Replicable quality across every channel: the Busara research agenda on methods in the Global South

No. 005

Winnie Mughogho, Nicholas Owsley, Tom Wein

Busara

Replicable quality across every channel: the Busara research agenda on methods in the Global South

Winnie Mughogho[1], Nicholas Owsley[2], Tom Wein[3]

Author Affiliations:
[1] The Busara Center for Behavioral Economics. Corresponding author:
winnie.mughogho@busaracenter.org
[2] The Busara Center for Behavioral Economics; University of Chicago
[3] The Dignity Project

Conflicts of Interest: There are no Conflicts of Interest to declare for this study

The Busara Blue Paper Series is a collection of working papers on topics relevant to experimental and behavioral economics in the Global South. All papers comply with Busara's Open Science and research ethics policies, have IRB approval, and all study materials and data are to be published on the Open Science Framework www.osf.io. Papers are circulated for discussion and comment - feedback on the paper directly to the corresponding author is welcomed.

Replicable quality across every channel: the Busara research agenda on methods in the Global South

Abstract

Any act of measurement brings with it concerns about data quality and replication. Yet very little of this work has focused in detail on the specific measurement and data quality concerns related to conducting research in the Global South. Busara operates both remote and in-person data collection, employing everything from photovoice to laboratory 'games' to list experiments. Across both of these domains, we will build on our long tradition of careful testing of measures and techniques to ensure high levels of access, response, attention and comprehension. We will examine what methodological practices work best for various populations, especially those with the least social power, in the Global South, to maintain data quality (this differs from our closely related program of work on cross-cultural validation of behavioral science constructs). Busara is well-positioned to do this, and to disseminate protocols for the use of these methods.

JEL Codes: B41; C80; D91

Key Words: replication, generalization, methods, data quality, attention, behavioral science

Global debate

There is great power to the availability of data, and the power of the digital revolution has reduced the barriers to data collection for many researchers and institutions (World Bank, 2021). That has begun to correct some of the many imbalances in knowledge production, from who gets studied and has their voice heard to who gets the funding and other privileges needed to publish empirical research (Chelwa, 2017; Briggs & Weathers, 2016; Ndlovu-Gatsheni, 2021). Research employing all kinds of methods can be transformative.

Our agenda

Busara Center for Behavioral Economics works with researchers and organisations to advance and apply behavioral science in pursuit of poverty alleviation. In three Busara 'Blue Papers', we are laying out our three year (2022-2025) research agenda on methods, cross-cultural validation and research ethics in line with that mission. This is part of our commitment to rigorous, respectful research in the Global South. We would be glad to discuss collaborations on any of these core questions, and we welcome feedback on these agendas.

There is wide agreement on the power of research among researchers. However, there is less debate than might be expected (at least in the published literature) about how to go about it. Critical thinking about measurement is often neglected (Boumans, 2015). Even comparatively practical handbooks on the conduct of experiments focus mainly on choosing an overall method, with little exploration of the challenges of measurement (Gertler et al, 2016; Khandker et al, 2010).[1] Many studies proceed using methods and measures that were developed in the Global North for use on WEIRD populations (Henrich et al, 2010). There is an extremely active methodological debate in the positivist social sciences when it comes to statistical analysis and interpretation after studies are completed,[2] but much less new published work on the mechanics and reliability of the incentivised lab games (Amir et al, 2012) and survey outcomes (Broockman et al, 2017) that many studies employ.[3] Though much piloting is done, especially when experiments take place in new contexts, this is not an explicit topic of study and publication (Glöckner & Hilbig, 2011; Hilbig, 2010). This is despite the fact that small changes to laboratory or survey procedure can have a major impact on results (Cilliers et al, 2015).

Particularly absent in this debate is the study of quantitative and qualitative measurement among populations in the Global South. Virtually all of the methods used by the behavioral sciences were first developed for WEIRD populations in the Global North (Haushofer et al, 2014). Just as those populations differ systematically in their behaviors, preferences and

---

[1] One valuable exception is the 'Handbook of Field Experiments' (Duflo & Banerjee, 2017), which has notable sections on construal, preference elicitation, mechanisms and related topics.
[2] See for instance recent debates on difference-in-difference methods (Baker, 2019).
[3] This compares to a longer tradition of more careful work on psychometrics in psychology - c.f. Dunlap, 1937.

psychology (Henrich et al, 2010; Muthukrishna et al, 2020), they very likely differ also in what techniques help drive high quality data.

There is however a burgeoning gray literature on these topics, often among populations in the Global South, shared on the institutional blogs of various implementers of research (e.g. Mathur, 2020; Rozo, 2021; Canagarajah et al, 2021). Much is also learned through ad hoc piloting, shared among academics through informal networks and rarely published for all to access. This push to improve quality of research has now begun to wrestle with the need to adopt different approaches for various populations in the Global South. In Busara's view, this is a potentially major source of underexplored variance and error in current research. It is the types of questions addressed in this gray literature that we propose to extend and bring structure to under this research agenda.

Meanwhile, for practical reasons, research methods are evolving fast. There has been an accelerating trend during the pandemic to try out new methods, with researchers and institutions forced to innovate in order to continue their work (Richardson et al, 2021; Hensen et al, 2021). The pandemic has also made much of that work exceptionally urgent (Van Bavel et al, 2020; Abaluck et al, 2021). This move has brought to the fore old questions of recruitment, administration, response rates, data quality, representativeness, research ethics, and much more (Cardel et al, 2020; Tuttle, 2020; Holmes et al, 2020; Lobe et al, 2020), for both quantitative and qualitative research methods - questions with which the social sciences have always wrestled (Lewis-Beck et al, 2003).

With remote research already rising in prevalence before the pandemic (Buhrmester et al, 2018), much of the most recent research literature has focused on the use of Amazon's Mturk. Considerable worries expressed about the quality of data being collected through that platform (Chmielewski & Kucker, 2020; Kennedy et al, 2020; Matherly, 2019; Aguinis et al, 2020). This has led many researchers to seek alternative panels that combine similar ease of use with greater assurances regarding data quality (Palan & Schitter, 2018; Peer et al, 2017; Peer et al, 2021). There is additionally considerable careful work, dating back a little further (Down & Duke, 2003) on SMS and other phone-based methods (Alam et al, 2014; Gibson et al, 2017; Lau et al, 2018; Lau et al, 2019a; Greenleaf et al, 2017), with valuable practical advice on what works in designing them (Lau et al, 2019b; Van Der Heijden, 2017; Kopper & Sautman, 2020). We note that the availability of participants in these panels is often more limited for countries in the Global South, and they are generally designed to work within the technological infrastructure of Global North countries.

When it comes to in-person data collection, recent methodological literature has been as varied as the many methods that domain contains, but it has included debates about the interpretation of interviews and ensuring rigor in qualitative data collection (Hughes et al, 2020; Hammersley, 2020; Edwards & Holland, 2020; Jenner & Myers, 2019), interrogating the unstructured interactions around quantitative data collection (Ongena & Dijkstra, 2020; Schaeffer, 2020; Conrad & Schober, 2020; Ting & Fitzgerald, 2019), reflections on positionality and the identities of researchers (Islam, 2020; Rogers, 2020; Pritchard, 2019),

and how differences between researchers and participants may be overcome in order that all sections of the population are included (Brooks et al, 2019; Kianersi et al, 2019; Kruger et al, 2019; Mao & Feldman, 2019), as well as the role of psychometrics (Wijsen & Borsboom, 2021). Only a handful of the papers cited in the preceding two paragraphs are focused on populations in the Global South.

There are also many relevant recent studies of methodological concerns that apply to both remote and in-person collection and design, such as extensive discussion of response options (Terentev & Maloshonok, 2019; Montagni et al, 2019; Höhne & Krebs, 2021; Rasmussen et al, 2020; Wang & Krosnick, 2020), respondent attention (Silber et al, 2019; Nichols & Edlund, 2020) and translation (Sarac & Koc, 2020). Yet once again, few of these focus on populations with the least social power, or more generally on populations in the Global South.

These concerns over data quality sit alongside ongoing anxiety in the field about wider failings of replicability across the behavioral and experimental social sciences (Simons, 2014; Klein et al, 2015). There has been a wide ranging push to drive up standards of experimental research in general, through principles of open science, robustness, pre-registration and similar improvements (Munafò et al, 2017; Blair et al, 2019). Important work has been done on researcher degrees of freedom (Veldkamp et al, 2017).

So far, research into the methods employed in the study of the positivist social sciences has been limited. What work there is rarely focuses on the populations with the least social power in the Global South. Compared to the work on replication and how researchers should analyse their data, there has not been an equivalent effort to examine the detail of the measures we use. A community of research implementers has begun to correct for this through occasional blog posts on methodological procedure in the Global South, but there is far more to do. This research agenda will formalise and extend that effort.

The story of methodological research at Busara

When the pandemic hit, like all research organisations, Busara converted to remote research - to continue our projects, and to ensure that we could keep hearing from those most in need during this crisis. We immediately launched a program of methods research, testing out the quality of the responses we were receiving. In that work we examined how SMS and phone modes compared across incentives, reminders, completion rates, and response quality. We have since launched a series of studies within the Busara KITE app, our remote data collection tool, learning more about what combination of incentives, reminders, and survey design principles succeed best for that research mode. Technologies of this kind have great promise in allowing us to reach participants with lower costs, provide feedback and co-create research alongside - if done right. Our Lab team has been learning a great deal across all our studies about

what works well in boosting data quality - something that we want to record and study more intentionally.

This work has built on a long succession of work examining methods at Busara. We have recently examined the inclusivity of our research processes, with a particular focus on gender, in both Kenya and Nigeria. We have launched an extensive program of contextualisation of standard behavioral games, tests and survey items, ensuring that they are comprehensible and relevant to our research participants. In 2018, in partnership with Ajua, we conducted an extensive assessment of data quality in SMS surveys. Much of this work is in the process of being shared through our contribution to the Science of Behavior Change Open Instruments repository.

In addition, we at Busara have often worked with academic partners to design and implement novel methods to address the most challenging specific questions in the laboratory. These have included developing new measures of ethnic preferences (Blum et al, 2021), stress (Haushofer et al, 2015b), noise and cognitive function (Dean, 2021), mental health (Haushofer et al, 2020; Baranov et al, 2020), social preferences (Jang & Lynham, 2015), temperature (Almås et al, 2019), legacies of violence (Amanela et al, 2020), civic engagement (Wein et al, 2020) and much else.

Throughout this time, Busara has been experimenting with new methods of data collection, and validating those methods, from mobile laboratories set up in hostels and buses to phone sensing, interactive voice response, participatory methods and qualitative co-design. At all times, we have tried to ground our methods in the lived realities of our participants. However, we have not always organised this work into a focused progression of addressing an organised taxonomy of pressing methodological questions. This agenda aims to bring order and focus to the great deal of learning we have done so far.

Busara's contribution

Busara's role in this global debate will focus on piloting and experimentally testing methodological innovations with our 75,000-strong pool of respondents, recruited and maintained in India, Kenya, Uganda and beyond. This will begin through the use of KITE, Busara's tool for remote data collection, and will later transition to incorporating learnings across other remote and in-person data collection methods, including both qualitative, quantitative and lab-based methods.

We will begin by developing contextually appropriate measures of access, attention and comprehension. Access, attention and comprehension are, we believe, reasonable proxies for the overall construct of high-quality data. Once we are confident in these outcome

measures, we will pilot and test improvements to boost those scores, ensuring high levels of participant access, response, attention, and accuracy. All this will be done for both our qualitative and quantitative research. As well as covering different methods, we will also focus on heterogeneity of results and pay special attention to neglected and minoritised segments of the population, including women, those in informal settlements, and ethnic or other minorities, ensuring that their voices are heard and examining how their experiences may differ. Throughout this process, Busara will share standardised protocols for use across our own projects and by the wider research community, where the data supports the use of such standardised approaches across different populations.

In support of our qualitative work, as well as doing the 'contextualisation' of methods and adapting for data quality, we will additionally develop and test protocols for quality in analysis, especially in transcription and coding, and learn lessons from a variety of novel qualitative methods. In doing so, we will have a special focus on adapting qualitative methods - especially group methods - to the limitations and opportunities of remote research in low-connectivity contexts, given the special value those methods have historically gained from face-to-face contact. This is likely to include exploring better ways to approach novel qualitative methods, including co-design, story completion, diaries and photovoice.

Since KITE is a new tool, and since there is considerable interest in the use and reliability of digital remote data collection methods of this kind, we will publish a specific review of methodological work and participant data related to KITE.
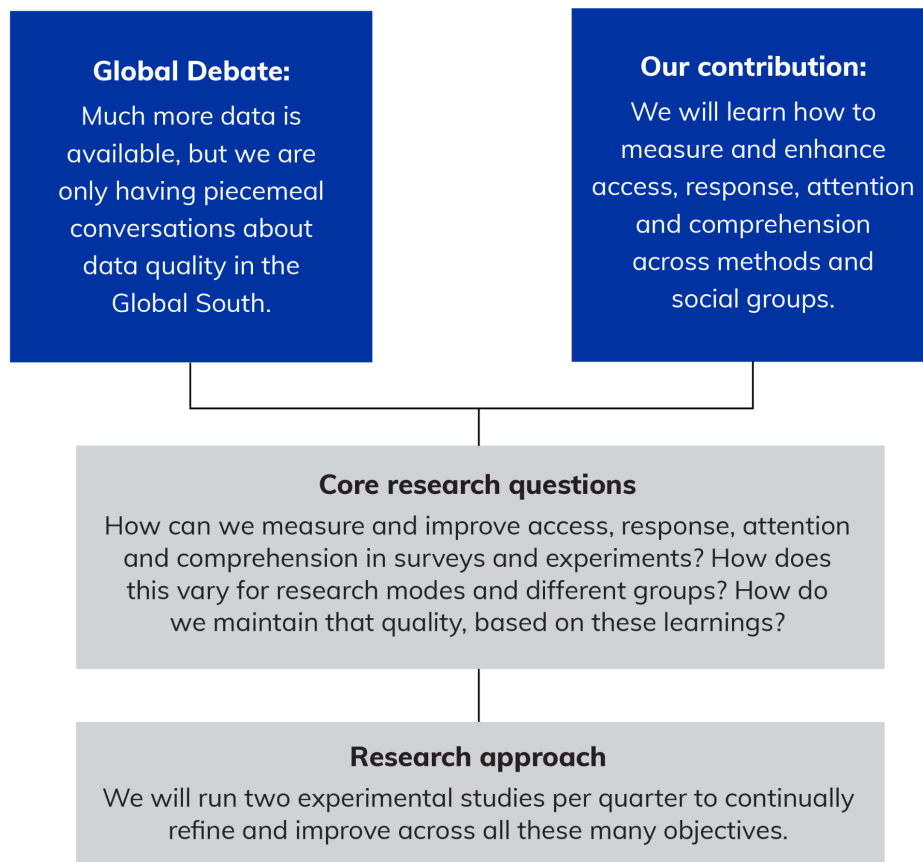
As a busy research institution that has run more than 500 behavioral science projects, and may be running upwards of 75 studies at any one time, and working across many different methods, Busara is ideally placed to create and share this sort of knowledge. Simple answers to questions of methods, translated into simple-to-administer protocols, should realise important gains in efficiency, reliability and quality across the institution and the research field. They should do so in ways that shift the focus of research towards appropriateness for populations in Global South contexts.

This work proceeds in close concert with a similar program of work on cross-cultural validation, which is intimately related. We can differentiate between them thusly: cross-cultural validation focuses on answering a set of research questions about specific *constructs* that are central to behavioral science. By contrast, our methods research program will build knowledge about all research methods, with a continual focus on data quality.

Core questions

Over the three years of this research agenda, from 2022 until 2025, we will comprehensively answer the following core questions, conducting multiple studies to address each one:

1. How can we contextualize, measure and improve access, response, attention, comprehension and depth of responses in surveys, experiments and qualitative research for populations in the Global South?
2. How can we quantify data quality in both qualitative and quantitative research?
3. How do the answers to these questions vary across gender, racial, national and economic groups? How can we build inclusion across different modes of research, for those populations with the least social power, and make certain these modes do not perpetuate exclusion?
4. How do the answers to these questions vary across common research methods, including in-person and remote data collection?
5. How do the answers differ for the most sensitive and difficult to approach research topics?
6. What protocols and practices should we adopt to maintain high data quality across methods and groups?

Global Debate:
Much more data is available, but we are only having piecemeal conversations about data quality in the Global South.

Our contribution:
We will learn how to measure and enhance access, response, attention and comprehension across methods and social groups.

Core research questions
How can we measure and improve access, response, attention and comprehension in surveys and experiments? How does this vary for research modes and different groups? How do we maintain that quality, based on these learnings?

Research approach
We will run two experimental studies per quarter to continually refine and improve across all these many objectives.

Research approach

This agenda will be best served by the prototyping and testing of different approaches to measure and improve data quality. We will therefore aim to run two experimental studies per quarter to continually refine and improve across these many objectives. Frequently these will be embedded in other existing studies, including those from our cross-cultural validation agenda. Initially the work will focus on refining data quality in KITE, our remote data collection app, before moving to other methods and data collection modes as the agenda develops.

Initially you can expect from Busara studies of incentives to install the KITE app and to take surveys. You can expect careful testing of attention checks, and of the role of incentives and reminders in boosting attention. You can expect new studies of list experiments and other approaches to asking the most sensitive questions, to see which perform best for populations in the Global South. To begin with, we will mainly be asking these questions of participants drawn from the Busara Lab Pool, which is principally composed of low income people in Nairobi. As the agenda develops, we will examine different segments of the population across the countries that Busara serves (including India, Nigeria, Ethiopia, Uganda and Tanzania), and trialling what we've learned from KITE in our laboratory, phone surveys, face-to-face surveys and qualitative methods. Busara will publish regular blog posts and working papers dedicated to sharing the learnings from this agenda.

Conclusion

Whose voices get heard? Does the use of certain modes perpetuate a gap in who we can reach? When we get a chance to hear them, are we hearing open, unfiltered views - or did people feel obliged to give certain answers when speaking to us? How can we create opportunities for people to be as open as they can be in sharing their lives with researchers? These are some of the most fundamental moral and technical questions that researchers face.

Through this agenda we will be doing the detailed, sometimes undervalued, work of making sure our methods are just right. We do so in service to the academic and policy community, which rightly asks how certain we can be in each of our findings. We do so too in service to the participants who give their time and views, and who deserve to know that we have done all we can to support them to share those as openly as we would wish to. Methods affects everything. Our hope is that by improving methods in our own work and by communicating how other research implementers may in turn improve their own methods, we will help generate better data for better policies in pursuit of human flourishing.

Call for collaborators

These are issues on which we welcome collaborators. If you have feedback on these papers, which are periodically updated, or if you are interested in discussing, supporting or participating in our research agenda on cross-cultural research, methods or research ethics, we'd love to hear from you. Please contact Anisha Singh on anisha.singh@busaracenter.org

References

Abaluck, J., Kwong, L. H., Styczynski, A., Haque, A., Kabir, M. A., Bates-Jefferys, E., Crawford, E., Benjamin-Chung, J., Benhachmi, S., & Raihan, S. (2021). Normalizing Community Mask-Wearing: A Cluster Randomized Trial in Bangladesh. National Bureau of Economic Research. DOI 10.3386/w28734

Aguinis, H., Villamor, I., & Ramani, R. S. (2020). MTurk Research: Review and Recommendations. SAGE Publications Sage CA: Los Angeles, CA.

Alam, I., Khusro, S., Rauf, A., & Zaman, Q. (2014). Conducting surveys and data collection: From traditional to mobile and SMS-based surveys. Pakistan Journal of Statistics and Operation Research, 169–187. https://doi.org/10.18187/pjsor.v10i2.758

Almås, I., Auffhammer, M., Bold, T., Bolliger, I., Dembo, A., Hsiang, S. M., Kitamura, S., Miguel, E., & Pickmans, R. (2020). Destructive Behavior, Judgment, and Economic Decision-making under Thermal Stress. National Bureau of Economic Research. https://doi.org/10.3386/w25785

Amanela, S., Ayee, T. F., Buell, S., Escande, A., Quinlan, T., Rigterink, A. S., Schomerus, M., Sharp, S., & Swanson, S. (n.d.). The mental landscape of post-conflict life in northern Uganda. Part 1: Research on behaviour and post-conflict life in northern Uganda – the research design. Working Paper, 88. https://securelivelihoods.org/wp-content/uploads/SLRC-Uganda-Report-01-ResearchDesign-WP88-Jul20.pdf

Amir, O., Rand, D. G., & Gal, Y. K. (2012). Economic Games on the Internet: The Effect of $1 Stakes. PLOS ONE, 7(2), e31461. https://doi.org/10.1371/journal.pone.0031461

Baker, A. C. (2019, September 25). Difference-in-Differences Methodology. Andrew Baker. https://andrewcbaker.netlify.app/2019/09/25/difference-in-differences-methodology/

Baranov, V., Haushofer, J., & Jang, C. (2020). Can Positive Psychology Improve Psychological Well-Being and Economic Decision-Making? Experimental Evidence from Kenya. Economic Development and Cultural Change, 68(4), 1345–1376. https://doi.org/10.1086/702860

Blair, G., Cooper, J., Coppock, A., & Humphreys, M. (2019). Declaring and diagnosing research designs. American Political Science Review, 113(3), 838–859. https://doi.org/10.1017/S0003055419000194

Blum, A., Hazlett, C., & Posner, D. N. (2020). Measuring Ethnic Bias: Can Misattribution-Based Tools from Social Psychology Reveal Group Biases that Economics Games Cannot? Political Analysis, 29(3), 385–404. https://doi.org/10.1017/pan.2020.37

Boumans, M. (2015). Science Outside the Laboratory: Measurement in Field Science and Economics. Oxford University Press.

Briggs, R. C., & Weathers, S. (2016). Gender and location in African politics scholarship: The other white man's burden? African Affairs, 115(460), 466–489. https://doi.org/10.1093/afraf/adw009

Brooks, R., Lainio, A., & Lažetić, P. (2020). Using creative methods to research across difference. An introduction to the special issue. International Journal of Social Research Methodology, 23(1), 1–6. https://doi.org/10.1080/13645579.2019.1672281

Broockman, D. E., Kalla, J. L., & Sekhon, J. S. (2017). The Design of Field Experiments With Survey Outcomes: A Framework for Selecting More Efficient, Robust, and Ethical Designs. Political Analysis, 25(4), 435–464. https://doi.org/10.1017/pan.2017.27

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. Perspectives on Psychological Science, 13(2), 149–154. DOI: 10.1177/1745691617706516

Canagarajah, R., Oyegunle, S., & Waweru, L. (2021, August 13). Unlikely tools for unlikely times. The Busara Blog. https://medium.com/busara-center-blog/unlikely-tools-for-unlikely-times-3d4d48b90387

Cardel, M. I., Manasse, S., Krukowski, R. A., Ross, K., Shakour, R., Miller, D. R., Lemas, D. J., & Hong, Y.-R. (2020). COVID-19 impacts mental health outcomes and ability/desire to participate in research among current research participants. Obesity, 28(12), 2272–2281. https://doi.org/10.1002/oby.23016

Chelwa, G. (2017). Does Economics Have an 'Africa Problem'? Some Data and Preliminary Thoughts. African Perspectives on Global Corruption Conference, Johannesburg, February, 22. https://wiser.wits.ac.za/system/files/seminar/Chelwa2017.pdf

Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. Social Psychological and Personality Science, 11(4), 464–473. https://doi.org/10.1177/1948550619875149

Cilliers, J., Dube, O., & Siddiqi, B. (2015). The white-man effect: How foreigner presence affects behavior in experiments. Journal of Economic Behavior & Organization, 118, 397–414. https://doi.org/10.1016/j.jebo.2015.03.015

Conrad, F. G., & Schober, M. F. (2021). Clarifying question meaning in standardized interviews can improve data quality even though wording may change: A review of the evidence. International Journal of Social Research Methodology, 24(2), 203–226. https://doi.org/10.1080/13645579.2020.1824627

Dean, J. (2021). Noise, cognitive function, and worker productivity. Working Paper. https://joshuatdean.com/wp-content/uploads/2021/03/NoiseCognitiveFunctionandWorkerProductivity.pdf

Down, J., & Duke, S. (2003). SMS polling. A methodological review. ASC, 277–286. https://www.researchgate.net/profile/Joel-Down/publication/268813162_SMS_Polling_A_Methodological_Review_in_ASC_2003_THE_IMPACT_OF_TECHNOLOGY_ON_THE_SURVEY_PROCESS_Edited_by_R_Banks_et_al_Compilation_C_2003_Association_for_Survey_Computing/links/547818e30cf2a961e484930d/SMS-Polling-A-Methodological-Review-in-ASC-2003-THE-IMPACT-OF-TECHNOLOGY-ON-THE-SURVEY-PROCESS-Edited-by-R-Banks-et-al-Compilation-C-2003-Association-for-Survey-Computing.pdf

Duflo, E., & Banerjee, A. (2017). Handbook of Field Experiments. Elsevier.

Dunlap, J. W. (1937). Review of Psychometric Methods.

Edwards, R., & Holland, J. (2020). Reviewing challenges and the future for qualitative interviewing. International Journal of Social Research Methodology, 23(5), 581–592. https://doi.org/10.1080/13645579.2020.1766767

Esopo, K., Mellow, D., Thomas, C., Uckat, H., Abraham, J., Jain, P., Jang, C., Otis, N., Riis-Vestergaard, M., Starcev, A., Orkin, K., & Haushofer, J. (2018). Measuring self-efficacy, executive function, and temporal discounting in Kenya. Behaviour Research and Therapy, 101, 30–45. https://doi.org/10.1016/j.brat.2017.10.002

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). Impact Evaluation in Practice, Second Edition. World Bank Publications.

Gibson, D. G., Pereira, A., Farrenkopf, B. A., Labrique, A. B., Pariyo, G. W., & Hyder, A. A. (2017). Mobile phone surveys for collecting population-level estimates in low-and middle-income countries: A literature review. Journal of Medical Internet Research, 19(5), e139. DOI: 10.2196/jmir.7428

Glöckner, A., & Hilbig, B. E. (2011). Methodology in judgment and decision making research. Judgment & Decision Making, 6(8). http://journal.sjdm.org/11/m00/m00.pdf

Greenleaf, A. R., Gibson, D. G., Khattar, C., Labrique, A. B., & Pariyo, G. W. (2017). Building the evidence base for remote data collection in low-and middle-income countries:

Comparing reliability and accuracy across survey modalities. Journal of Medical Internet Research, 19(5), e140. DOI: 10.2196/jmir.7331

Hammersley, M. (2021). The 'radical critique of interviews': A response to recent comments. International Journal of Social Research Methodology, 24(3), 393–395. https://doi.org/10.1080/13645579.2020.1841881

Haushofer, J., Collins, M., de Giusti, G., Njoroge, J., Odero, A., Onyango, C., Vancel, J., Jang, C., Kuruvilla, M., & Hughes, C. (2014). A Methodology for Laboratory Experiments in Developing Countries: Examples from the Busara Center (SSRN Scholarly Paper ID 2155217). Social Science Research Network. https://doi.org/10.2139/ssrn.2155217

Haushofer, J., Mudida, R., & Shapiro, J. (2020). The Comparative Impact of Cash Transfers and a Psychotherapy Program on Psychological and Economic Well-being. National Bureau of Economic Research. https://www.nber.org/papers/w28106

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. Nature, 466(7302), 29–29. https://doi.org/10.1038/466029a

Hensen, B., Mackworth-Young, C. R. S., Simwinga, M., Abdelmagid, N., Banda, J., Mavodza, C., Doyle, A. M., Bonell, C., & Weiss, H. A. (2021). Remote data collection for public health research in a COVID-19 era: Ethical implications, challenges and opportunities. Health Policy and Planning, 36(3), 360–368. https://doi.org/10.1093/heapol/czaa158

Hilbig, B. E. (2010). Precise models deserve precise measures: A methodological dissection. Judgment and Decision. Judgment and Decision Making, 5(4), 272–284. https://doi.org/10.1.1.177.1293

Höhne, J. K., & Krebs, D. (2021). Mismatching middle options: Consequences for attitude measurement in smartphone surveys. International Journal of Social Research Methodology, 24(3), 381–386. https://doi.org/10.1080/13645579.2020.1803621

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Silver, R. C., & Everall, I. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: A call for action for mental health science. The Lancet Psychiatry. https://doi.org/10.1016/S2215-0366(20)30168-1

Hughes, K., Hughes, J., & Cocq, F. P.-L. (2020). Introduction: Making the case for qualitative interviews. International Journal of Social Research Methodology, 23(5), 541–545. https://doi.org/10.1080/13645579.2020.1766756

Islam, M. (2020). Reflection note: Confessions of a Muslim researcher – considering identity in research. International Journal of Social Research Methodology, 23(5), 509–515. https://doi.org/10.1080/13645579.2019.1700726

Jang, C., & Lynham, J. (2015). Where do social preferences come from? Economic Letters, 137. https://doi.org/10.1016/j.econlet.2015.09.004

Jenner, B. M., & Myers, K. C. (2019). Intimacy, rapport, and exceptional disclosure: A comparison of in-person and mediated interview contexts. International Journal of Social Research Methodology, 22(2), 165–177. https://doi.org/10.1080/13645579.2018.1512694

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. (2020). The shape of and solutions to the MTurk quality crisis. Political Science Research and Methods, 8(4), 614–629. https://doi.org/10.1017/psrm.2020.6

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2009). Handbook on impact evaluation: Quantitative methods and practices. World Bank Publications.

Kianersi, S., Luetke, M., Jules, R., & Rosenberg, M. (2020). The association between interviewer gender and responses to sensitive survey questions in a sample of Haitian women. International Journal of Social Research Methodology, 23(2), 229–239. https://doi.org/10.1080/13645579.2019.1661248

Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., & Brumbaugh, C. C. (2014). Investigating variation in replicability. Social Psychology, 45(3), 142–152. https://doi.org/10.1027/1864-9335/a000178.

Kopper, S., & Sautmann, A. (2020). Best practices for conducting phone surveys. Abdul Latif Jameel Poverty Action Lab (J-PAL). https://www.povertyactionlab.org/blog/3-20-20/best-practices-conducting-phone-surveys#:~:text=Keep%20the%20informed%20consent%20script,and%20duration%20of%20the%20survey.

Kruger, L. J., Rodgers, R. F., Long, S. J., & Lowy, A. S. (2019). Individual interviews or focus groups? Interview format and women's self-disclosure. International Journal of Social Research Methodology, 22(3), 245–255. https://doi.org/10.1080/13645579.2018.1518857

Lau, C. Q., Cronberg, A., Marks, L., & Amaya, A. (2019). In search of the optimal mode for mobile phone surveys in developing countries. A comparison of IVR, SMS, and CATI in NIgeria. Survey Research Methods, 13(3), 305–318. https://doi.org/10.18148/srm/2019.v13i3.7375

Lau, C. Q., Lombaard, A., Baker, M., Eyerman, J., & Thalji, L. (2019). How representative are SMS surveys in Africa? Experimental evidence from four countries. International Journal of Public Opinion Research, 31(2), 309–330. https://doi.org/10.1093/ijpor/edy008

Lau, C. Q., Sanders, H., & Lombaard, A. (2019). Questionnaire Design in Short Message Service (SMS) Surveys. Field Methods, 31(3), 214–229. https://doi.org/10.1177/1525822X19862506

Laumas, A., Owsley, N., & Haldea, P. (2020, July 2). How Preeti was born [Medium]. Busara Center. https://medium.com/busara-center-blog/how-preeti-was-born-af4583208fc9

Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). The Sage encyclopedia of social science research methods. Sage Publications.

Lobe, B., Morgan, D., & Hoffman, K. A. (2020). Qualitative Data Collection in an Era of Social Distancing. International Journal of Qualitative Methods, 19,. https://doi.org/10.1177/1609406920937875

Mao, J., & Feldman, E. (2019). Class matters: Interviewing across social class boundaries. International Journal of Social Research Methodology, 22(2), 125–137. https://doi.org/10.1080/13645579.2018.1535879

Matherly, T. (2019). A panel for lemons? Positivity bias, reputation systems and data quality on MTurk. European Journal of Marketing, 53(2), 195-223. https://doi.org/10.1108/EJM-07-2017-0491

Mathur, M. (2020, April 3). 3 steps to designing an effective phone survey that reaches respondents. IDinsight Ideas and Insights. https://www.idinsight.org/article/3-steps-to-designing-an-effective-phone-survey-that-reaches-respondents/

Mavletova, A., & Couper, M. P. (2015). A meta-analysis of breakoff rates in mobile web surveys. In Toninelli, D. et al. (2015). Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies, 81–98.

Montagni, I., Cariou, T., Tzourio, C., & González-Caballero, J.-L. (2019). "I don't know", "I'm not sure", "I don't want to answer": A latent class analysis explaining the informative value of nonresponse options in an online survey on youth health. International Journal of Social Research Methodology, 22(6), 651–667. https://doi.org/10.1080/13645579.2019.1632026

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD)

Psychology: Measuring and Mapping Scales of Cultural and Psychological Distance. Psychological Science, 31(6), 678–701. https://doi.org/10.1177/0956797620916782

Ndlovu-Gatsheni, S. J. (2021). The cognitive empire, politics of knowledge and African intellectual productions: Reflections on struggles for epistemic freedom and resurgence of decolonisation in the twenty-first century. Third World Quarterly, 42(5), 882–901. https://doi.org/10.1080/01436597.2020.1775487

Nichols, A. L., & Edlund, J. E. (2020). Why don't we care more about carelessness? Understanding the causes and consequences of careless participants. International Journal of Social Research Methodology, 23(6), 625–638. https://doi.org/10.1080/13645579.2020.1719618

Ongena, Y. P., & Dijkstra, W. (2021a). Advances in research on survey interview interaction. International Journal of Social Research Methodology, 24(2), 177–179. https://doi.org/10.1080/13645579.2020.1824625

Ongena, Y. P., & Dijkstra, W. (2021b). Advances in research on survey interview interaction. International Journal of Social Research Methodology, 24(2), 177–179. https://doi.org/10.1080/13645579.2020.1824625

Owsley, N., Adhiambo, J., Owelle, E., Ngina, I., & Amanela, S. (2019, February 20). Off The Record #3: Are cognitive biases universal across cultures? [Medium]. Busara Center. https://medium.com/busara-center-blog/are-cognitive-biases-universal-across-cultures-6b8f7791eb6c

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology, 70, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A., & Damer, E. (2021). MTurk, Prolific or panels? Choosing the right audience for online research. Choosing the Right Audience for Online Research (January 10, 2021). Litman, Leib and Moss, Aaron and Rosenzweig, Cheskie and Robinson, Jonathan, Reply to MTurk, Prolific or panels? Choosing the right audience for online research (January 28, 2021). http://dx.doi.org/10.2139/ssrn.3775075

Pritchard, E. (2019). Female researcher safety: The difficulties of recruiting participants at conventions for people with dwarfism. International Journal of Social Research Methodology, 22(5), 503–515. https://doi.org/10.1080/13645579.2019.1590896

Rasmussen, M. L., Graefenstein, S., Singleton, A., Halafoff, A., & Bouma, G. (2020). Methodological challenges of designing a survey to capture young people's (non-binary) affiliations in relationship to religion, sexuality and gender. International Journal of Social Research Methodology, 23(6), 695–709. https://doi.org/10.1080/13645579.2020.1763692

Richardson, J., Godfrey, B., & Walklate, S. (2021). Rapid, remote and responsive research during COVID-19. Methodological Innovations, 14(1), 20597991211008580. https://doi.org/10.1177/20597991211008581

Rogers, B. A. (2021). Researching while queer: A research note about a genderqueer lesbian conducting qualitative research in the southeastern United States. International Journal of Social Research Methodology, 24(1), 31–37. https://doi.org/10.1080/13645579.2020.1760551

Rozo, S. (2021, July 21). Tips for Collecting Surveys of Hard-to-Reach Populations. Innovations for Poverty Action. https://www.poverty-action.org/blog/tips-collecting-surveys-hard-reach-populations

Sarac, M., & Koc, I. (2021). Sources of nonresponse error in the translation process of survey instruments: The impact of language mismatch and on-the-spot translation on the quality of birth date data. International Journal of Social Research Methodology, 24(3), 313–325. https://doi.org/10.1080/13645579.2020.1785088

Schaeffer, N. C. (2021). Interaction before and during the survey interview: Insights from conversation analysis. International Journal of Social Research Methodology, 24(2), 181–202. https://doi.org/10.1080/13645579.2020.1824626

Silber, H., Danner, D., & Rammstedt, B. (2019). The impact of respondent attentiveness on reliability and validity. International Journal of Social Research Methodology, 22(2), 153–164. https://doi.org/10.1080/13645579.2018.1507378

Simons, D. J. (2014). The Value of Direct Replication. Perspectives on Psychological Science, 9(1), 76–80. https://doi.org/10.1177/1745691613514755

Terentev, E., & Maloshonok, N. (2019). The impact of response options ordering on respondents' answers to rating questions: Results of two experiments. International Journal of Social Research Methodology, 22(2), 179–198. https://doi.org/10.1080/13645579.2018.1510660

Ting, C., & Fitzgerald, R. (2020). The work to make an experiment work. International Journal of Social Research Methodology, 23(3), 329–345. https://doi.org/10.1080/13645579.2019.1694621

Tuttle, K. R. (2020). Impact of the COVID-19 pandemic on clinical research. Nature Reviews Nephrology, 16(10), 562–564. https://doi.org/10.1038/s41581-020-00336-9

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90(4), 293. https://doi.org/10.1037/0033-295X.90.4.293

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., & Druckman, J. N. (2020). Using social and behavioural science to support COVID-19 pandemic response. Nature Human Behaviour, 4(5), 460–471. https://doi.org/10.1038/s41562-020-0884-z

Van Der Heijden, P. (2017). The practicalities of SMS research. International Journal of Market Research, 59(2), 157–172. https://doi.org/10.2501/IJMR-2017-013

Veldkamp, C. L. S., Bakker, M., van Assen, M., Crompvoets, E. A. V., Ong, H. H., Soderberg, C. K., & Wicherts, J. M. (2017). Restriction of opportunistic use of researcher degrees of freedom in pre-registrations on the Open Science Framework. Preprint. https://doi.org/10.31234/osf.io/cdgyh

Wang, R., & Krosnick, J. A. (2020). Middle alternatives and measurement validity: A recommendation for survey researchers. International Journal of Social Research Methodology, 23(2), 169–184. https://doi.org/10.1080/13645579.2019.1645384

Wein, T., Owsley, N., Escande, A., Wahome, L., Wanjeri, P., Okaba, C., Kimani, J., Mwesige, M., & Wanja, D. (2019). Primed to Engage: Rights and capability language encourage civic attention, but not civic engagement, in a laboratory experiment in Uganda. Working Paper. https://drive.google.com/open?id=19cYnvZX2B3UG9_18I4tSUKIvAhusdbSc

Wijsen, L. D., & Borsboom, D. (2021). Perspectives on Psychometrics Interviews with 20 Past Psychometric Society Presidents. Psychometrika, 86(1), 327–343. https://doi.org/10.1007/s11336-021-09752-7

World Bank. (2021). World Development Report 2021: Data for Better Lives. The World Bank. https://www.worldbank.org/en/publication/wdr2021