH 2021), this study aims to examine whether the ethnicity of the interviewer biases responses, and thus whether choosing an interviewer of the same ethnicity as the respondent would produce results more closely corresponding to realities on the ground.

Literature review

Ethnicity-of-interviewer effects have been well-studied in the United States, mainly investigating effects between black and white interviewers and interviewees. Evidence from other parts of the world is much more limited, with only a few studies published in the last decade, for example on a selection of African countries (Adida et al. 2016), Timor Leste (Himelein 2015) and the Arabian Peninsula (Gengler, Le, and Wittrock 2019). To our knowledge, no systematic study has been undertaken in a refugee or humanitarian context.

Ethnicity-of-interviewer effects have been shown to be more pronounced for sensitive or taboo topics, such as sexual activity, fraud, and racism (Krumpal 2013). Specifically considering race, studies find that the extent of bias depends largely on the racial sensitivity of the question (Cotter, Cohen, and Coulter 1982; Heelsum 2013; Holbrook, Johnson, and Krysan 2019; Kappelhof 2017; Schaeffer 1980). United States-based studies indicate that both black and white respondents adjust their answers relative to the ethnicity of the interviewer to avoid offence (Campbell 1981; West and Blom 2017). Questions relating to ethnicity may see larger interviewer effects as “respondents relate questions about interethnic relations directly to the relationship with the interviewer” (Heelsum 2013). Thus, such questions cause the interviewee to stress similarities and to avoid criticisms. Little attention has been paid thus far to ethnicity-of-interview effects on questions not directly related to ethnicity (Holbrook, Green, and Krosnick 2003).

Social desirability is the tendency for people to offer responses that are based upon the perceived cultural expectations of others, or refers to conditions in which respondents feel pressure to answer survey questions in socially acceptable ways (Dillman and Christian 2005). Overall, the presence of an interviewer has long been shown to elicit more socially desirable answers and acquiescence (a tendency to agree with the interviewer) on the part of the interviewee (Dillman and Christian 2005).

The cross section between social desirability and race has also been studied, with respondents often anticipating the interviewer’s opinion based on their race and adjusting their responses accordingly. For example, Adida et al. (2016) report in a study focusing on 14 African countries that respondents interviewed by non-co-ethnics consistently gave more socially desirable answers that presented their ethnic group in a positive light. Similar bias has also been found in studies on immigrant populations in Europe (Heelsum 2013). Indeed, one explanation for the presence of social desirability bias is that

interviewer ethnicity provides a visual or audible indication of their opinion on certain questions (Kühne 2020; Welford et al. 1995). Thus, we would expect social desirability bias only for questions on which the interviewer's opinion can be assumed based on characteristics such as ethnicity, and not for neutral questions where such factors would not be informative. This is supported, for example, by a study from Timor Leste which found larger effects on subjective perception questions rather than objective, factual questions (Himelein 2015).

Of particular relevance for the refugee context are those studies that take into account the underprivileged minority position of a particular ethnic group. Ethnicity-of-interviewer effects may be particularly strong here, since discrimination that minority groups face increases pressure to socially conform to the interviewer (Holbrook, Green, and Krosnick 2003). Mistrust or uneven power relations may also affect the answers respondents are comfortable giving (Heelsum 2013), which is why qualitative field research in particular emphasises relationship building as a key research step (Funder 2005). Adida et al. (2016) found that politically salient non-co-ethnic dyads—that is, groups that have a history of conflict—show larger interviewer effects than dyads that are not caught up in political tensions. This is a particularly important consideration in this study, given Rohingya refugees often experience strained relations with Bangladeshi host communities.

Hypotheses and Research Objectives

Following West and Blom's (2017) application of the total survey error framework in the context of interviewer effects, we can distinguish three areas of bias when conducting surveys: coverage errors, stemming from how interviewers sample from the population; nonresponse bias, resulting from how successfully interviewers make contact with and elicit responses from interviewees; and measurement errors, which relate to the content of the responses provided. This research focuses on measurement errors, since coverage and nonresponse errors can largely be controlled in refugee camp settings where surveys are primarily conducted face-to-face and nonresponse rates are low.

Our hypotheses are as follows:

1. Rohingya refugees report more positive opinions about the humanitarian response to Bangladeshi interviewers compared to Rohingya interviewers.
2. The difference in perception scores given to Rohingya and Bangladeshi interviewers will be greatest on questions relating to ethnicity or race, such as questions on abuse or mistreatment by particular ethnic groups.

3. Rohingya respondents interviewed by an interviewer from the same ethnic group will be more likely to admit to behavior perceived to be socially undesirable, such as selling aid to buy other items.

Research design

This study compares survey results of Rohingya refugees reported to Bangladeshi and Rohingya interviewers. We demonstrate which questions have the largest differences between interviewer groups, identifying topics that may be particularly prone to bias.

The questionnaire was designed in collaboration with Rohingya interviewers and researchers from the International Organisation for Migration's Communication with Communities team (IOM CwC), drawing from themes commonly explored by perception monitoring work in the humanitarian sector. The survey consisted of 40 questions, most of which were binary in nature.

Both Bangladeshi and Rohingya enumerators received the same training on the survey and how to use the data collection software. Bangladeshi interviewers were part of a 100-member enumeration team that regularly conducts surveys and assessments in the camps. Rohingya interviewers were less experienced in quantitative enumeration than their Bangladeshi counterparts, and so received training on both qualitative and quantitative data collection methods through IOM CwC. It is unclear to what extent, if at all, the Rohingya interviewers' additional qualitative training impacted how the interviews were conducted and therefore the results of the research.

Sampling

We used purposive sampling to select 5 out of the 34 camps in the Kutupalong–Balukhali expansion site based on camp size, population density, and level of refugee community interaction with host communities. Given the high population density and lack of census data in the camps, a geo-information-systems-based sampling approach was used (Eckman and Himelein 2019). A minimum of 125 coordinates per camp and interviewer group were randomly generated, and interviews were conducted at these locations. With a minimum sample size of 125 per camp, the margin of error amounts to a maximum of 10 percentage points at a 95% confidence level per camp. The margin of error for overall estimates of all five camps combined amounts to a maximum of 4.5 percentage points at a 95% confidence level (Thompson 1987).

Table 1. Interviews per interviewer group and camp.

	Rohingya interviewers	Bangladeshi interviewers	Total
Camp 15	145	127	272
Camp 18	142	130	272
Camp 20	152	126	278
Camp 2e	137	132	269
Camp 9	141	124	265
Total	717	639	1,356

Sample

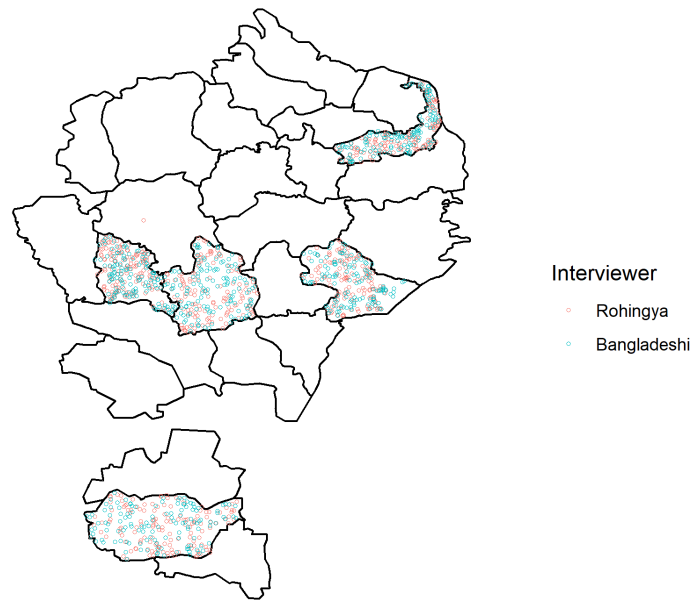


Figure 1. Map of Kutupalong–Balukhali expansion site, showing sample strategy in the five selected camps

Analytical methods

Multivariate non-parametric tests were employed to test the significance of the differences in responses between the two interviewer groups. We used the npvm package in R which employs non-parametric versions of well-known multivariate tests² (Burchett et al. 2017).

Subsequently, univariate significance tests were employed to discern which individual questions show a significant difference and in what magnitude. For ordinal Likert scale questions, we used a Wilcoxon rank sum test, and for binary data, Fisher's exact test. In both cases due to multiple comparisons, we used a Benjamini–Hochberg procedure to decrease false discovery rate.

² ANOVA type, Wilks' Lambda type, Lawley Hotelling type, and Bartlett Nanda Pillai type

Table 2. Results of non-parametric tests comparing two interviewer ethnicities.

	Test Statistic	Degrees of freedom group 1	Degrees of freedom group 2	P-value
ANOVA type test p-value	247.652	19.478	23,801.91	<0.00001
McKeon approx. for the Lawley Hotelling Test	99.883	28	1,195	<0.00001
Muller approx. for the Bartlett-Nanda-Pillai Test	99.802	28.022	1,194.98	<0.00001
Wilks Lambda	99.883	28	1,195	<0.00001

Results

The effect sizes presented below are expressed as a probabilistic index $P(x>y)$ (Acion et al. 2006); that is, the probability that a randomly selected individual interviewed by a Rohingya interviewer provides a higher rating than a randomly selected individual interviewed by a Bangladeshi interviewer. If there was no effect of interviewer ethnicity, we would expect these probabilities to be around 50%, meaning neither group is more likely to give a higher score than the other.

All four non-parametric multivariate tests show significant p-values, as presented in [Table 2](#). Thus, our results show a clear significant difference between responses elicited by different interviewer groups.

Univariate results

Satisfaction / Dissatisfaction with aid services

We see statistically significant and substantively large gaps between the perceptions collected by Bangladeshi and Rohingya interviewers for questions relating to satisfaction with aid services. People interviewed by Rohingya interviewers provide considerably more negative responses than people interviewed by Bangladeshi interviewers. Regarding education, food, health, and shelter services, the majority of people asked by Rohingya interviewers rated the quality of aid services either very bad or bad (between 63% and 87%), while the majority of those asked by Bangladeshi interviewers rated them good or very good.

Univariate Wilcoxon rank sum tests on each aid quality question show significant differences between the two interviewer groups ([Table 3](#)). Substantively, the differences between the groups are large—a randomly chosen person interviewed by a Rohingya interviewer has only between a 15% and 21% probability of giving a higher score (i.e., reporting higher levels of quality), compared to a 79–85% probability for a randomly chosen individual interviewed by a Bangladeshi interviewer.

Level of improvement of aid services

Similarly, on questions addressing degradation or improvements in aid services, people asked by Bangladeshi interviewers are more positive than people interviewed by Rohingya ([Figure 3](#)) and are more likely to select the neutral

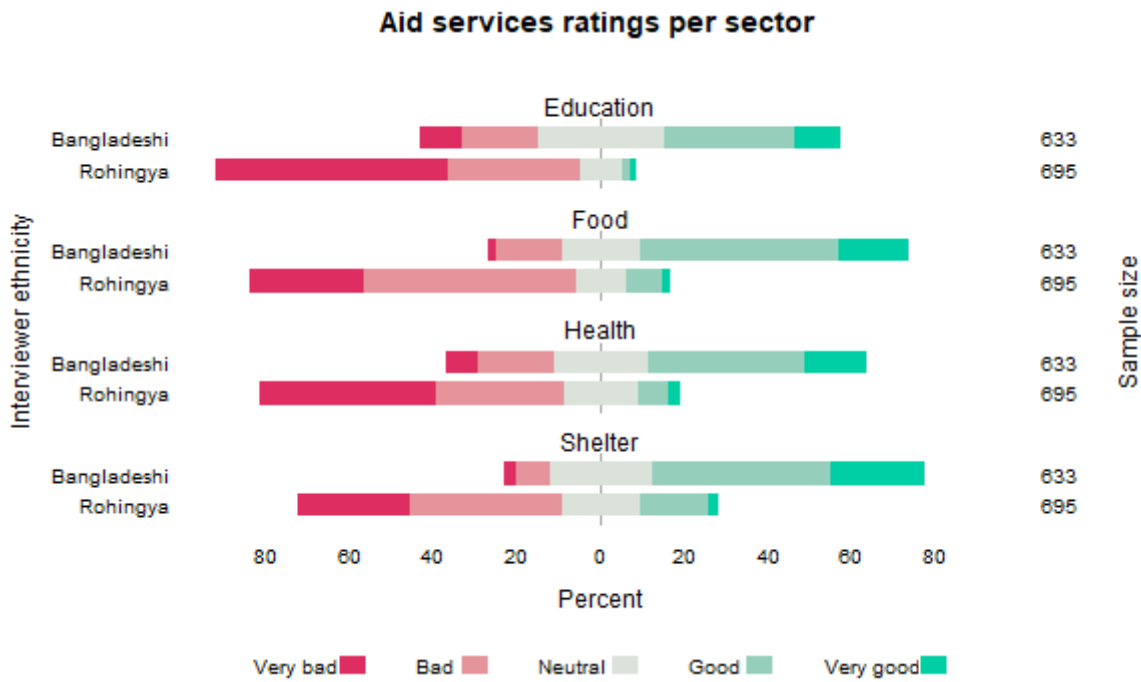


Figure 2. Quality of aid services.

Note that some respondents did not provide answers to these questions; hence, sample sizes are slightly smaller than the totals reported in [Table 1](#).

Graphs made using the HH package in R (Heiberger 2022).

Table 3. Service quality, results of Wilcoxon rank sum tests.

Service quality	Probabilistic Index	Adjusted p-value
Education	0.156	<0.00001
Health	0.206	<0.00001
Shelter	0.18	<0.00001
Food	0.15	<0.00001

option (except for water and sanitation [WASH] and nutrition services). These response patterns indicate a reluctance of the respondent to give a negative answer to Bangladeshi interviewers.

Univariate Wilcoxon rank sum tests on each improvement question show significant differences between the two interviewer groups for all but two questions ([Table 4](#)). Substantively, the question on improvements in protection services exhibits the largest difference between the groups—here, a randomly chosen individual interviewed by a Bangladeshi interviewer has a 76% probability of giving a higher score, compared to 24% for a randomly chosen individual interviewed by a Rohingya interviewer.

General perception questions

On general binary perception questions, the direction of the difference between interviewer groups depends on the sensitivity of the question, and whether certain responses would be more socially desirable than others vis-à-



Figure 3. Changes in aid services

Table 4. Service improvement, results of Wilcoxon rank sum tests.

Service improvement	Probabilistic Index	Adjusted p-value
Protection	0.243	<0.00001
Health	0.302	<0.00001
Education	0.304	<0.00001
Shelter	0.335	<0.00001
Non-food items	0.368	<0.00001
Psycho-social support	0.373	<0.00001
Food	0.38	<0.00001
Site management	0.383	<0.00001
Site development	0.408	<0.00001
Liquid petroleum gas	0.418	<0.00001
Nutrition	0.504	0.799
WASH	0.506	0.743

vis a Bangladeshi interviewer. Bangladeshi interviewers recorded higher levels of agreement ($P(x>y)^3=25-45\%$, $p<0.0001$) for questions which present humanitarian actors and the humanitarian response in a positive light.⁴ The opposite is true for questions which examine issues reflecting negatively on the humanitarian response.⁵ These questions show higher levels of agreement from individuals interviewed by Rohingya interviewers ($P(x>y)=51-77\%$, $p<0.0001$). Univariate Fisher's exact tests ([Table 5](#)) confirm that the difference is significant for 20 out of 24 questions.

The magnitude of the difference between the groups is particularly pronounced for questions relating to threatening or abusive behavior by humanitarians (see [Figure 4](#)). While 17% of those interviewed by a Rohingya interviewer (119 individuals) reported having witnessed someone in their community be threatened by a Bangladeshi humanitarian, no one reported this behavior to Bangladeshi interviewers. Similarly, 32% of people interviewed by a Rohingya interviewer witnessed rude behavior by Bangladeshi humanitarians, compared to only 2% of those interviewed by Bangladeshis. Contrary to our expectations, people interviewed by Rohingya were also more likely to report rude or threatening behavior by Rohingya humanitarians, though as seen in [Figure 4](#), reports of threatening behavior by Rohingya humanitarians are few. This could be understood as a partial contradiction to hypothesis 2, as those interviewed by Rohingya were more likely to report rude or threatening behavior by both Bangladeshi and Rohingya humanitarians. Alternatively, this may show increased comfort in reporting abuse to interviewers of the same ethnicity, regardless of the source of this behavior.

3 x = Rohingya interviewer, y = Bangladeshi interviewer

4 Including questions such as whether refugees would approach aid workers to report sensitive complaints, can ask aid workers about aid services or believe their opinions are considered in programming.

5 This includes whether refugees observed extortion by aid workers, filed a complaint against aid agencies, or sold excess food aid.

Table 5. Binary perception questions.

Perception question	Probabilistic Index	Adjusted p-value
Do you receive any food you don't eat?	0.772	<0.00001
Would you go to an aid organisation with a sensitive problem?	0.253	<0.00001
Have you seen a Bangladeshi humanitarian be rude to anyone in your community?	0.653	<0.00001
Did you sell any food items?	0.69	<0.00001
Can you ask an aid provider to provide the items you need?	0.347	<0.00001
Have you seen anyone ask for money in exchange for help?	0.615	<0.00001
Have you given feedback/a complaint?	0.657	<0.00001
Do humanitarians consider your opinion?	0.363	<0.00001
Have you seen a Bangladeshi humanitarian threaten anyone in your community?	0.583	<0.00001
Have you seen a member of the host community be rude to anyone in your community?	0.596	<0.00001
Have you seen a member of the host community threaten anyone in your community?	0.581	<0.00001
Have you seen a humanitarian ask for money in exchange for help?	0.581	<0.00001
Did you receive enough shelter materials?	0.38	<0.00001
Were women and girls able to use the latrines at night without being afraid?	0.377	<0.00001
Have you seen a Rohingya humanitarian be rude to anyone in your community?	0.571	<0.00001
Can everyone in your family sleep peacefully at night?	0.448	<0.00001
Have you seen a Rohingya humanitarian threaten anyone in your community?	0.53	<0.00001
Did you receive shelter items you didn't need?	0.519	<0.00001
Has anyone asked you what goods or services you need?	0.45	<0.00001
Was aid explained to you?	0.451	0.0001
Did you sell any shelter materials?	0.518	0.065
Can you borrow food from anyone in your sub-block?	0.487	0.091
Has anyone asked you if you have problems with aid?	0.486	0.146
Can you borrow money from anyone in your sub-block?	0.514	0.295

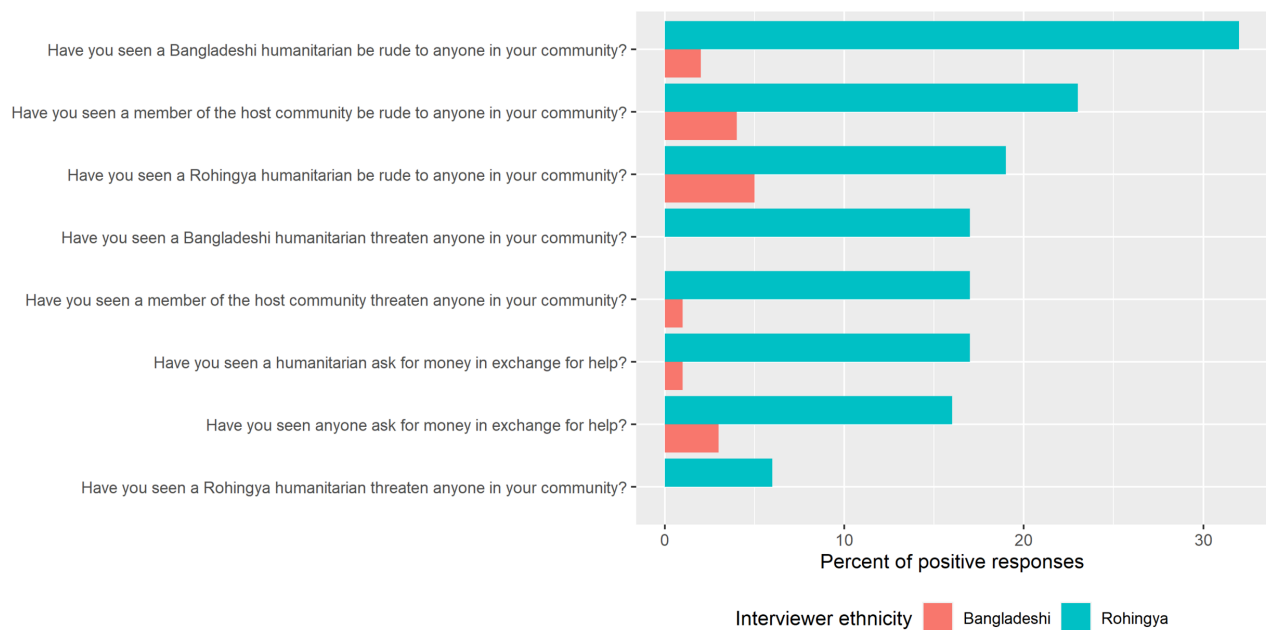


Figure 4. Questions relating to abusive behavior

Graphs made using the ggplot2 package in R (Wickham 2016).

In summary, a clear difference between answers provided to the two interviewer groups can be observed. Out of the 40 questions asked, only six do not show significant difference by interviewer ethnicity.⁶

Discussion and Conclusion

This study examines whether evidence of ‘ethnicity-of-interviewer’ effects can be found among survey responses by Rohingya refugees in Bangladesh —that is, whether responses to surveys collected by Bangladeshi interviewers are significantly different to those collected by Rohingya interviewers.

We find considerable evidence to support hypothesis 1, that *Rohingya refugees report more positive opinions to Bangladeshi interviewers compared to Rohingya interviewers*. When looking at perceived quality of humanitarian services ([Figure 2](#)), Rohingya refugees express far more positive views to Bangladeshi interviewers than to interviewers of their own ethnic group. Our findings are in line with previous research showing that questions which respondents relate directly to the interviewer may exhibit larger ethnicity-of-interviewer effects (see Adida et al. 2016; Heelsum 2013). The magnitude of the bias being greatest for questions relating to service quality is a clear indication that Rohingya refugees are less willing or comfortable criticizing services to Bangladeshi interviewers who can be thought of as representing the humanitarian response.

Hypothesis 2 that *the difference in perception scores given to Rohingya and Bangladeshi interviewers will be greatest on questions relating to ethnicity or race, such as questions on abuse or mistreatment by particular ethnic groups* receives some support as well: we see large differences between groups for questions on negative relations between Bangladeshi humanitarians and Rohingya refugees, and no significant difference on questions relating to relations within the Rohingya community.⁷ However, many questions which do not relate to race or ethnicity exhibited particularly large differences (see [Table 5](#)), and we see a significant difference on questions about rude or threatening behavior from Rohingya humanitarians as well. As such, hypothesis 2 cannot be confirmed.

Hypothesis 3, that *Rohingya respondents interviewed by an interviewer from the same ethnic group will be more likely to admit to behavior perceived to be socially undesirable, such as selling aid to buy other items*, can be confirmed. Some of the largest differences can be seen for questions relating to behavior that respondents may want to conceal, such as receiving or selling unwanted food items, and reporting issues with humanitarian aid (see [Table 5](#)).

⁶ Two of these questions refer to whether an individual can borrow food or money from members of their own community in times of need. Despite the large difference we see concerning selling food aid, there was no significant difference on whether people sell shelter materials. Finally, there was no significant difference on the question “Last month, has anyone come to ask whether you have any problems with receiving aid?”.

⁷ For example, whether Rohingya can borrow food or money from people within their community.

Despite the statistically significant differences for many of the questions, it should be noted that the size of the effect, and therefore, its substantive significance is small for some questions. Particularly on the questions “Do you receive shelter materials you don’t need?” and questions about rudeness or threatening behavior from Rohingya humanitarians, people interviewed by Rohingya interviewers have only between a 52% and 57% probability of giving a higher score than someone interviewed by a Bangladeshi interviewer. That is, we should be conservative about any conclusions derived regarding these questions.

The stark difference in responses to the majority of the survey questions demonstrates the extent to which response biases have impacted our understanding of the needs, preferences, and experiences of Rohingya refugees in Cox’s Bazar thus far. In addition to protection risks given the clear reluctance to speak out to Bangladeshi interviewers on abuse or mistreatment by humanitarian staff, humanitarians should also be concerned that Rohingya refugees appear less comfortable or willing to criticize the response to Bangladeshi interviewers. Ethnic strife in refugee contexts is not unique to Bangladesh but threatens to significantly affect the accuracy of data collected from refugee communities around the world. In order to collect accurate information, properly and systematically considering ethnicity as an influencing factor is key.

Acknowledgements

Contributors:

Christian Els conceptualized the study, devised the sampling strategy, conceptualised the statistical analysis plan, analysed the data, drafted and revised the paper.

Hannah Miles conceptualized the statistical analysis plan, analysed the data, and wrote and revised the paper.

Cholpon Ramizova conceptualized the study, designed data collection tools with Rohingya partners, and monitored data collection.

All authors worked for Ground Truth Solutions at the time of publication. This research was undertaken as part of Ground Truth Solutions’ mission to make humanitarian aid more people-centred and to include the views of people affected by crisis in humanitarian programming.

We thank our partners from the International Organization for Migration’s Communication with Communities team, including Danny Coyle, who were instrumental in designing the survey questions. We also thank their Needs and Population Monitoring Unit for providing logistical and enumeration

support. Their collective insight into the Rohingya community, including their linguistic and cultural norms, was crucial in creating a survey that would best capture their feedback.

We also thank our partners from Statistics without Borders, Tracey Di Lascio-Martinuk and Shalini Gopalkrishnan, who supported us in our analysis techniques. The time you took reviewing our approach and suggesting further avenues of analysis was greatly appreciated.

Submitted: March 02, 2022 EDT, Accepted: August 10, 2022 EDT

REFERENCES

- ACAPS / IOM. 2021. "Our Thoughts: Rohingya Share Their Experiences and Recommendations - Bangladesh." ACAPS / IOM. <https://reliefweb.int/report/bangladesh/our-thoughts-rohingya-share-their-experiences-and-recommendations>.
- Acion, Laura, John J. Peterson, Scott Temple, and Stephan Arndt. 2006. "Probabilistic Index: An Intuitive Non-Parametric Approach to Measuring the Size of Treatment Effects." *Statistics in Medicine* 25 (4): 591–602. <https://doi.org/10.1002/sim.2256>.
- Adida, Claire L., Karen E. Ferree, Daniel N. Posner, and Amanda Lea Robinson. 2016. "Who's Asking? Interviewer Coethnicity Effects in African Survey Data." *Comparative Political Studies* 49 (12): 1630–60. <https://doi.org/10.1177/0010414016633487>.
- Burchett, Woodrow W., Amanda R. Ellis, Solomon W. Harrar, and Arne C. Bathke. 2017. "Nonparametric Inference for Multivariate Data: The R Package npmv." *Journal of Statistical Software* 76 (1): 1–18. <https://doi.org/10.18637/jss.v076.i04>.
- Campbell, Bruce A. 1981. "Race-of-Interviewer Effects among Southern Adolescents." *Public Opinion Quarterly* 45 (2): 231–44. <https://doi.org/10.1086/268654>.
- Cotter, Patrick R., Jeffrey Cohen, and Philip B. Coulter. 1982. "Race-of-Interviewer Effects in Telephone Interviews." *Public Opinion Quarterly* 46 (2): 278–84. <https://doi.org/10.1086/268719>.
- Dillman, Don A., and Leah Melani Christian. 2005. "Survey Mode as a Source of Instability in Responses across Surveys." *Field Methods* 17 (1): 30–52. <https://doi.org/10.1177/1525822x04269550>.
- Eckman, Stephanie, and Kristen Himelein. 2019. "Methods of Geo-Spatial Sampling." In *Data Collection in Fragile States*, 103–28. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-25120-8_7.
- Funder, Mikkel. 2005. "Bias, Intimacy and Power in Qualitative Fieldwork Strategies." *The Journal of Transdisciplinary Environmental Studies* 4 (1).
- Gengler, Justin J., Kien T. Le, and Jill Wittrock. 2019. "Citizenship and Surveys: Group Conflict and Nationality-of-Interviewer Effects in Arab Public Opinion Data." *Political Behavior* 43 (3): 1067–89. <https://doi.org/10.1007/s11109-019-09583-4>.
- Ground Truth Solutions. 2020. "Bangladesh: Strengthening Accountability to Affected People." Ground Truth Solutions. <https://reliefweb.int/report/bangladesh/bangladesh-strengthening-accountability-affected-people-q1-2020>.
- Heelsum, Anja van. 2013. "The Influence of Interviewers' Ethnic Background in a Survey Among Surinamese in the Netherlands." In *Surveying Ethnic Minorities and Immigrant Populations: Methodological Challenges and Research Strategies*, edited by Joan Font and Mónica Méndez, 111–30. Amsterdam University Press. <https://doi.org/10.1017/9789048519187.006>.
- Heiberger, R.M. 2022. "HH: Statistical Analysis and Data Display: Heiberger and Holland." *R Package Version 3.1-49*. <https://CRAN.R-project.org/package=HH>.
- Himelein, Kristen. 2015. *Interviewer Effects in Subjective Survey Questions: Evidence from Timor-Leste*. Policy Research Working Papers. The World Bank. <https://doi.org/10.1596/1813-9450-7208>.
- Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. "Telephone Versus Face-To-Face Interviewing Of National Probability Samples With Long Questionnaires: Comparisons Of Respondent Satisficing And Social Desirability Response Bias." *Public Opinion Quarterly* 67 (1): 79–125. <https://doi.org/10.1086/346010>.

- Holbrook, Allyson L., Timothy P. Johnson, and Maria Krysan. 2019. "Race- and Ethnicity-of-Interviewer Effects*." In *Experimental Methods in Survey Research*, 197–224. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119083771.ch11>.
- Kappelhof, Joost. 2017. "Survey Research and the Quality of Survey Data Among Ethnic Minorities." In *Total Survey Error in Practice*, 235–52. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119041702.ch11>.
- Krumpal, Ivar. 2013. "Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review." *Quality & Quantity* 47 (4): 2025–47. <https://doi.org/10.1007/s11135-011-9640-9>.
- Kühne, Simon. 2020. "Interpersonal Perceptions and Interviewer Effects in Face-to-Face Surveys." *Sociological Methods & Research* 1 (36). <https://doi.org/10.1177/0049124120926215>.
- REACH. 2021. "Participation of Rohingya Enumerators in Data Collection Activities: Findings of a Pilot Assessment in Cox's Bazar, Bangladesh." https://www.impact-repository.org/document/reach/15f51f89/reach_bgd_brief_rohingya_enumerator_pilot_april2019_1.pdf.
- Schaeffer, Nora Cate. 1980. "Evaluating Race-of-Interviewer Effects In a National Survey." *Sociological Methods & Research* 8 (4): 400–419. <https://doi.org/10.1177/004912418000800403>.
- Thompson, Steven K. 1987. "Sample Size for Estimating Multinomial Proportions." *The American Statistician* 41 (1): 42–46. <https://doi.org/10.2307/2684318>.
- West, Brady T., and Annelies G. Blom. 2017. "Explaining Interviewer Effects: A Research Synthesis." *Journal of Survey Statistics and Methodology* 5 (2): 175–211. <https://doi.org/10.1093/jssam/smw024>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-24277-4>.
- Wolford, Monica L, Ronald E Brown, Anne Marsden, James S Jackson, Chase Harrison, and Monica L Wolford. 1995. "Bias in Telephone Surveys of African Americans: The Impact of Perceived Race of Interviewer on Responses." #5110. Institute for Social Research.