EDITED BY
PATRICK S. FORSCHER AND MARIO SCHMIDT

A BETTER HOW

NOTES ON DEVELOPMENTAL META-RESEARCH

A BETTER HOW

NOTES ON DEVELOPMENTAL META-RESEARCH

EDITED BY

PATRICK S. FORSCHER AND

MARIO SCHMIDT

First published in Kenya 2024 by Busara.
38, Apple Cross Road, Lavington, Nairobi, Kenya.

All Busara publications are available on the Busara website: www.busara.global

To cite this book:
Forscher P.S. & Schmidt, M. (eds). (2024). A better how: notes on developmental meta-research. Busara. DOI: doi.org/10.62372/ISCI6112

To cite a chapter from this book:
<Author names, Author names.> <Essay title>. (2024). In P.S. Forscher & M. Schmidt (eds), A better how: notes on developmental meta-research (pp XX-XX). Busara. DOI: doi.org/10.62372/ISCI6112

ISBN: 978-9914-766-06-6

The views presented in this publication are those of the author(s) and do not necessarily reflect the positions of our diverse set of funders and supporters. The editors appreciate the support and help of Michael Onsando who played a crucial role in publishing this volume.

Edited by Patrick S. Forscher and Mario Schmidt
Copyedited by Aarti Nair
Designed by Micheal Bagorogoza

TABLE OF CONTENTS

Part 2

How to improve the evidence-to-policy pipeline

Part 3

How to improve the ways in which we conduct research

INTRODUCTION

Notes on developmental meta-research

Patrick Forscher (Busara) and Mario Schmidt
(Busara & Max Planck Institute for Social Anthropology)

The past decade has raised concerns about how research is conducted, evaluated, and disseminated. Fuelled by the replication crisis in psychology and allied disciplines, these concerns have spawned a movement that unites scholars from across the globe: the open science movement. The movement has produced and popularized a huge array of innovations to enhance the replicability of research and has even caught the notice of several large institutional actors. For example, the Biden administration declared 2023 the 'year of open science' in the United States. Internationally, UNESCO has collated and issued comprehensive recommendations for how member states can use and incorporate open science into policymaking. This interest has set in motion a broader movement and a dedicated academic subfield to improve how research is used in society. This movement, the meta-research movement, is the subject of this book.

Meta-research, sometimes called meta-science, is research focused on investigating the research process itself, often aiming to make concrete improvements. These improvements have, to date, primarily focused on improving the fundamental soundness of academic research. Due to its roots in the behavioral sciences, these improvements are also often behaviorally informed. For example, the meta-research innovation called 'preregistration' involves a precommitment to a particular way of analyzing data before seeing it. This innovation is designed to reduce the risks that the analyst intentionally or unintentionally changes the analysis plan after seeing how the data turns out to suit their preferred interpretation. Although this innovation focuses on the basic soundness of research, the improvements sought by meta-researchers can, in principle, involve anything – including

the problem areas that are the traditional focus of the global development community, such as north-south power differences, building healthy research ecosystems, and the treatment of participants or beneficiaries.

Our team has spent the past two years investigating the potential intersections between meta-research and global development. We discovered that many communities within global development have already been concerned about how to improve how research is conducted and used without necessarily realizing that they have a common cause with other, similar communities. Likewise, many meta-researchers would benefit from learning how research is used outside academic contexts. We believe this intersection between meta-research and global development is a fruitful one and has the potential to constitute an entirely new subdiscipline, which we call ***developmental meta-research***. This subdiscipline turns meta-research's behaviorally-informed critical lens toward topics that have traditionally been the focus of global development practitioners. We believe the resulting subdiscipline can fruitfully inform intersections between development practitioners and meta-researchers and mobilize a new community around improving how research is done in development.

> **Meta-research, sometimes called meta-science, is research focused on investigating the research process itself, often aiming to make concrete improvements**

This volume brings together various contributors from across development and meta-research. These contributors span different sectors, institutions, countries, and problem areas. What unites them is a shared focus on improvement: in the distribution of power, how evidence is used to inform policy, and regarding the overall conduct of research. We have categorized the contributions into three clusters, each focusing on a different dimension of what it means to strive for a 'better how' by considering how research is produced, implemented, and disseminated. Contributions to our first part, ***How to improve collaboration across the Global North-South divide,*** discuss ways local research infrastructures can be strengthened through collaborations across the Global North-South divide, and critically interrogate the potentials and pitfalls of science across economic, political, and linguistic barriers. Contributions in this section ask uncomfortable questions about who benefits from collaborations across these barriers and how the research capacity in low- and middle-income countries can improve through such collaborations. In our first contribution to this section, *Advantages and challenges of Global North/Global South research collaborations: an emphasis on sub-Saharan African research infrastructure*, Dana M. Basnight-Brown uses her own experiences of working across the Global North/Global South divide to offer advice on how to avoid misunderstandings and mismatched priorities that often emerge in collaborations between high-income countries and low-and middle-income countries. Shifting the focus away from Africa, our next contribution – Amirun Haqqim bin Eldeen Husaini and Miguel Silan's *'Should you study abroad?' The mechanisms and utility of*

educational emigration from Southeast Asia – sheds light upon the political-economic effects of educational migration on local research infrastructures in Southeast Asia. Complexifying one-sided answers to the consequences of educational migration, the authors argue for a more nuanced view on the topic, highlighting both potential benefits and dangers for local, underfinanced research infrastructures. The successive contribution *Building local networks for open science: a case study of the Chinese Open Science Network (COSN)*, was written by a group of authors from the *Chinese Open Science Network* (Zhiqi Xu, Yue Wang, Liangyou Zhang, Wenqianglong Li, Hu Chuan-Peng, Chenghao Zhou and Xi Chen). It shares crucial insights on how comparatively cost-effective ways could improve local research infrastructures if enough early career researchers team up together and embrace the principles of open science. Among the initiatives discussed are translating important English-speaking articles into Chinese, tutorials, conference meetings, and increasing the use of social media platforms. How such efforts toward a more just and open science could look if implemented transnationally is explored in the contribution by Alma Jeftic, Marc Yancy Lucas, Nadia Corral-Frías, and Flavio Azevedo. Their paper *Bridging the majority and minority worlds: liminal researchers as catalysts for inclusive open and big team science* introduces two collaborative efforts – FORRT (The Framework for Open and Reproducible Research Training) and ABRIR (Advancing Open and Big-Team Reproducible Science through Increased Representation) – and illustrates the pivotal, yet underestimated, role of what they call 'liminal researchers' who have intimate knowledge of different research environment in the Global South and North. Early career researchers, however, also depend on more senior colleagues for advice, mentorship, and guidance on navigating science and its relationship to society, development, and their lives. Sharing her own experience with founding and leading the African-led research mentorship organization *Eider Africa* in her contribution *Contextually grounded research in postgraduate research training in Africa: why and how,* Aurelia Munene calls for more contextually grounded research that dismantles colonial knowledge hierarchies and helps African scholars to move from the margins to the center of global knowledge production. Along similar lines but specifically focusing on clinical psychology, Helen Niemeyer and Louis Schiekiera's paper *How inclusive and equitable is research in clinical psychology that focuses on the Global South?* criticizes that clinical studies in the Global South are rarely informed by local knowledge and often only implement superficial adaptations, such as changing names in narratives. Considering the potentially dramatic consequences of mental health studies, Niemeyer and Schiekiera suggest integrating local researchers into teams and taking their knowledge seriously. The last contribution of our first part, Maya Ranganath's *Collaborating to support a more inclusive evidence ecosystem,* introduces the *Center for Effective Global Action's* initiative *Collaboration for Inclusive Development Research* that aims to produce knowledge about the inclusion of African scholars in global development along different stages with the ultimate aim to improve research, impact the evidence-to-policy pipeline, and design actionable guidelines to ultimately transform the development space into a more diverse, equitable, and inclusive ecosystem.

Our second part ***How to improve the evidence-to-policy pipeline*** zooms in on the question of

how we can ensure that the evidence that informs policy is sound and inclusive, that it addresses issues of concern to the ultimate users of evidence, and that the ecosystem for synthesizing evidence into insights is healthy and robust. In their contribution to the edited volume *Behavioral public policy for global challenges*, Sanchayan Banerjee and Matteo M. Galizzi focus on how behavioral interventions can be scaled up to target the global problems of our time. Suggesting three different ways to accomplish this – expanding the toolkit of behavioral public policy, assessing the heterogeneity in treatment effects of behavioral interventions, and implementing systematic, reproducible, and transparent multi-country experimentations – they urge all of us not to lose sight of the bigger picture. Focusing on a significant 'big team science' effort across several African countries which he led, our next author, Adeyemi Adetula, centers his contribution *Building research capacity in Africa via big team science: challenges and lessons learned from the ManyLabs Africa initiative* around the challenges his team faced when trying to boost local interest and skills in open science practices and when generalizing from sub-Saharan African scholarship to the West rather than the reverse. Among these challenges, Adetula highlights political-economic incentives for African researchers to focus on their careers and infrastructural problems, such as poor internet access, insufficiently staffed labs, and expensive ethics approvals. Tommie Yeo Thompson, Winnie Mughogo, and Anisha Singh's *The gender data gap in development policy research* zooms in on what we still too often forget: changing the behavior of men and women does not just require different interventions. We must also be careful to do justice to the gendered dimension of human behavior when we choose our sample and measurement tools. Enlarging our toolbox is something our next two contributions suggest as well. In their paper *Assessing deliberative polling methodology for decolonizing and indigenizing research*, Dennis Chirawurah and Niagia F. Santuah discuss the decolonizing potential of the Deliberative Polling Method to open up the space for local and contextual knowledge to emerge. Starting from the premise that whenever policy choices have a consequence for local communities, these communities should be consulted on their opinion about it, the Deliberative Polling Method urges researchers and those implementing policies to embrace the tacit and explicit knowledge of indigenous people if they want interventions that have a lasting impact. *Evaluating interventions: a practical primer for specifying the smallest effect size of interest*, written by a large team of authors spearheaded by Hannah K. Peetz and Maximilian A. Primbs, urges us to move beyond conventionally determined ways of interpreting the effect sizes used to measure the results of quantitatively assessed interventions, such as Cohen's benchmarks. As an alternative, they consider a different kind of benchmark that better allows for considering contextual and other kinds of knowledge: the smallest effect size that important stakeholders, such as researchers, consider to be meaningful in a specific context. Introducing three ways to determine the smallest effect size of interest – person as effect size, cost-benefit analyses, and the minimal important difference – they suggest moving beyond mindless benchmarks when judging the efficacy of interventions. The last contribution of our second section, Anushka Ghosh's *Utility of meta-research for Global South policymaking: a reflection on education research*, scrutinizes the potential of meta-research to influence decisions made by researchers and policymakers in the wider field of education.

Reflecting upon her own experience as a researcher from the Global South, she urges us to make use of meta-research not only to improve research and policies but also as a tool to open up space for critical distance and reflection upon political hierarchies and colonial legacies still impacting how research and policy is done.

> **Quality of research, in other words, is not only a necessary part of our way to accomplish development goals but, to a significant degree, also depends on thinking about what constitutes good research and the ecosystems that are necessary to support it. It depends, in other words, on developmental meta-research.**

Papers in our third and last section ***How to improve how we conduct research*** take a closer look at how we can improve research epistemologically, ethically, and methodologically. In his contribution, *The democratizing effects of doubting*, Adam Moe Fejerskov shares his uneasiness about global development's fascination with and focus on certainty, suggesting that we would all benefit from embracing a more humble attitude that leaves space for doubt. One potential way to increase one's embrace of doubt might be to engage in more intimate ethnographic methods, a form of research advocated for by Ben Jones and Ben Eyre together with their team of 'citizen ethnographers' (Sharon Acio Enon, Dorah Adoch, Vicky Alum, Joel Ekaun Hannington, Jimmy Ezra Okello, Robert Oluka, James Opolo, and Ann Gumkit Parlaker). Their paper *'Here, my degree does not matter, you are the teacher': ethnography, citizen ethnography, and researching research on global development* argues for the need to train and mentor local citizens in ethnographic methods to increase the context-sensitivity of development projects. Mario Schmidt's contribution *Do the randomized know they are randomized? A critique of the turn towards randomization in high-stakes development cooperation initiatives* picks up on Fejerskov's embrace of doubt and Eyre's call for more ethnography by highlighting respondents' (mis)understanding of the randomization process during an unconditional cash transfer program in western Kenya with ethnographic detail. Also, taking a Kenyan cash transfer program as its empirical starting point, our next contribution *Manage relationships when starting and ending research with human participants* by Joel Wambua, Anisha Singh, Kelvin Kihindas, Irene Gachungi, and Patrick S. Forscher, discusses ways to improve the relationship between researcher and researched. The authors particularly focus on the moments of community entry and exit, and caution us that research projects characterized by large gaps between the knowledge and expectations of researchers and those of the researched are particularly susceptible to cause harm if approached naively. Jason M. Chin's paper *Why applied psychologists should consider their work's value-laden context* also focuses on a specific case, namely a consequential law and psychology study that occurred in the context of an inquiry by the Australian government called the Royal Commission into Institutional Responses to Child Sexual Abuse (RCIRCSA). Chin questions the neglect of non-epistemic values coupled with an overreliance on scientific conventions when evaluating interventions, such as specific standards of evidence

sufficiency (for instance, an alpha level of 0.05), irrespective of the value-laden context these interventions appear in. Our next two authors, Symen Auke Brouwers and Floriza Freire Gennari, also argue for more context-sensitivity in global development. In their contribution *Cultural context and ecological validity in global development research*, they argue that we should never naively assume that an initiative that worked in one place will also work elsewhere even if the two areas are similar. Rather, they suggest a critical analysis of context to optimize ecological validity through a mixed-methods approach combining qualitative and quantitative approaches. Our last contribution, Joel Wambua's *Consent, open-ended questions, and feedback loops: empirical insights into research ethics in the Global South*, shares the first results of an experiment in Nairobi testing the effects of minor improvements to research protocols – particularly focusing on changes to consent procedures, the inclusion of open-ended questions, and different types of feedback sent to the participants after the study – and how these change participants' perceptions of ethical practice. The experiment conclusively finds no change as a result of any of the interventions, suggesting that, in this context at least, small changes are insufficient to impact participant experience – even though participants consistently ask for improvements in consent, the inclusion of open-ended questions, and feedback, including in this experiment. Wambua urges us to focus more directly on what participants mean when they ask for improvements to their research experience, as small tweaks to the research experience may be insufficient to fully respond to these requests

If there is one thing we wish our readers to learn from this edited volume, it is that meta-research and global development are neither opposed nor mere add-ons to each other. Instead, meta-research and global development are integral to one another. It is this integration that forms the ***developmental meta-research*** subfield that we have identified over the past two years and in the preparation of this volume. Today, we not only witness increasing calls for evidence in the development cooperation sector but also observe the production of thousands of scientific studies evaluating interventions. We can thus no longer afford the luxury to either 'do' meta-research or 'do' development. If we do not reflect on how we construct robust and supportive research ecosystems that support researchers in all parts of the world, build the tools and workflows so that those researchers can plan and implement sound and useful research, and create the infrastructure and partnerships to ensure that research can be used in the places where it matters the result will be a development sector that relies on hearsay and conventions. Quality of research, in other words, is not only a necessary part of our way to accomplish development goals but, to a significant degree, also depends on thinking about what constitutes good research and the ecosystems that are necessary to support it. It depends, in other words, on developmental meta-research.

We have put together this edited volume that zooms in on the relationship between meta-research and global development to reflect upon these urgent questions. The volume tries to catalyze debates about decolonizing development and science in the Global South, boost the open science movement and its principles globally, nurture emergent local

research ecosystems, and ask critical questions about methods and ethics. We hope that the contributions will spark a debate about the potential of developmental meta-research, opening new collaborations across subfields, geographies, and sectors.

PART ONE

How to improve collaboration across the Global North-South divide

Advantages and challenges of Global North/Global South research collaborations: an emphasis on sub-Saharan African research infrastructure

Dana M. Basnight-Brown (United States International University – Africa)

During the last decade, researchers have brought greater attention to the fact that individuals from the Global North dominate many research endeavors (for example, authorship, participant samples, editorial boards, and journal ownership, cf. Arnett, 2009; Lin & Li, 2023; Rad et al., 2018). Using the African continent as an example, Thalmayer et al. (2021) report that less than 1% of first authors in top psychology journals were from Africa, despite the continent contributing to roughly 17% of the global population. This is particularly striking as many studies within the behavioral sciences claim to report generalizable findings of human behavior, which has implications for educational systems, policy, healthcare outcomes, and overall well-being. However, with such a limited slice of the human population participating in such studies, it is impossible to accurately understand these issues. This is detrimental to our understanding of various scientific outcomes, affecting research growth and capacity building in parts of the world seeking to advance their research footprint. Despite an increase in international collaborations in recent years, there is still a lot of work to be done to better understand the advantages and challenges of low/middle-income countries (LMIC) and high-income countries (HIC) collaborations. To remedy this important problem, we need more robust international collaborations that truly engage researchers, research administrators, and local samples from world regions that are under-represented in the scientific literature. Yet, when collaborations of this nature are created, cultural, governmental, and institutional processes and priorities often differ across collaborators, primarily from HIC versus LMIC contexts. Many times, those from HICs have a more well-established research infrastructure and better-funded research programs, and, as a result, the impression is that all research processes operate similarly.

The advantages and challenges that emerge in HIC-LMIC collaborations can lead to misunderstandings and mismatched priorities and cause problems in achieving research goals. This is the main problem that I seek to address by highlighting some of the advantages and challenges that dominate international research collaborations. Many of my examples come from psychological research findings, as that research domain has a lot of implications for better understanding human behavior globally, which coincides with my research expertise. I have spent over a decade working at an institution of higher education in an LMIC (Kenya), where I worked with international collaborators at various non-profit research organizations and was directly involved in developing the first Institutional Review Board (IRB) at a local institution. I hope that highlighting some of these issues will be

informative and beneficial to everyone engaged in international collaborations and that awareness of the existing differences will foster stronger research partnerships in the future.

International research collaborations and team science initiatives

Many scholars agree that international research collaborations and team science initiatives play a dominant role in shaping how science will change in future decades. As Adams (2012) points out, international research collaborations and papers have risen dramatically in recent decades, with growth showing no signs of slowing down. In the United States, for example, Web of Science data indicate that partnerships with China, the UK, and Germany comprise a large portion of their international research output. In Kenya, the top international collaborators were the United States, the UK, and China. In addition to the rise in international collaborations, there was a surge of large multi-authored research initiatives, often referred to as 'Big Team' science (Adams 2012). Although these two are not identical, they are frequently correlated. Big Team Science (BTS) projects are likely to involve those from multiple countries, and both are potential drivers for shifting scientific output. Recently, researchers have begun to document the challenges of team science, describing how they require a lot of management resources, need to find ways to be sustainable, and could be sensitive to unaccountable leadership (Coles et al., 2023; Forscher et al., 2023). Despite these difficulties, BTS efforts have the potential to help us solve the scientific challenges of our time, as they allow us to pool resources, whether it be intellectual, financial, or time, resulting in stronger and more impactful scientific outcomes.

Advantages of HIC and LMIC international collaborations

As noted earlier, most published scientific literature focuses on research samples and authors from Western societies. Henrich et al. (2010) report that '96% of psychological samples come from countries with 12% of the world's population'. To see whether the data changed a decade later, Thalmayer et al. (2021) investigated the same journals, revealing that roughly 11% of the global population was represented in the literature, indicating that there is still a long way to go in meeting more equitable and diverse standards in the study of human behavior. As a result, the field of psychology, as well as other scientific disciplines, needs to rely on more sustainable and ethical international collaboration in the future, which raises the question of the advantages of LMIC-HIC collaborations.

A better understanding of human behavior

Any study of human behavior focused on a limited population results in a weak and potentially inaccurate understanding of psychological processes that affects our sense of developmental milestones, decision-making, cognitive processing, and social interactions. For example, in my research, an investigation of how East African multilinguals express positive and negative emotions in their first, second, and third languages was different from outcomes reported in the literature that focused on North American bilinguals, as well as different from patterns observed in Asian populations (Basnight-Brown et al., 2022). These outcomes were useful as they contributed to our understanding of language processing

and expanded it to include researchers and participants with vibrant linguistic experiences to share. In addition, obtaining a better understanding of human behavior through more diverse international collaborations with those in LMICs has advantages for students in those regions as well. For example, many students often comment on the difficulties in obtaining scientific literature that pertains to local issues and populations, even finding that so-called 'international' textbooks are still lacking in their coverage of locally relevant research. As one of my students aptly put it: 'We always complain that reading materials don't give relevant examples from an African context, and so we need to start creating such materials, and research plays a huge role in this' (USIU-Africa student, 2020). Expanding our research landscape to include researchers from less-represented geographic regions has the potential to strengthen our understanding of various scientific phenomena, something that can only benefit science overall.

New insights from local collaborators

When collaborating with people from other geographic regions and cultures, there is the opportunity to gain meaningful insights that expand one's understanding of how research might be done differently and, more importantly, may reframe goals of the research project. For example, within East Africa, much emphasis is placed on the United Nation's Sustainable Development Goals (SDGs, United Nations, n.d.). Understandably, many global leaders rely heavily on the higher education sector to incorporate these goals into their teaching curriculum and research priorities. In my experience, many students in East Africa are very familiar with these goals and attempt to incorporate them into their research. This is supported by research that revealed that participants from Africa and Asia were most likely to use SDGs in their work and reported that SDGs play an important role in student leadership, policy development, and community engagement at their institutions (Schrieber et al., 2022).

Despite some regions of the world having more familiarity and dedicating greater attention to the SDGs, progress on them globally has not been at the pace initially envisioned by the UN. This led the International Science Council (2023) to suggest that a BTS approach is likely needed to achieve these goals. As their report mentions, this would be particularly advantageous for the Global South, supporting the idea that HIC-LMIC collaborations will likely be integral to advancing these goals. Greater emphasis on the SDGs in many LMICs suggests that researchers from HICs may need to adjust their projects' research agenda and goals better to match the priorities of collaborators in other regions. Not only is this important for the strength of the collaboration, it may also affect whether IRB approval is provided. For example, at USIU-Africa, the IRB Standard Operating Procedures (2019) state that research must be 'guided by humanitarian and equity-based concerns, not by a pursuit of knowledge for its own sake.' In addition, they outline that research projects must be relevant to the community being studied and that they must show a direct community benefit. This is important as it may push researchers in HICs to expand their thinking to more carefully consider how a certain research project can genuinely benefit a population.

In addition to a shift in how research projects are conceived and contextualized, insights from local collaborators also provide essential information to learn from. As Adetula et al. (2022) suggest, psychological effects and findings that originated in Africa should be replicated outside of Africa, resulting in more bi-directionality in the research approach. In the past, findings that originated in the Global North have been 'exported' to Africa to see whether they replicate there. However, there is no reason why research cannot be reciprocal, allowing both HIC and LMIC researchers to learn from findings in both contexts.

> **By expanding research samples, a better understanding of human behavior can be obtained while also gaining new insights from local collaborators who may push teams to rethink how a particular project may impact a specific community.**

Embracing open science practices

Another advantage of international collaborations is that they have the potential to increase openness to 'Open Science' practices, such as open-access publishing and open data. Of course, adopting an open science approach is not reliant on HIC-LMIC partnerships. Yet, because many LMICs rely more on open science, collaboration can be a natural outcome. For instance, open-access publishing distributes knowledge faster and more widely while keeping costs down. Since many LMIC institutions do not have large budgets to subscribe to journals that are behind expensive paywalls, there are often more significant incentives to utilize open-access opportunities. Across Africa, some open-access processes emerged from government directives. For example, in 2019, the Ethiopian government developed an open-access policy requiring all federally funded research to be published in open-access journals. They were the first African country to require this, and perhaps even more striking is that this decision was made before the National Science Foundation's (NSF) directive to adopt a similar procedure for federally funded research in the United States.

As Grahe et al. (2020) describe, another advantage of open science is that it provides more opportunities to advance diversity in research. Although this is likely true, as Onie (2020) aptly points out, the benefits of open science should not be based on shifting already well-established research systems to LMIC researchers but also on seeking to design new ones that fit the challenges and issues within the culture. As he describes, national policies can play an important role in this process as local researchers and institutions should decide what to implement. Using Indonesia as an example, Onie (2020) explains that open data has not been well received within the country, often because some researchers have used the data to publish papers without giving credit and seeking permission from the original authors. A national data repository was created to curb this behavior, allowing each dataset to be labeled by the owner(s). Processes were also put in place to train individuals on how to access and use the data ethically.

In summary, several key advantages of HIC-LMIC research collaborations have the potential to benefit researchers from all parts of the globe. By expanding research samples, a better understanding of human behavior can be obtained while also gaining new insights from local collaborators who may push teams to rethink how a particular project may impact a specific community. In the following section, some of the challenges of HIC-LMIC collaborations will be discussed.

Challenges of HIC and LMIC international collaborations

Despite the advantages of HIC-LMIC collaborations, challenges emerge in these partnerships – particularly due to differences in processes or cultural factors. Some of the most pressing challenges focus on areas of ethics approvals, limited resources and training available, funding issues, and communication differences.

Ethics approval

In many LMICs, research infrastructure may still be in the early stages of development, such that formalized ethics boards may be new or non-existent. In Kenya, for example, Institutional Review Boards (IRB) have only been formalized at institutions in recent years after the government issued a directive to require this of all universities. Although some public institutions had existing review boards, this was a new initiative for many Kenyan institutions. Despite this advancement, many African countries do not have IRBs, so research is conducted without this oversight. Navigating the different regulations across countries, especially when extensive multi-site studies are being performed, can be challenging for all researchers, but primarily for those from HICs accustomed to well-established research offices with well-documented IRB procedures. In addition to nascent IRB practices and policies, obtaining IRB approval in some locations can be time-consuming and costly. Often, relatively new research offices in LMICs are not well-staffed, therefore, it can take a long time for researchers to receive IRB approval. Second, institutions in Kenya and several surrounding East African countries often charge IRB fees for local students, faculty, and international collaborators, all of which can be expensive for researchers.

Finally, as described earlier, the goals of the research office or local organization in some LMIC contexts may be more community-focused or applied, factors that could influence whether ethics approval is issued for a project. This is important for all stakeholders, as the project must provide some type of benefit for the population. For example, in the study of child development, Weber et al. (2021) recently described how UNICEF's Care for Child Development Program has been criticized for trying to implement caregiving practices and values that originated in HICs to children in LMIC environments, with little regard for the communities' goals for their children or local cultural practices. While it is important to remember that the IRB is not there to serve as a roadblock but rather to protect, it is also important to remember that IRBs unfamiliar with a specific culture may not accurately evaluate all risks that come with a certain project or program. For this reason,

HIC collaborators will need to rely on LMIC ethics boards to meet IRB standards within a country and cultural standards of goodwill, even if those processes may be more time-consuming or different.

Resources: funding, space, and training

Institutions in all parts of the world fight for limited resources to maintain their research programs, yet due to the economic situation in many LMICs, resources related to funding opportunities can be extremely limited. In Africa, governments only marginally fund Research and Development (R&D), such that recent data indicate that Africa's overall R&D funding stands at roughly 0.42% of gross domestic product (GDP), with no country on the continent even reaching 1% (Kigotho, 2021). For reference, the World Bank reports that the global R&D expenditure stood at 2.63% for 2020. As a result, many Sub-Saharan African countries rely on funding opportunities from international sources, suggesting that international collaboration is an essential component of their research infrastructure.

In addition to funding challenges, many LMIC institutions are plagued by limited resources in the form of space, research infrastructure (whether personnel or equipment), and training opportunities, which can be a major hurdle for conducting basic research requiring lab space. Another resource related challenge refers to the level of support that researchers receive from research offices, as these are usually not as well-staffed as compared to those at research-intensive institutions in HICs. This creates more constraints for faculty, who do not have the same level of support in terms of grant writing, drafting budgets, and generating funding opportunities. This is supported by an investigation of the research administration profession, where Kerridge and Scott (2018) observed that regions of the world that included Africa (coded as 'rest of the world' in their study) had the greatest number of individuals who worked in the RA profession while simultaneously having another role at the institution. These institutions also had the fewest full-time research administrators, indicating that institutions in many LMICs have limited research capacity, a major challenge for administrators and faculty.

Overall, many institutions in LMICs face enormous challenges concerning the research resources, infrastructure, and training available to them. Partners from HICs must be aware of this to avoid entering partnerships and getting frustrated when these limitations affect the project's success. Therefore, it is important to discuss these issues before developing collaborations so that measures to accommodate all stakeholders can be put in place early.

Differences in expectations and communication

The potential for HIC and LMIC partners to differ in their expectations for a research project is another challenge worth examining, one that focuses more on the soft skills of research partnerships. These differences are likely to be more striking when individuals from different cultures work together, so it is important to set the tone for the project by carefully listening to input and experiences from everyone involved. This might require asking in the

early stages of development, 'What will success look like for each partner?' This is crucial, as for some, it may involve the creation of something, perhaps an intervention or a program, while for others, the main goal may be to engage students in research.

Using the field of psychology as an example, expectations may also not line up due to differences in how researchers from a HIC versus those from a LMIC view a discipline. In Kenya, the term 'psychological science' is not widely used to describe the psychology discipline. Many African institutions consider the field to be part of the humanities, whereas in the United States, some areas of psychological science are recognized as STEM fields. This difference obviously affects how the discipline is viewed and prioritized within each country.

Other challenges may emerge in how individuals communicate. As Meyer (2014) points out, countries differ in whether they are high- or low-context, with the former being more relationship-oriented. Individuals operating in these contexts may say yes but mean no. It has been suggested that one way to minimize this communication challenge is to ask open-ended questions. In addition, others have pointed out that trust is often hard to build (Sloan & Alper, 2014). This is likely due to several factors, one being that some populations in LMICs have historically been exploited for the advancement of research, which naturally erodes trust for future participants and generations. Second, corruption in daily life can be a challenging issue that is more prevalent in some countries than in others; therefore, people are less likely to believe that individuals will do as they say.

In summary, various advantages and challenges occur when researchers from LMICs and HICs form research partnerships. Due to differing perspectives, experiences, cultural norms, and processes that dominate the research context for each stakeholder, researchers may require a new level of perseverance, planning, and creativity. However, for those who are up to the challenge, the outcomes can be extremely rewarding and can have a tremendous impact on changing the way research is conducted and creating a research environment that is truly more inclusive.

Conclusions

Through highly collaborative efforts by those in HIC and LMIC contexts, we can gain new insights from local collaborators, learn to reframe research questions, and approach research endeavors with creativity. However, to achieve this, partners must ensure that research projects are relevant to the local culture and engage the community in beneficial ways. As Mughogho et al. (2023) stress, local researchers should be treated as full partners in the research process. This introduces an important point, as many of the issues raised in this commentary focus on things that researchers from HICs need to consider. Yet, it is equally important to recognize that researchers from LMICs need to take ownership of how the power they have influences the research landscape globally and in their geographic regions. As African scholars Nhemachena and Mawere (2022) highlight, many researchers in Africa utilize ideas produced by those from North America and Europe, trying to apply those findings to African societies without creating their own. They argue that researchers

on the African continent need to do more for their own theory creation, indicating that everyone must share the responsibility for research growth and development.

In closing, I hope the information provided in this commentary will spark fruitful conversation on how the advantages and challenges addressed can strengthen research partnerships. Increased contributions from LMIC research partners are expected to be a vital part of international research development moving forward. For this reason, working in highly collaborative international teams will only increase in importance, and we will all be better for it.

References

Adams, J. 2012. The rise of research networks. *Nature* 490, 335-6. https://doi.org/10.1038/490335a

Adetula, A., Forscher, P.S., Basnight-Brown, D.M., Azouaghe, S., & IJzerman, H. 2022. Psychology should generalize from – not just to – Africa. *Nature Reviews Psychology* 1(7), 370-1. https://doi.org/10.1038/s44159-022-00070-y

Arnett, J.J. 2009. The neglected 95%: a challenge to psychology's philosophy of science. *American Psychologist* 64(6), 571-4. https://doi.org/10.1037/a0016723

Basnight-Brown, D.M., Ayoob, A., & Altarriba, J. 2022. Emotion processing in a highly proficient multilingual Sub-Saharan African population: a quantitative and qualitative investigation. *Languages* 7, 280-97. https://doi.org/10.3390/languages7040280

Coles, N., DeBruine, L., Azevedo, F., Baumgartner, H. & Frank, M. 2023. Big team science challenges us to reconsider authorship. *Nature Human Behaviour* 7, 665-7. https://doi.org/10.1038/s41562-023-01572-2

Forscher, P.S., Wagenmakers, E-J., Coles, N.A., Silan, M.A., Dutra, N., Basnight-Brown, D.M., & IJzerman, H. 2023. The benefits, barriers, and risks of big team science. *Perspectives on Psychological Science* 18(3), 607-23. https://doi.org/10.1177/17456916221082970

Grahe, J., Cuccolo, K., Leighton, D.C., & Cramblet-Alvarez, D. 2020. Open science promotes diverse, just, and sustainable research and educational outcomes. *Psychology Learning & Teaching* 19, 5-20. https://doi.org/10.1177/1475725719869164

Henrich, J., Heine, S.J., & Norenzayan, A. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Institutional Review Board. 2019. *Institutional Review Board Standard Operating Procedures*. Nairobi: United States International University – Africa.

International Science Council. 2023. *Flipping the science model: a roadmap to science missions for sustainability*. Paris: International Science Council. https://council.science/publications/flipping-the-science-model

Kerridge, S., & Scott, S.F. 2018. Research administration around the world. *Research Management Review* 23, 1-34.

Kigotho, W. 2021 (June 17). Countries spend less than 1% of GDP on research. *University World News*. https://www.universityworldnews.com/post.php?story=20210616151534847

Lin, Z., & Li, N. 2023. Global diversity of authors, editors, and journal ownership across subdisciplines of psychology: Current state and policy implications. *Perspectives on Psychological Science* 18(2), 358-77. https://doi.org/10.1177/17456916221091831

Meyer, E. 2014. *The culture map*. Public Affairs Publishing.

Mughogho, W., Adhiambo, J., & Forscher, P.S. 2023. African researchers must be full participants in behavioral science research. *Nature Human Behaviour* 7, 297-9. https://doi.org/10.1038/s41562-023-01536-6

Nhemachena, A., & Mawere, M. 2022. Academics with clay feet? Anthropological perspectives on academic freedom in twenty-first century African universities. *Journal of African American Studies* 26, 142-65. https://doi.org/10.1007/s12111-022-09584-4

Onie, S. 2020. Redesign open science for Asia, Africa, and Latin America. *Nature* 587 (7832), 35-7. https://doi.org/10.1038/d41586-020-03052-3

Rad, M.S., Martingano, A.J., & Ginges, J. 2018. Toward a psychology of *Homo sapiens*: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences, USA* 115(45), 11401-5. https://doi.org/10.1073%2Fpnas.1721165115

Schrieber, B., Moscaritolo, L. B., Perozzi, B., & Luescher, T. M. 2022 (September 17). Student affairs and services commit to advancing SDGs. *University World News*. https://www.universityworldnews.com/post.php?story=20220913110205173

Sloan, S.S., & Alper, J. 2014. *Culture matters: international research collaboration in a changing world*. National Academies Press.

Thalmayer, A.G., Toscanelli, C., & Arnett, J.J. 2021. The neglected 95% revisited: is American psychology becoming less American? *American Psychologist* 76, 116-29. https://doi.org/10.1037/amp0000622

The World Bank. n.d. *Research and development expenditure (% of GDP)*. Retrieved October 24, 2022, from https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?most_recent_value_desc=true

United Nations. n.d. *The 17 goals*. Retrieved February 2, 2023, from https://sdgs.un.org/goals

Weber, A.M., Diop, Y., Gillespie, D., Ratsifandrihamanana, L., & Darmstadt, G.L. 2021. Africa is not a museum: the ethics of encouraging new parenting practices in rural communities in low-income and middle-income countries. *BMJ Global Health* 6(7), 1-7. https://doi.org/10.1136/bmjgh-2021-006218

‘Should you study abroad?’ The mechanisms and utility of educational emigration from Southeast Asia

Amirun Haqqim bin Eldeen Husaini (The University of Nottingham, Malaysia Campus & University of the Bundeswehr, Munich) and Miguel Silan (Université Lumière Lyon 2 & Annecy Behavioral Science Lab)

Psychology in Southeast Asia (SEA) has been widely neglected by the global scientific community despite its rich historical background, growing economy, and potential. This lack of recognition on the international stage may be attributed to implicit biases resulting from a prolonged dependence on research from WEIRD (Western, Educated, Industrialized, Rich, and Democratic) societies. Alternatively, it may be rooted in limitations within SEA’s research capacity, hindering the global dissemination of insights and knowledge from this region. Researchers in the region encounter various challenges that shape the landscape of the local research capacity, including educational emigration, political instability, bureaucratic biases, and funding inadequacies.

In this article, we aim to discuss whether educational emigration benefits or hinders the development of research capacity in SEA. First, we explicate potential factors that contribute to the educational emigration of students from SEA within the field of research in the psychological sciences, and then, secondly, we explore the impact of educational emigration on local research capacity.

WEIRD & the need for strong local research capacity

The enduring dominance of WEIRD societies within the psychological sciences raises substantial concerns regarding the impacts of the homogeneity observed in sample populations, theoretical frameworks, and researchers within the discipline. This emphasis on WEIRD societies influences the quality of internationally published research and contributes to the general neglect of other ‘non-WEIRD’ societies (Thalmayer et al., 2021). Advocates contend that fostering diversity in the scholarly landscape of the psychological sciences necessitates cultivating research capacity within these non-WEIRD societies and promoting their local talent. This strategic approach seeks to ensure that knowledge and insights from these regions are better represented on the global scientific stage.

What is the general research capacity in SEA?

Countries within SEA predominantly feature developing economies. This is primarily evident in the significant percentage of globally recognized psychologists from SEA who received their education and training outside of SEA (Leung, 2007). Given the emphasis on socio-economic development within certain countries within SEA, like Vietnam or

Cambodia, the focus on sectors like agriculture or manufacturing within these countries typically limits the focus on research in the social sciences (Saich et al., 2008). Additionally, several global economic policies in the past have unequally benefited nations outside of SEA and impeded higher education systems in SEA (Gupta, 2015). Rich and colleagues (2020) support these claims as they argue that psychology as a discipline in SEA lacks 'equal development', partly tied to socioeconomic and political factors and their 'third-world country' label. Existing limitations manifest in the region's overall research output. For instance, an investigation into psychological research in Indonesia revealed only four articles in the PsycINFO database that are authored or co-authored by Indonesian scholars (Kiling & Bunga, 2015). Additionally, Adair and colleagues (2009) explored the publication trends in 25 general psychology journals over three years and found a tendency for authors from SEA to publish in low-impact journals.

However, recent socio-economic growth within the region, characterized by increased participation in primary and secondary education, has ushered in significant transformations in the region's higher education sector (Kuroda et al., 2018). Local governments have invested substantially in higher education to address challenges arising from this growth (Arokiasamy et al., 2009). Indeed, growth in higher education within SEA is evident in recent reports, which found that 52.4% of students in higher education enrolment globally are from Asia and SEA (Lin & Pleskovic, 2008). The region's research capacity reflects such momentum within the higher education sector. One study found, for example, significant growth rates in the number of peer-reviewed international publications produced in SEA (Hien, 2010).

> **Students may choose to emigrate from SEA for their higher education, given the relatively higher socioeconomic status of WEIRD societies than societies in SEA, leading to potentially greater graduate salaries and satisfaction. Therefore, it is relatively uncertain whether educational emigration can be used as a tool for researchers from non-WEIRD societies to address the gaps in their domestic research capacity or if it is rather a byproduct of the commercialization of higher education.**

Furthermore, there is a growing trend toward international publishing and adherence to recognized standards within higher education institutions in SEA. Scholars in Indonesia, for instance, have increasingly targeted international journals for publication (Kiling & Bunga, 2015). Such increased attention toward the higher education sector within SEA has corresponded with the observed increase in the quality of educational and research infrastructures in the region (Bhandari & Blumenthal, 2011). These efforts from educational institutions, government bodies, and local talent reflect a notable internal movement toward enhancing regional research output. Trends in the rate of Asian and Southeast Asian

first authorship, single-nation authorship, and cross-cultural research reflect such efforts and trends, suggesting SEA to be much more autonomous and distinctive in recent years (Haslam & Kashima, 2010). Thus, although limitations exist, the research capacity within SEA has been significantly growing and seems to be continuously increasing alongside commitments to the higher education sector.

Given the prominence and pertinence of growth in the region's research capacity, discussions about how local entities can contribute to such progress arise. One suggested approach involves education policies supporting the educational emigration of local talent. Despite past concerns of 'brain drain', recent literature argues that educational emigration can enhance the educational investments of sending countries (Baas, 2019). However, this claim on the benefits of educational emigration hinges on the assumption that the sending country lacks the research infrastructures for adequate local researcher training. Given recent developments in research capacity and infrastructures in SEA, the following question then arises: Is a lack of proper research infrastructures still a valid motivating factor for educational emigration from SEA?

What drives educational emigration?

Many students from non-WEIRD countries pursue their higher education in WEIRD countries (Rizvi, 2005). This is evident in the substantial proportion of international undergraduate students in WEIRD societies, such as the United States, originating from non-WEIRD regions, notably Asia (Adewale et al., 2018). Indeed, it is generally supported that students from Asia tend to emigrate to countries such as the United States, United Kingdom, Australia, and countries in Europe for their higher education (Lin & Pleskovic, 2008). At the individual level, factors driving educational emigration for students in SEA include personal growth aspirations and the desire to contribute to their home countries after graduation, amongst others (Singh & Jack, 2018). Students have additionally been found to choose a host country to migrate to based on the economic factors within the host country, the reputation of its institutions, as well as the existing limitations within the institutions of their home country (William & Van Dyke, 2008). Thus, in the context of the psychological sciences, institutions in WEIRD societies, characterized by their high-income status, global university rankings, and dominance in the psychological sciences, have a comparative advantage over institutions in non-WEIRD societies. This has led to economically developed countries in North America, Europe, and Oceania being central to the global higher education system in the psychological sciences and becoming common destinations for educational migration. However, it is argued that the growing number of economically prosperous societies and high-ranking institutions within SEA allows local talents to develop their skills locally or within the region. Thus, given the growing higher education sector in SEA, the rationale for utilizing educational emigration to Eurocentric centers to develop the skills of local talent and their capacity to contribute to the local research landscape becomes less viable.

However, despite local opportunities, the commercialization of higher education may still be a motivating factor for students from SEA to emigrate to such countries. Gupta (2015) argues that there is a contemporary industrialization of higher education, emphasizing marketing

within institutions and framing educational migration as a global commercial strategy rather than a means to address research capacity gaps. This marketization encourages institutions to maximize their commercial value by emphasizing graduates' salaries and alumni satisfaction. Students may choose to emigrate from SEA for their higher education, given the relatively higher socioeconomic status of WEIRD societies than societies in SEA, leading to potentially greater graduate salaries and satisfaction. Therefore, it is relatively uncertain whether educational emigration can be used as a tool for researchers from non-WEIRD societies to address the gaps in their domestic research capacity or if it is rather a byproduct of the commercialization of higher education. The subsequent subsection will explore this topic by delineating specific limitations within SEA's research infrastructures that constrain its research capacity and by assessing whether educational emigration is a viable strategy to overcome these limitations.

Can we use educational emigration?

Research capacity is inherently shaped by socio-economic, political, and bureaucratic infrastructures (Hien, 2010). Specific factors influencing political infrastructures, such as histories of colonization or trauma (for example, Cambodia, Vietnam) and ecologies prone to disasters (for example, Philippines, Indonesia), can significantly impact the political landscape surrounding research through neglect of the educational and research infrastructures (Islam et al., 2020). For instance, Laos and Myanmar have a historically low research output tied to their prolonged histories of political revolutions and ethno-religious conflict (ibid.). Additionally, authoritarian tendencies in several SEA countries may restrict information availability, resulting in publicly accessible information deemed politicized or untrustworthy (Schomerus & Seckinelgin, 2015). Challenges escalate when academic institutions are closely tied to government bodies or are heavily reliant on government support, as in Malaysia (Ismail et al., 2019). Such dependence on unstable or unreliable political structures in the region fosters complex research landscapes with bureaucratic practices lacking rigor.

Furthermore, disjointed bureaucratic processes and inadequate oversight of aid policies can contribute to limitations in the funding landscape in SEA. For example, institutional challenges in Indonesia hinder the participation of non-profit organizations in government-sponsored research which impacts overall research funding (Carden, 2018). Moreover, a recent analysis of top journals in Psychology, Sociology, and Anthropology for Malaysia revealed biases in funding practices, which particularly neglect Chinese, Indian, and non-peninsular Malaysian populations (Silan et al., 2024). Such biases reflect potential bureaucratic biases or poor planning in Malaysian funding practices. Indeed, Azizi Ismail (2008) supports this claim as he found that Malaysian universities generally lack a common approach to funding strategy, which opens the potential for such biases to exist in funding bodies. Poor funding strategies in SEA have additionally led to heavy dependence on international funding within the region, potentially misaligning research goals with local needs (Islam et al., 2020). Such over-reliance on international financing underscores challenges in the region's political and socio-economic infrastructures surrounding research.

Addressing the limitations in research infrastructures within SEA necessitates heavy investment from local governments or support from international agencies. In the context of improving research capacity within SEA, educational migration does not arise as a universally effective solution. However, specific countries, notably less developed ones like Laos and Cambodia, participate in international enhancement programs that leverage educational migration to improve local research skills (Hong & Songan, 2011). Therefore, the value of educational migration for enhancing research capacity in SEA may be more relevant for less developed countries in the region, focused more in nurturing local talent rather than in the direct development of research infrastructure.

How to use educational emigration

For SEA to strategically leverage educational emigration to enhance its research capacity, it must encourage the return of internationally trained talents. In the absence of assurance regarding the return of such students, the transnational movement of potentially highly skilled individuals can be labeled detrimental to national development and shifts the implications of educational emigration from a focus on developing research capacity to be more aligned with the widely recognized phenomenon known as 'brain drain.' Brain drain refers to the international emigration of human capital, particularly highly educated individuals (Stankovic et al., 2013). While there has been a more recent shift towards terminologies such as 'brain gain' in the literature, emphasizing potential benefits associated with the movement of highly qualified individuals, the term 'brain drain' persists with negative connotations. However, scholars argue that if migration is not an absolute certainty and individuals eventually return to their home country, the transfer of human capital can yield a net positive impact for the sending country (Lien & Wang, 2005).

Fundamentally, countries in SEA aspiring to utilize educational emigration to improve their local research landscape must create incentives for their emigrants to return, such as job opportunities or improved quality of life. Alternatively, local governments can introduce return migration schemes that facilitate the development of local talent through international education, contingent on their commitment to return to their home country (Delgado, 2020). Various schemes in SEA, such as the Return Service Agreement (RSA) in the Philippines, necessitate graduates from specific programs to serve communities within the country for a designated period in exchange for tuition fee waivers for their education abroad. If implemented in other countries, such schemes may leverage educational emigration while specifically attending to local needs and issues.

Recommendations and conclusion

The importance of academic institutions in SEA in addressing existing gaps in their research capacity is evident. Continuing the current internal momentum toward growth within academic institutions and local talent requires incentivizing international research publications and the internationalization of higher education, one pathway of which is through educational emigration. However, beyond educational emigration, local governments should also promote more stable and reliable research environments by investing in research infrastructures, addressing inherent political constraints, and improving funding strategies. Local governments must prioritize such investments to enhance research capacity in SEA.

This imperative of local government support for improving research capacity may similarly apply to other non-WEIRD regions, for example Africa, where similar limitations in research capacity exist, and the key barriers to improving general research capacity primarily involve a lack of adequate research infrastructures, policy, or policy adherence (Chiware & Skelly, 2022).

The current essay explored educational emigration within the context of the psychological sciences in SEA. The essay illuminates a complex research landscape shaped by various factors that may influence researchers from SEA to participate in educational emigration. Principally, the essay posits that educational emigration is not a fundamental tool to be used to improve research capacity in SEA. Although certain developing countries may utilize educational emigration to nurture local talent, several challenges facing research capacity, such as political instability, bureaucratic biases, and funding inadequacies, rely on local government initiatives. This essay calls for a nuanced approach to improving the research landscape in SEA and other non-WEIRD societies, where educational emigration alone will not suffice. Local governments within these countries are urged to build upon existing internal momentum and advocate for more conducive environments for growth in research capacity. In this sense, educational emigration becomes a complementary tool within a broader strategy aimed at nurturing the general research landscape.

References

Adair, J.G., Kashima, Y., Maluf, M.R., & Pandey, J. 2009. Beyond indigenization: international dissemination of research by majority-world psychologists, *Papers from the International Association for Cross-Cultural Psychology Conferences* 2(1), 59-68. https://doi.org/10.4087/HHQF6392

Adewale, T., D'Amico, M.M., & Salas, S. 2018. 'It's kinda weird': hybrid identities in the international undergraduate community. *Journal of International Students* 8(2), 861-83. https://doi.org/10.32674/jis.v8i2.118

Arokiasamy, L., Ismail, M., Ahmad, A., & Othman, J. 2009. Background of Malaysian private institutions of higher learning and challenges faced by academics. *Journal of International Social Research* 2(8), 60-7.

Azizi Ismail, N. 2008. Information technology governance, funding and structure: a case analysis of a public university in Malaysia. *Campus-Wide Information Systems* 25(3), 145-60. https://doi.org/10.1108/10650740810886321

Baas, M. 2019. The education-migration industry: international students, migration policy and the question of skills. *International Migration* 57(3), 222-34. https://doi.org/10.1111/imig.12540

Bhandari, R., & Blumenthal, P. 2011. Global student mobility and the twenty-first-century silk road: national trends and new directions. In *International students and global mobility in higher education: National trends and new directions*. Palgrave Macmillan, pp. 1-23.

Carden, F. 2018. *Policy and causality: a learning approach*. Working Paper 28, Knowledge Sector Initiative, https://www.ksi-indonesia. org/old/in/file_upload/Policy-and-Causality-A-Learning-Approach-06Feb2018170706.pdf.

Chiware, E.R., & Skelly, L. 2022. Open science in Africa: what policymakers should

consider. *Frontiers in Research Metrics and Analytics* 7, 950139. https://doi.org/10.3389/frma.2022.950139

Delgado, L. 2020. Return migration schemes for international higher education students. Impact for sending and receiving countries. *International Journal of Humanities, Social Sciences and Education* 7(9), 1-11. https://doi.org/10.20431/2349-0381.0709009

Gupta, T. 2015. The genesis and growth of commercialization of higher education. *International Journal of Multidisciplinary Approach & Studies* 2(6), 2348-70.

Haslam, N., & Kashima, Y. 2010. The rise and rise of social psychology in Asia: a bibliometric analysis. *Asian Journal of Social Psychology* 13(3), 202-7. https://doi.org/10.1111/j.1467-839X.2010.01320.x

Hien, P.D. 2010. A comparative study of research capabilities of East Asian countries and implications for Vietnam. *Higher Education* 60(6), 615-25. https://doi.org/10.1007/s10734-010-9319-5

Hong, K.S., & Songan, P. 2011. ICT in the changing landscape of higher education in Southeast Asia. *Australasian Journal of Educational Technology* 27(8), 1276-90. https://doi.org/10.14742/ajet.893

Islam, S., Devine, J., & Alcaraz, G. 2020. *Opportunities and challenges for research translation in South and Southeast Asia*. Washington, DC, Research Technical Assistance Centre.

Ismail, S., Mustafa, N.A., & Jalil, F. 2019. Kajian undang-undang tadbir urus universiti bagi pembangunan masyarakat Malaysia. *Akademika* 89, 93-103.

Kiling, I.Y., & Bunga, B.N. 2015. Identify Indonesian psychology publications in PsycINFO Database: a scoping review. *ANIMA Indonesian Psychological Journal* 30(3), 139-47. https://doi.org/10.24123/aipj.v30i3.543

Kuroda, K., Sugimura, M., Kitamura, Y., & Asada, S. 2018. *Internationalization of higher education and student mobility in Japan and Asia*. Background paper prepared for the 2019 Global Education Monitoring Report.

Leung, K. 2007. Asian social psychology: achievements, threats, and opportunities. *Asian Journal of Social Psychology,* 10(1), 8-15. https://psycnet.apa.org/doi/10.1111/j.1467-839X.2006.00205.x

Lien, D., & Wang, Y. 2005. Brain drain or brain gain: a revisit. *Journal of Population Economics* 18, 153-63. https://doi.org/10.1007/s00148-003-0174-x

Lin, J.Y., & Pleskovic, B. (eds) 2008. *Annual World Bank conference on development economics 2008, regional: higher education and development*. World Bank Publications.

Rich, G.J., Jaafar, J.L., & Barron, D. (eds) 2020. *Psychology in Southeast Asia: sociocultural, clinical, and health perspectives*. Routledge.

Rizvi, F. 2005. *International education and the production of cosmopolitan identities*. Paper presented at the Transnational Seminar series at the University of Illinois at Urbana-Champaign.

Saich, A., Perkins, D., Dapice, D., Vu, T.T.A., Nguyen, X.T., & Huynh, T.D. 2008. *Choosing success: the lessons of East and Southeast Asia and Vietnam's future*. Harvard Vietnam Program, John F. Kennedy School of Government.

Schomerus, M., & Seckinelgin, H. 2015. *Evidence-based policymaking in Myanmar? Considerations of a post-conflict development dilemma*. JSRP Paper 25, The Asia Foundation.

Silan, M., Aouadi, M., Balla, N., El Zein, T., Haqqim, A., Lougraini, M., Madyaningrum, M., Marasigan, A., & Singdee, C. 2024. *ECoH landscaping review – Southeast Asia and the*

Arab region. Unpublished Manuscript.

Singh, J.K.N., & Jack, G. 2018. The benefits of overseas study for international postgraduate students in Malaysia. *Higher Education* 75, 607-24. https://doi.org/10.1007/s10734-017-0159-4

Stankovic, M., Angelova, B., Janeska, V., & Stankovic, B. 2013. Science and innovation policy in Southeast Europe: brain drain as brain gain. *International Journal of Technological Learning, Innovation and Development* 6(3), 262-82. https://doi.org/10.1504/IJTLID.2013.056341

Thalmayer, A.G., Toscanelli, C., & Arnett, J.J. 2021. The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist* 76(1), 116-29. https://doi.org/10.1037/amp0000622

Williams, R., & Van Dyke, N. 2008. Reputation and reality: ranking major disciplines in Australian universities. *Higher Education* 56, 1-28. https://doi.org/10.1007/s10734-007-9086-

Building local networks for open science: a case study of the Chinese Open Science Network (COSN)

Zhiqi Xu (Erasmus University Rotterdam), Yue Wang (The Chinese University of Hong Kong), Liangyou Zhang (Utrecht University), Wenqianglong Li (University of Oxford), Chenghao Zhou (New York University), Xi Chen (OPPO) and Hu Chuan-Peng (Nanjing Normal University)

Open Science embodies a global effort to democratize the processes of creating, evaluating and disseminating scientific knowledge. Central to this movement are objectives that align closely with meta-research: enhancing collaboration within the research community and bridging the gap between science and society (Science Europe, n.d.). Despite the admirable intentions of making research more accessible, transparent, and collaborative, a critical challenge arises from the predominantly developed-country origins and leadership (Jin et al., 2023). This has led to an unintended marginalization of communities in the Global South, further exacerbating the capacity gap within the international scholarly community. This chapter focuses on the *Chinese Open Science Network* (COSN), a leading grassroots group promoting Open Science in the Chinese-speaking community. COSN's objectives include engaging young researchers, spreading Open Science knowledge, bridging language gaps in science, and aiding the growth of the Open Science movement (Jin et al., 2023). COSN's role in grassroots Open Science highlights it as a potential model for empowering emerging academic communities, showing that local efforts can significantly contribute to the global movement and foster a more inclusive scientific community.

We will explore COSN's multifaceted approach to strengthening academic ecosystems through local networks, such as bridging language and cultural gaps, expanding community engagement, and enhancing national collaboration using digital infrastructures. The focus will also be on engaging early-career researchers and students in capacity building, emphasizing grassroots involvement, and utilizing digital community-building platforms. Additionally, the chapter discusses COSN's frugal operation strategy for sustainable capacity building, including volunteer workforce management and funding and cost management. This comprehensive analysis underscores COSN's pivotal role in steering scientific research toward enhanced global inclusivity and collaboration, closely aligning with international development objectives and worldwide capacity-building efforts (Kent et al., 2022).

Strengthening academic ecosystems through local networks

Since its inception in 2016, COSN has significantly advanced Open Science within the Chinese-speaking research community (Hu et al., 2016; Jin et al., 2023). The transition of COSN from a small interest group to a widely recognized network can be attributed to its focused efforts in strengthening scholarly ecosystems through localized networks. COSN's strategies have been multifaceted, including attempts to overcome language and cultural barriers, expand community engagement, foster interdisciplinary, national collaboration, and provide education and training resources. These efforts, which include translating critical resources, organizing events, and utilizing digital platforms, underscore COSN's commitment to capacity building and Open Science advocacy.

Bridging language and cultural gaps: introducing open science to the Chinese community

During the emergence of the 'replication crisis', active discussions flourished on English social media platforms (for example, X) and online social events (for example, Meet the Editor). In contrast, fewer discussions occurred on the Chinese internet due to language barriers. To raise awareness of Open Science among Chinese-speaking researchers, early members of COSN published a Chinese journal article, introducing the 'replication crisis' in psychology to the Chinese-speaking community for the first time (Hu et al., 2016). This article, followed by a consequential workshop, brought the replication crisis and the Open Science Movement to the attention of Chinese-speaking communities. This move connected Chinese academics with the international community and intensified their interest to engage with the Open Science Movement.

Recognizing the language barrier in accessing global Open Science resources, COSN's initiative involved translating English resources into Chinese, evolving into a featured column titled 'OpenTransfer.' To date, COSN has translated 38 key resources from English to Chinese, including book chapters, journal articles, and blogs. Table 1 shows the seven most-viewed translated articles. This initiative has not only made valuable resources accessible to the Chinese-speaking community but also served as a bridge between international Open Science developments and local researchers. This approach has proven instrumental for networks aiming to globalize local research communities.

RANKING	YEAR	TITLE	PAGE VIEWS
No. 1*	2018	Facial materials and toolkit collection	14801
No. 2	2023	Sample size justification: what is a reasonable sample size?	14655
No. 3	2019	Elife | Ten common statistical errors in paper writing and peer review	10683
No. 4*	2018	Collection of body images, sounds, tools, and Chinese vocabulary	8786
No. 5	2018	Essential literature resource list for psychological measurement	6415
No. 6	2018	Open data | Collection of open datasets in psychology	5921
No. 7*	2019	Meta-analysis series of R packages	5097

TABLE 1. The seven most-viewed 'OpenTransfer' articles (2024 Jan). Note: * *stands for partial translation.*

Beyond translating English materials into Chinese, COSN has actively encouraged local community members to publish tutorial papers in Chinese. Many accomplished local scholars typically publish their work, especially cutting-edge methodological studies, in English to align with the prevalent academic evaluation metrics that favor English publications. To reconnect these globalized scholars with the local community and serve the community's needs, COSN has partnered with Chinese journals to organize several special issues on complex scientific concepts like Bayesian inference (cf. Hu, 2023) and meta-analysis in Open Science (cf. Liu et al., 2021). This initiative also challenges the existing norms of scholarly evaluation, emphasizing the value of native language contributions in diversifying and enriching scientific discourse.

Expanding community engagement through resource sharing and focused events

The growing interest in open science has led to an increased need for information among Chinese scholars, such as adhering to international open science protocols and staying updated with the latest developments in methodologies. To meet these needs, COSN initiated an Official Account on WeChat, the most popular Chinese social media platform. This account serves as an outlet to disseminate Open Science principles and practices, keeping the Chinese-speaking community informed about the latest developments and progress in methodologies. Through WeChat and other digital platforms, COSN shares

free and systematic study materials, practical research knowledge and skills previously available only in English. Additionally, COSN provides timely updates on Open Science developments and research techniques. This resource-sharing initiative has attracted over 1,000 actively participating researchers and students and has expanded its reach to over 31,100 subscribers on WeChat (as of 24th May 2024).

The table below summarizes the five types of resource-sharing initiatives and events utilized by COSN, which include 'OpenTutorials,' 'OpenTalk,' 'OpenTransfer,' 'OpenMinds,' and 'OpenPlus.' Following this summary, Figure 1 illustrates the growth in popularity of these COSN resource-sharing projects and events on WeChat, demonstrating a significant expansion in reach since their initiation.

PROJECT	FUNCTION	EXAMPLE OF INTERNATIONAL COUNTERPARTS	RECORDS
OpenTransfer	Translation of resources about open science		38 resources translated
OpenMinds	Journal club dedicated to open science	ReproducibiliTea (http://reproducibilitea.org)	76 sessions
OpenTalks	Talk series on methods and skills	RIOT science club (http://riotscience.co.uk)	66 talks
OpenTutorials	Tutorials on methods and skills	Tutorials on methods and skills	18 tutorials
OpenPlus	Panel discussions on topics related to research life and careers		Five panels

TABLE 2. Summary of 'Open' series and their main functions. Note: Adapted from Jin et al., 2023, p. 8. Data updated until January 2024.

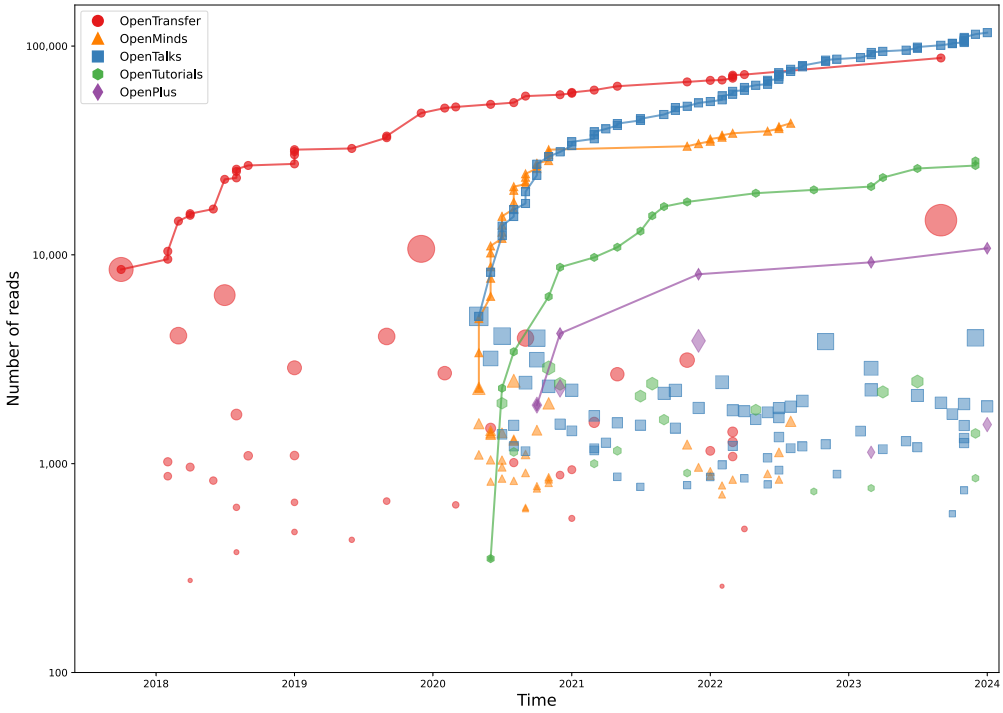

FIGURE 1. Summary of 'Open' series and their main functions

Besides sharing resources through online outlets, organizing focused events has been a cornerstone of COSN's strategy. The inaugural workshop 2016, held in Xi'an as a pre-conference workshop before the Chinese Psychological Society's annual meeting, marked the integration of Open Science into academic discourse. Subsequent workshops in 2017 and 2019 were also organized as part of the Chinese Psychological Society's annual meeting agenda. The growing participation has solidified COSN's reputation as a community builder within Chinese academia. Adapting to the COVID-19 pandemic, COSN transitioned to online events, thereby expanding its reach and inclusivity. In 2023, after the pandemic, COSN organized three additional offline workshops and co-hosted the CogSci2023 Shanghai Meetup, demonstrating its significant impact. In summary, COSN organizes workshops in both online and offline formats (see Table 3 for a summary of offline workshops), adapting to various environments and circumstances and thereby establishing close connections with local networks.

Education and training programs to empower the local community

Education and training are pivotal to promoting open science and enhancing research skills and knowledge. Recognizing this, COSN has developed a hands-on tutorial project, 'OpenTutorials,' and a talk series, 'OpenTalk', for research progress and skills sharing. The

YEAR	CONFERENCE	TOPIC
2016	Chinese Psychological Society's annual meeting	Reproducibility and open science
2017	Chinese Psychological Society's annual meeting	Reproducibility and open science
2019	Chinese Psychological Society's annual meeting	Reproducibility and open science
2023	CogSci2023 Shanghai meetup	Computational psychiatry competition
2023	COSN summer hackathon	Open science
2023	CogSci2023 Shanghai meetup	Statistics skill: Bayesian factor, Bayesian HDDM

TABLE 3. Six focused events organized by COSN between 2016-2023

array of tutorials and talks conducted by COSN covers a broad spectrum of topics within Open Science, Psychology, and Neuroscience. A notable example is the widely-read tutorial 'Overcoming the Seven Challenges in Accessing Chinese Literature in Psychological Research', which is an example in this context.

An overview of 'OpenTutorials' (see Table 4) showcases COSN's commitment to enhancing the research capabilities of Chinese-speaking communities. The program covers statistics, digital literacy, and Open Science practices. The real-world impact of these tutorials is significant, as evidenced by a COSN paper author successfully applying tutorial skills on preregistration in the Open Science Framework to their research. This case exemplifies the program's effectiveness in providing researchers with essential knowledge and tools for best practices in Open Science.

Additionally, COSN's journal club, 'OpenMinds,' serves as an interactive platform for continuous learning and discussion, particularly among undergraduate and graduate students. COSN has successfully conducted 18 tutorials, delivered 66 talks, and facilitated

YEAR	TITLE
2020	Introduction to Linux, Docker, and Git
2020	A step-by-step guide to preprocessing functional magnetic resonance imaging (fMRI) and T1 structural data using fmriprep
2020	Sharing the journey from preregistration to article publication
2020	Overcoming the seven challenges related to literature in psychological research
2020	An in-depth guide to fMRIprep (multimodal MRI data preprocessing tool)
2021	Using Rstudio in Docker
2021	Online tools for brain morphology: MicroDraw and BrainBox
2021	Introduction and FAQs on HDDM (hierarchical drift-diffusion model)
2021	Bayes factors: what they are, how to calculate them, and how to avoid pitfalls
2021	C-PAC: a flexible and user-friendly batch processing tool for resting-state functional magnetic resonance imaging (rfMRI)
2021	A guide of workflow for reproducible research
2022	Introduction to Nilearn
2022	Introduction to SPRiNT: spectral parameterization methods
2023	IDE, giving you a helping hand!
2023	Online psychological experiments: an overview and implementation
2023	Advanced compilation of ERP waveforms using the ERPLAB toolbox
2023	User guide for JASP and its applications in reproducible research
2023	CentileBrain model: standardized modeling of brain morphometry

TABLE 4. Overview of 'OpenTutorials' (2024 Jan)

76 journal club sessions as of January 2024.

COSN's provision of training resources and community-building platforms is also instrumental in establishing internationally recognized scientific standards in China. By fostering collaboration, knowledge-sharing, and adherence to best practices, COSN aligns local scientific endeavors with global benchmarks. These education and training initiatives are crucial to promoting standard adoption within the Chinese scientific landscape, thereby enhancing the credibility of scientific work in China and supporting the broader international scientific community's efforts in knowledge replication and validation.

Enhancing collaboration nationally and across disciplines: utilizing local digital infrastructures for research connectivity

In China, researchers often face communication and collaboration challenges due to a lack of centralized coordination and training. By leveraging digital infrastructure, particularly platforms like WeChat, COSN enhances interactions among researchers, librarians, funders, and publishers, fostering national collaboration across various scholarly groups. This includes organizing online events such as talks with the Eilfe editorial board to familiarize Chinese scholars with publication processes and disseminating calls for collaboration to domestic and international labs and databases.

The dissemination of collaboration information, especially internationally, has encouraged more Chinese researchers to engage in collaborative endeavors within the grassroots Big-Team Science Movement (Forscher et al., 2023). For example, COSN has introduced several impactful projects, notably The Psychological Science Accelerator, alongside innovative platforms like StudySwap and RRR. These initiatives exemplify COSN's commitment to advancing collaborative research efforts, fostering knowledge exchange, and promoting reproducibility and reliability in scientific studies. By facilitating the exchange of knowledge and fostering a spirit of cooperation, COSN contributes to building robust networks and strengthening the involvement of Chinese researchers in large-scale scientific collaborations.

Furthermore, COSN extends its initiatives beyond psychology and cognitive neuroscience, emphasizing interdisciplinary collaboration and the promotion of local digital infrastructures such as ChinaXiv and the Science Data Bank. The promotion of these platforms is pivotal in encouraging Chinese-speaking scholars to embrace Open Science practices such as preprints and data archiving, thereby increasing the visibility of local digital resources. Furthermore, these initiatives highlight the importance of interdisciplinary work, merging diverse fields to foster a more integrated scholarly community. This strategy, which promotes interdisciplinary dialogue and the use of local open-access digital resources, aligns with global trends that recognize the essential role of digital platforms in expanding open access to research materials. Such efforts contribute significantly to enriching the scholarly ecosystem, echoing the international movement towards more accessible and collaborative research environments (Suber, 2012).

Engaging early-career researchers and students in capacity-building

COSN has developed a robust strategy to engage Early-Career Researchers (ECRs) in Open Science, recognizing their crucial role in building scholarly capacity. Through the innovative use of digital platforms, a grassroots emphasis, and empowering initiatives, COSN and ECRs have formed a mutually beneficial relationship. While COSN is nurturing the next generation of Open Science practitioners, ECRs contribute significantly to COSN's growth.

Emphasizing grassroots and diverse involvement

COSN strongly emphasizes grassroots involvement, focusing on undergraduate students and ECRs who may lack formal research training support (Restivo, 2005). By targeting ECRs and student groups, COSN aims to sow the seeds of Open Science and foster community-wide awareness and adoption in a bottom-up manner. This strategy is pivotal for cultivating a new generation of researchers equipped with Open Science principles from the outset of their careers. For example, most of COSN's WeChat account readers fall within the 18-35 age range, which aligns well with the demographic of the targeted groups COSN aims to support (see Figure 2). Furthermore, there is a higher concentration of regular readers of COSN's contents from underdeveloped regions in China (see Figure 3). These distributions may indicate that COSN's efforts effectively reach and assist the intended audience.

Using local and global digital platforms to unite ECRs across communities

COSN has successfully accessed the ECR group and generated efficient collaborations through WeChat. Initially, COSN used its official WeChat account as an information outlet

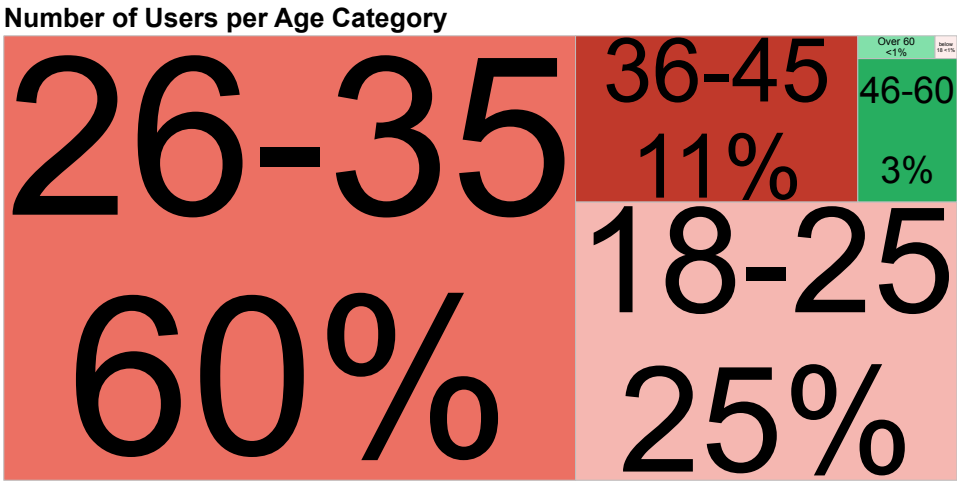

FIGURE 2. Number of readers per age category (WeChat Data, Jan 2024)

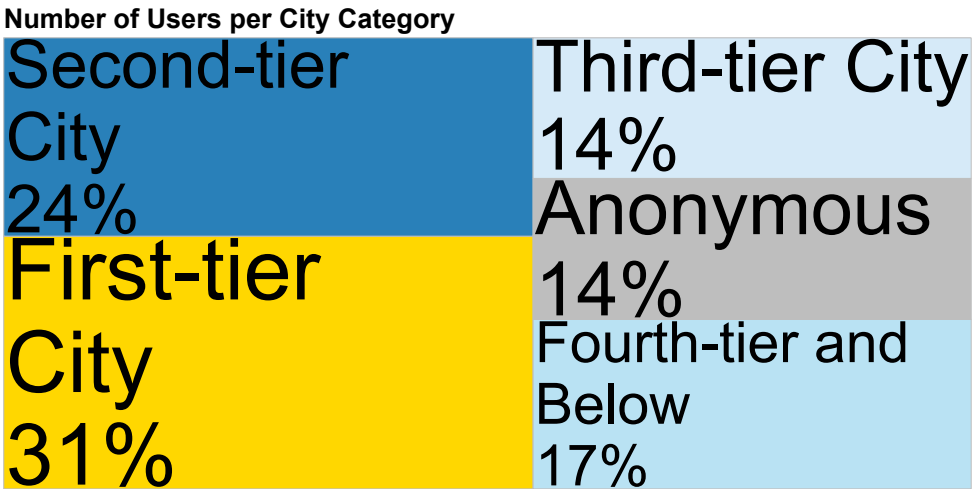

FIGURE 3. Number of readers per location (WeChat data, Jan 2024)

to disseminate Open Science principles and practices and to serve interested audiences. Subsequently, COSN managed to attract many ECRs and students among the audience to join its volunteer team, who have contributed immensely to COSN's outputs and daily operations. Furthermore, COSN has created multiple WeChat chat groups for different interest groups. Instant communications through such channels have accelerated team formation and collaboration among ECRs and between ECRs and student groups.

In addition, COSN disseminates its content in Chinese and English via other Chinese and international social media and video-streaming platforms to expand its reach. Online event planning also intentionally accommodates participants across various time zones. Table 5 (Jin et al., 2023) below summarizes these efforts.

This strategic decision to leverage WeChat's extensive user base and other digital platforms has reached a wider audience and fostered a community around Open Science across different scholarly groups. Chinese-speaking ECRs and students are connected regardless of nationality, place of residence, career level, field of study, and sociodemographic factors.

PLATFORM	FUNCTION	LINK OR ACCOUNT
Website	Main portal	https://open-sci.cn
WeChat	Primary Chinese social media platform	Account ID: OpenScience
X	Portal for international engagement	Handle: @ChineseOpenSci
Bilibili	Outlet for sharing event recordings	https://space.bilibili.com/252509184
Github	Code storage and sharing	https://github.com/OpenSci-CN

TABLE 5. Platforms used by COSN. Note: Adapted from Jin et al., 2023, p. 7. Data updated until January 2024.

Moreover, COSN's presence on multiple platforms embeds these ECRs and students within the international community.

Providing ECRs with a platform for research exchange and career growth

COSN's 'OpenTalk' is a speaking platform that invites diverse speakers to share their research and engage in discussions, with a particular emphasis on promoting the work of ECRs. This initiative allows ECRs to self-nominate and participate actively, including those with recently published impactful work. As of January 2024, ECRs have delivered 47 of 66 talks in this series, enhancing their professional development and visibility in the scientific community. Additionally, COSN brings in senior researchers from various backgrounds to facilitate connections between ECRs and established academics.

COSN's 'OpenPlus' panel series serves as a dynamic forum for ECRs to discuss various academic and research-related topics, including career decisions between academia and industry (see Table 6). It provides a platform for sharing experiences and gaining insights from industry leaders and researchers, assisting academic youth in making informed career choices. The events, often coinciding with major holidays, foster a festive and communal atmosphere. For instance, during the Laba Festival on January 18, 2024, COSN organized an OpenPlus event focused on the experiences of newly independent Principal Investigators (PIs) from Chinese institutes and abroad. These events facilitate collaboration between academia and industry and strengthen COSN's relationships with researchers.

Providing personal value for contributors

COSN's management, primarily composed of ECRs dedicated to Open Science, plays a crucial role in organizing events tailored to the needs of their Chinese-speaking peers.

YEAR	PANEL TITLE
2020	Halloween special: trick or treat – From academia to industry
2020	Year-end special: to be or not to be – Navigating the choices of 2020
2021	Christmas special: merry industry
2023	COSN tea party – Reflecting, preserving, and moving forward
2024	Laba tea party – Talking about the growth of academic youth as winter transitions into spring

TABLE 6. Overview of 'OpenPlus' (2024 Jan)

This leadership structure, profoundly empathetic and experienced, is key to effectively engaging ECRs in Open Science practices. Participation in COSN's activities offers significant benefits for ECRs. For instance, Hu Chuan-Peng, through his involvement in COSN, published several academic papers, contributed to Big-Team Science, and wrote Chinese tutorials, enhancing his career prospects. Similarly, Haiyang Jin, the organizer of OpenMind, and other key members have experience in professional development. Students also find growth opportunities. Yuanrui Zheng, for example, joined COSN as an editor of its WeChat Official Account after engaging with the tutorials. His involvement in COSN, starting as an undergraduate from an underdeveloped region, led to contributions to academic publications and a successful application for a master's project at Vrije University of Amsterdam, the Netherlands.

COSN has established a model for engaging and empowering ECRs in open science through these strategies. Their approach demonstrates how targeted initiatives can significantly enhance the capacity of ECRs, preparing them to be future leaders in the open science movement.

Frugal operation for sustainable capacity building

COSN's resource acquisition approach, particularly in managing human resources and costs and securing funding, is a key factor in sustaining its initiatives and operations over time.

Achieving personnel continuity through a frugal volunteer workforce

COSN is a loosely structured academic organization with a predominantly student and ECR volunteer workforce. These volunteers engage in various projects without the need to manage daily tasks, typically contributing around one hour per week. In contrast, project leaders oversee ongoing initiatives with an average of three hours weekly. The ' affordable

‘ principle is central to COSN’s operations, ensuring tasks are manageable and within the team’s capacity.

Faced with the challenge of personnel changes as members progress in their careers, COSN employs strategies for stability and effectiveness. They streamline operations by using templates for WeChat posts and other activities, such as speaker invitations, and by limiting monthly committee meetings to thirty minutes to reduce time demands on members.

COSN also undertakes systematic recruitment and comprehensive training of new members to maintain continuity and adapt to evolving team dynamics. This process ensures the retention of essential skills and knowledge, allowing for a swift response to changes within the team. Additionally, COSN is committed to fostering a culture of development, enhancing the organization’s capacity to promote its volunteers’ growth. This commitment to personnel development equips team members to address Open Science challenges effectively, contributing to COSN’s long-term goals and aligning with international standards in organizational management and capacity building in the scientific community.

Strategic funding acquisition and cost management through collaborations

Securing funding for Open Science is a significant challenge, especially with limited official support (see UNESCO, 2023). While developed countries have recently begun supporting Open Science projects, like open-source software and neuroinformatic databases (cf. Gau et al., 2021; Halchenko et al., 2021), developing countries still struggle to access such funding equitably. COSN has, therefore, diversified its funding sources, exploring government grants, private foundations, corporate sponsorships, and crowdfunding. This approach was effective in 2021 when the non-profit institute Tianqiao and Chrissy Chen Institute for Brain Disease (Shanghai) sponsored COSN’s WeChat Official Account, helping cover operational costs such as a Zoom subscription. Additionally, COSN strengthens its financial position through collaborations, pooling resources, and expertise with like-minded organizations. A notable success was a $2500 grant secured at the CogSci2023 Shanghai Meetup.

Alongside seeking diverse funding sources, COSN implements effective cost-reduction strategies. For instance, they have outsourced event logistics and venue provision to the institute co-organizer, Anhui Medical University, for the 2023 Hackathon. Moreover, contrary to the norm, most invited speakers do not receive honorariums. This voluntary contribution from the scholarly community significantly reduces costs for COSN, exemplifying the spirit of collaboration and mutual support within the Open Science field.

COSN’s efforts are ongoing regarding grant applications and proposal development. While a recent application to the Einstein Foundation did not yield the desired result, it significantly enhanced COSN’s expertise in crafting grant proposals. With an eye on the future, COSN aims to leverage its growing recognition in the Open Science community, incorporating new members with relevant experience to strengthen its funding acquisition strategies.

In navigating both personnel and funding challenges, COSN exemplifies frugal operation through its grassroots spirit. Core members volunteer for additional responsibilities in their spare time, optimizing resources and fostering a cost-effective, community-driven approach to achieving Open Science goals. This dedication inspires collective participation, leading to impactful work, and epitomizes COSN's core values.

Conclusion

In reflecting upon COSN's journey, we share four insights that may benefit scholars across the Global South interested in fostering Open Science practices (cf. Jin et al., 2023).

1. The cornerstone of effective engagement lies in identifying and addressing the unique needs of the scholarly community. This is vital for empowering ECRs and students who often grapple with resource constraints.
2. The emphasis on practical applications is a magnet for new members, showcasing Open Science's tangible advantages and fostering a vibrant, topic-centered community.
3. Forging connections with a spectrum of partners – from local grassroots initiatives to domestic and international organizations has magnified our initiative's impact and, consequently, enriched its resource base.
4. Maintaining affordability for both participants and contributors is vital. Adopting a frugal operational model, leveraging crowdsourcing, volunteer support, and cost-free digital tools, has ensured our team's sustainability.

Beyond these individual strategies, their synergistic integration,illustrated by COSN's fusion of resource-sharing with community-building,has significantly broadened our audience and enhanced service delivery. However, the diversity of academic communities underscores that there is no universal solution. Nevertheless, the core principle of serving community needs, informed by our experiences, offers a roadmap. Initiatives can begin on a small scale informally, as COSN did, and evolve to strengthen the global push towards a research ecosystem that is both more inclusive and equitable.

References

Forscher, P., Wagenmakers, E-J., Coles, N.A., Silan, M.A., Dutra, N., Basnight-Brown, D.M., & IJzerman, H. 2023. The benefits, barriers, and risks of big team science. *Perspectives on Psychological Science* 18(3), 607-623. https://doi.org/10.1177/17456916221082970

Gau, R., Noble, S., Heuer, K., Bottenhorn, K.L., Bilgin, I.P., Yang, Y.F., ... & Marinazzo, D. 2021. Brainhack: developing a culture of open, inclusive, community-driven neuroscience. *Neuron* 109(11), 1769-1775. https://doi.org/10.1016/j.neuron.2021.04.001

Halchenko, Y.O., Meyer, K., Poldrack, B., Solanky, D.S., Wagner, A.S., Gors, J., ... & Hanke, M. 2021. DataLad: distributed system for joint management of code, data, and their relationship. *Journal of Open Source Software* 6(63), 3262. https://doi.org/10.21105/joss.03262

Hu, C.-P., Wang, F., Guo, J., Song, M., Sui, J., & Peng, K. 2016. The replication crisis in psychological research. *Advances in Psychological Science* 24(9), 1504-1508. https://doi.org/10.3724/SP.J.1042.2016.01504

Hu, C.-P. 2023. Introduction to the special issue on 'Bayesian statistics and inference'. *Journal of Psychological Techniques and Applications* 11(9), 513.

Jin, H., Wang, Q., Yang, Y.-F., Zhang, H., Gao, M., Jin, S., ... & Chuan-Peng, H. 2023. The Chinese Open Science Network (COSN): building an Open Science community from scratch. *Advances in Methods and Practices in Psychological Science* 6(1), 1-17. https://doi.org/10.1177/25152459221144986

Kent, B.A., Holman, C., Amoako, E., Antonietti, A., Azam, J.M., Ballhausen, H., ... Weissgerber, T.L. 2022. Recommendations for empowering early career researchers to improve research culture and practice. *PLOS Biology* 20(7), e3001680. https://doi.org/10.1371/journal.pbio.3001680

Liu, Y., Chen, S., ... & Hu, C.-P. 2021. A standardized checklist on meta-analysis reporting in the Open Science era. *Scientia Sinica Vitae* 51(6), 764-78. https://doi.org/10.1360/SSV-2021-0009

Restivo, S.P. 2005. *Science, technology, and society: an encyclopedia*. Oxford University Press.

Science Europe. n.d. *Open science*. Retrieved from https://www.scienceeurope.org/our-priorities/open-science/

Suber, P. 2012. *Open access. MIT Press,* https://doi.org/10.7551/mitpress/9286.001.0001

UNESCO. 2023. *Open science outlook 1: status and trends around the world*. UNESCO. Retrieved from https://doi.org/10.54677/GIIC6829

Bridging the majority and minority worlds: liminal researchers as catalysts for inclusive open and big-team science

Alma Jeftic (University of Copenhagen & International Christian University, Tokyo), Marc Yancy Lucas (Universidad de Sonora, Hermosillo), Nadia Corral-Frías (Universidad de Sonora, Hermosillo) and Flavio Azevedo (University of Utrecht)

In the contemporary landscape of psychological sciences, the imperative to foster inclusive, socially just, and globally collaborative research endeavors has never been more pressing. Despite the potential for diversity and equity, systemic barriers and biases persist in open and big-team science projects, often relegating researchers from low- and middle-income countries (LMIC) to peripheral roles within international collaborations. This situation underscores a critical issue within psychological sciences: the inadequate representation of diverse participants and researchers. Such a shortfall not only exacerbates the replication crisis but also suggests that the crisis extends beyond mere methodological rigor. It points to a deeper systemic problem rooted in the insufficient diversity and inclusivity of research practices and theoretical perspectives. This multidimensional challenge highlights the necessity for an approach to scientific inquiry that integrates a broader range of cultural and experiential backgrounds into the fabric of psychological research. This paper proposes a path forward, stressing the critical role of *liminal scientists* who can bridge the gaps between the Global North and South, as well as between diverse academic structures.

Decoding the scientific jargon: unpacking the roles of liminal scientists in open and big-team science

We frequently use terms in this paper that are specific to scientific and academic discussions. To ensure that all readers, regardless of their background, can fully engage with our findings and arguments, we explain some key terms. We use the term 'liminal scientist' to describe individuals bridging the gaps between researchers, institutions, and academic interests in the Global North and those in the Global South. The concept draws on Victor Turner's (2017) notion of liminality, or the deliberate uncertainty some social actors are subjected to when crossing thresholds between culture, context, and place. Turner further articulates a social dichotomy that contrasts structures, representing top-down hierarchies of power and control, with anti-structures of creative bottom-up responses. The greater the experienced powerlessness, the more pressing the need for anti-structural processes of positive community activity (St John, 2008; Bigger, 2009). We perceive connections between this conceptual framework and the challenges faced by research in the majority world (also called the Global South) in the context of open and big-team science.

Liminal scientists have experienced academia from different perspectives and can translate that expertise into fostering open relationships by transferring knowledge, information, resources, and collaborative efforts between currently divided worlds. Individuals capable of creatively navigating and softening barriers between disparate academic realms but also able to challenge existing structures and alleviate the sense of isolation experienced by those navigating these divides stand as pioneering bridge-builders in academia, shaping the future of inclusive and global collaboration. Liminal researchers hold a unique position to foster diversity, inclusivity, and equity within the open and big-team science communities. They accomplish this by championing equitable contributions, ensuring the fair allocation of resources, advocating for proper academic recognition, and highlighting the critical importance of incorporating cultural, linguistic, and contextual insights into research. This inclusion is vital for producing valid and relevant research across different global contexts and propelling global scientific progress. Such a shift promises to improve equity in research processes and practices and to ensure that open and big-team science projects reflect the vast tapestry of human experience.

The more specific context of our arguments focuses on open and big-team science, even if we believe that some of our arguments might also apply to other contexts. Open science is an 'umbrella term reflecting the idea that scientific knowledge of all kinds, where appropriate, should be openly accessible, transparent, rigorous, reproducible, replicable, accumulative, and inclusive, all of which are considered fundamental features of the scientific endeavor' (Parsons et al., 2022). Open science is a scientific social movement aimed at making scientific research, data, and dissemination accessible to everyone, from citizen scientists to academics (Nosek & Bar-Anan, 2012). The major aspects of open science include open data which makes research data publicly available, open methodology which shares the methods used in research to enhance reproducibility, open source which provides free access to software and tools used in research, open access which allows research findings to be freely accessible to all, open peer review which fosters transparency in the review process, and open educational resources to foster learning and education (Munafò et al., 2017; Nosek et al., 2012, 2015). Together, these practices have the potential to help democratize scientific knowledge, thereby making science more inclusive, collaborative, and in line with social justice principles (Azevedo et al., 2019; Crüwell et al., 2019; Kathawalla et al., 2021).

Big-team science – large-scale collaborations of scientists working on a scholarly common goal and pooling resources across labs, institutions, disciplines, cultures, and countries – is more and more often considered a key part of open science (see, for example, Forscher et al., 2023, Korbmacher et al., 2023). Through coordinated efforts and by leveraging the strengths of numerous researchers from various disciplines and regions, big-team science projects can address complex scientific challenges that single researchers or groups might not be able to tackle adequately. This team-science method significantly enhances resource allocation, diversifies expertise, and improves the reproducibility of scientific findings, crucially aligning with the foundational goals of open science (Coles et al., 2023; Forscher et

al., 2023). Big-team science broadens the scope and scalability of investigations. It offers a crucial opportunity to progress the most challenging aims of open scholarship – specifically, enhancing diversity, equity, inclusion, accessibility, and social justice within research (Korbmacher et al., 2023). If big-team science fulfills its potential, it can render scientific inquiries more democratic, socially just, and globally inclusive.

Our paper also contrasts research structure, capacity, and goals between low- and middle-income countries (LMICs) and high-income countries (HICs). LMICs are defined by the World Bank as nations with lower per capita gross national income than HICs, which possess more substantial resources and infrastructure to support scientific research and development. This classification helps highlight the diverse economic realities and resource availabilities between these groups of countries. Additionally, we advocate for the incorporation of principles of DEIA – Diversity, Equity, Inclusion, and Accessibility – into open and big-team science. DEIA represents a set of social justice principles dedicated to ensuring fair treatment, opportunity, and advancement for all. In scientific research, embracing DEIA means identifying and eliminating barriers that inhibit full participation and creating environments where diverse researchers and their perspectives are valued and all individuals can thrive and contribute to their fullest potential.

> **Together, our stories highlight the disparities within academia and the potential for liminal researchers to act as bridges between divergent worlds, enriching psychological science with diverse insights, experiences, and methodologies.**

Reflexivity statement

We have included a reflexivity statement to acknowledge our diverse backgrounds as liminal researchers, underscoring how our intersections of privilege and disprivilege shape our perspectives and motivations in addressing the disparities within global scientific research. In writing this paper, we, as a diverse collective of scholars from Bosnia, Mexico, Brazil, and the United States of America, find ourselves at the intersection of various dimensions of privilege and disprivilege. Our academic journeys, marked by experiences in both HICs and LMICs, position us as liminal researchers navigating the complexities of global academia. Our insights are informed by the academic wheel of privilege (Elsherif et al., 2022; Middleton et al., 2024), which serves as a critical lens through which we examine our positions within the academic structure, highlighting the nuances of our privileges and the challenges we face.

Our unique experiences reflect a broad spectrum of societal, economic, and educational backgrounds, illustrating the multifaceted nature of academic privilege and disprivilege. For instance, Alma Jeftic's academic path, which traversed from the war-torn landscapes of the Balkans to prestigious institutions in Europe and Asia, embodies the resilience and

adaptability required to overcome substantial barriers. Her experiences of studying and researching in environments marked by conflict and scarcity and her subsequent navigation of global academic networks shed light on the unique perspectives and strengths that liminal researchers bring to the scientific community. Similarly, Flavio Azevedo's journey from experiencing poverty and violence in Brazil to becoming a neurodivergent early-career scholar in Germany, the UK, and the Netherlands underscores the challenges and triumphs associated with navigating academia from a position of multiple marginalities. Adding to this mosaic of experiences, Yancy Lucas, born and educated in the United States with European ancestry, navigates the academic landscape as an early career, non-tenured professor. Despite the privilege of his native language and background, Yancy confronts the complexities of teaching and research in a language and context that are not his own, reflecting the nuanced challenges of academic integration and identity. Nadia Corral-Frías, a tenured professor and a woman of color educated in a HIC, experiences her own unique set of advantages and hurdles. Working in Mexico, Nadia faces obstacles related to funding, access to research materials, and balancing caretaking duties with her academic responsibilities. Yet, her proficiency in English has enabled her to participate actively in the predominantly English-speaking academic world.

Together, our stories highlight the disparities within academia and the potential for liminal researchers to act as bridges between divergent worlds, enriching psychological science with diverse insights, experiences, and methodologies. As liminal researchers, we grapple with the dualities of our identities, navigating both the privileges and disprivileges of our social and academic parkours. This duality informs our understanding of the academic landscape and motivates our commitment to bridging the divide between the Global North and South. Throughout this paper, we aim to contribute to a broader dialogue on integrating diverse cultural, linguistic, and contextual insights into scientific inquiry.

Our understanding of these complexities and our positions as liminal researchers have directly influenced our involvement in initiatives such as the consortium *Advancing Big-team Reproducible Science through Increased Representation*, ABRIR (https://abrirpsy.org/), and *The Framework for Open and Reproducible Research Training*, FORRT (https://forrt.org/). FORRT is an interdisciplinary community of scholars from all career stages that aim to integrate open science principles into higher education as well as to advance research transparency, reproducibility, rigor, and ethics through pedagogical reform and meta-science (Azevedo et al., 2019). ABRIR, the *Advancing Big-Team Reproducible Science through Increased Representation* consortium, aims to identify challenges uniquely faced by underrepresented researchers in psychological sciences and to devise best practices that can be shared within the global academic community. Working on building these communities is central to our commitment to bridging the divides between HICs and LMICs. It embodies our efforts to translate the principles of open and inclusive science into actionable strategies that address the unique challenges faced by underrepresented researchers. By founding and participating in these organizations, we actively work to ensure that research practices

become more inclusive, directly addressing the disparities experienced throughout our academic journeys. Through ABRIR and FORRT, we advocate for a more equitable research environment and implement the necessary changes to make it a reality. These organizations are critical to our strategy of fostering genuine diversity, equity, inclusivity, and accessibility in open and big-team science, providing platforms that empower and afford opportunities for underrepresented researchers to contribute on an equal footing.

After discussing the barriers faced by researchers from LMICs, we discuss how to strengthen research capacity for LMIC researchers using ABRIR and FORRT as practical examples of how liminal researchers can actively shape the research landscape.

The current state of affairs in open and big-team science

As investigators with experience in HICs and LMICs, as well as open and big-team science networks, we have identified the critical importance of expanding psychological sciences to include researchers worldwide. Global North dominance stifles psychological science by hindering the creation of generalizable theories (IJzerman et al., 2021). There is also a moral imperative to research psychological phenomena that disproportionately affect LMICs (Adetula et al., 2022) as shown by, for instance, research addressing the effects of economic disadvantage on mental health and well-being in children from the majority world (Morrow et al., 2019) or research illustrating how religious and spiritual coping mechanisms may be impacted by resource struggles during crises like the COVID-19 pandemic in LMIC (Captari et al., 2022). Enhancing inclusivity is not only a matter of equity but also a strategic imperative. The understanding of diverse psychological processes cannot be solely derived from the perspectives prevalent in HICs (Adetula et al., 2022). Recognizing the urgent need to expand inclusivity within psychological science, we believe that the key to this endeavor is the role of liminal scientists in bridging the gap between LMICs and HICs, fostering a more comprehensive and inclusive approach to research (Forscher et al., 2023). Despite substantial evidence that most psychological science has been, and is currently, developed, tested, published, and applied in predominantly White and North American samples and HIC contexts (Arnett, 2008; Ghai, 2021), diversity has not greatly improved in over fifteen years (Thalmayer et al., 2021). The causes of this lack of change are multifaceted, and a comprehensive discussion does not fall within the scope of this chapter (see Coles et al., 2023). However, it is important to point out that structures have changed little, and linguistic hegemonic practices, such as reliance on papers published in English, remain the norm. Editorial boards continue to demonstrate limited diversity, which may lead to less acceptance of papers from LMIC and reduced citations (Liu et al., 2023). Taken together, these can be understood as a result of a widespread and systematic disinterest that hinders inclusion and status quo changes.

The challenge of including research and researchers from LMICs extends into the realms of open and big-team science networks. For instance, the *Psychological Science Accelerator* (PSA), a globally distributed network of researchers notable for its significant contributions

to fostering large-scale, worldwide research collaborations, has disclosed in its capacity reports a revealing snapshot of the current landscape. In both 2020 and 2023, researchers from HICs comprised nearly two-thirds of all PSA researchers, accounting for 65.88% and 65.25%, respectively. Not surprisingly, most of the PSA papers have been led by authors from the US and Western Europe (constituting 93.75% from 2020 to 2021 and 84.61% from 2022 to 2023). This discrepancy underscores a critical issue: inclusion does not necessarily equate to visibility, leadership, or equitable distribution of resources or prestige. Many researchers from LMICs find their contributions marginalized, rarely serving as either the first or last authors (Hedt-Gauthier et al., 2019; Korbmacher et al., 2023). Without a genuine and substantive effort to integrate LMIC researchers (and their perspectives) within open and big-team science, such imbalances are poised to persist and deepen. These risks entrench inequities in authorship, credit allocation, access to resources, and global academic recognition.

This disparity in representation transcends the boundaries of the PSA and similar open and large-scale scientific endeavors, suggesting deeper systemic and structural inequalities within the scientific community. Despite targeted efforts to bolster engagement within these networks, such as establishing PSA regional hubs, there has been no significant increase in participation from LMIC researchers (Corral-Frías et al., 2024). This stagnation may stem from a mutual lack of understanding between these networks and LMIC scientists concerning each other's scientific methodologies, cultural priorities, and infrastructural realities.

Liminal scientists, with their unique positioning and ability to navigate between diverse scientific and cultural contexts, are perhaps more equipped to serve as intermediaries, facilitating communication and understanding across divides. However, the challenges of bridging such varied environments with their distinct organizational cultures and expectations cannot be understated. In the following sections, we outline a strategy for enhancing the field of psychological sciences by harnessing the unique perspectives and capabilities of liminal scientists who, as research indicates (Kato, 2015; Chu et al., 2023), often surpass expectations in forging collaborative partnerships.

Moving forward

Throughout our academic journey, we attempted to serve communities as liminal scientists and, as such, often performed research while lacking resources and opportunities, demanding heightened creativity to navigate scarce financial, structural, and human resources. We learned that concerns pivotal to researchers in LMIC, such as authentic representation of samples, researchers, and research topics, often take a backseat to trends favored by the academic mainstream in HICs. The quest for genuine diversity, inclusion, equity, and accessibility in open and big-team science has elicited mixed reactions: enthusiasm from some, apathy from others, and even hostility from a few. These experiences mirror broader struggles for social justice, both within and outside academia. Encountering indifference

and hostility early in our open and big-team science endeavors deepened our resolve rather than diminishing our spirits. Motivated by these challenges, we have joined forces with consortia and groups aligned with our goals, such as ABRIR and FORRT. Collaborating with like-minded people within these communities has boosted our morale and strengthened our determination to effect meaningful change in psychological sciences. Working for and with these partnerships has fostered a sense of belonging outside traditional academic boundaries, highlighting that, for example, empathy and support from well-meaning individuals of privileged backgrounds can indeed amplify our effectiveness in bridging diverse worlds, ultimately enriching our collective endeavor toward a more inclusive and interconnected scientific community.

In this context, open and big-team science initiatives would greatly benefit from the active involvement of liminal researchers to effectively navigate the complex academic, administrative, and political landscapes encountered in engagements with LMICs. Familiarity with the cultural, linguistic, and historical background of an area or region, alongside knowledge of funding distribution mechanisms, drives a more efficient information transfer between institutions and individuals. By establishing contextually specific networks of research collaborators and working groups, these insights enable targeted and effective engagements, enhancing the collaboration process within open and big-team science initiatives.

Lived experiences from liminal researchers

In open and big-team science, liminal researchers frequently need clarification, particularly when HIC networks overlook or undervalue their lived research experiences. These experiences often encompass context-specific political and ethical issues, notably in areas such as trauma, conflict-related research, and interethnic relations. These are critically important yet sometimes dismissed as mere logistical challenges. Liminal researchers often confront skepticism about the efforts and resources they invest in accessing hard-to-reach populations, with their contributions frequently relegated to mere data collection. This undervaluation fosters a cycle of pessimism among LMIC researchers about the prospect of collaborating with HIC counterparts. Such collaborations are approached with increasing wariness, largely due to a history of inequitable credit distribution and resource allocation in these projects. In open and big-team science projects, there is a noticeable trend where scientists from more developed regions often lead the narrative, inadvertently relegating researchers from LMICs to roles that support the primary objectives without contributing to the intellectual core of the project. In such settings, liminal researchers frequently find themselves positioned as mere providers of resources or data, serving the interests of the project without due recognition or involvement in decision-making.

This dynamic can be interpreted as extractive and echoes colonialist practices, where the value and contributions of LMIC researchers are overshadowed by the priorities and interests of their counterparts from more privileged backgrounds. It highlights the urgent need for

a shift towards more equitable and inclusive research practices that recognize and utilize the full scope of expertise and perspectives that LMIC and liminal researchers bring to the scientific table. Hoekstra et al. (2018) highlight the versatility of liminal researchers, from customizing research tools to embody local relevance to infusing the research process with local perspectives during data interpretation. By effectively conveying the specific needs and insights of LMIC communities to broader research collaborations, liminal researchers pave the way for more equitable and reciprocal scientific exchanges. Such contributions foster greater engagement and investment from LMIC researchers and enrich the research outcomes with diverse perspectives and deeper contextual understanding.

Advancing equitable global collaborations

There are important processes to consider while establishing and managing global collaborations. Power dynamics should be considered in advance. Singh (2022) argues that some HIC researchers believe that the mere inclusion of LMIC researchers is enough; thus, authorship order or active involvement in projects is considered unimportant. We have found that in many cases, collaborative procedures are unclear or differ from region to region, limiting the involvement of researchers from LIMCs to data collection and middle authorship. Projects are, as shown by some recent multinational studies (see, for example, Buchanan, 2023; Cologna et al., 2024; Dorison et al., 2022; Psychological Science Accelerator Self-Determination Theory Collaboration, 2021; Vlasceanu et al., 2024; and Wang et al., 2021), initially often conceptualized by researchers in HICs, who also control the funding, allocate resources, and determine the scientific process as well as the publication strategy. As a result, the research design, including research questions, data collection methodologies, and analytical techniques, is often predominantly developed according to the interests, standards, and practices used in HICs.

Unsurprisingly, this arrangement not only hinders the contributions of LMIC researchers but also inadvertently strengthens the hierarchical division between HIC and LMIC collaborators, thus perpetuating and potentially exacerbating the global disparities in academic recognition and influence. To address these inequities, liminal researchers are well positioned to illustrate how procedural disparities in collaboration – from conceptualization to publication – constrain the engagement of LMIC researchers, resulting in limited opportunities for significant authorial and scholarly recognition. Liminal researchers advocate for establishing models that ensure equal contributions, highlighting the essential roles and importance of LMIC researchers and clarifying research collaboration conventions. Current research paradigms, primarily shaped by the academic systems and interests of HICs, often fail to consider the specific needs and challenges inherent to LMIC contexts. To achieve genuinely equitable collaboration, these standards must evolve to recognize and address the realities LMIC researchers face. This approach will help foster a more inclusive and diverse scientific community where all contributions are valued and influential. For open and big-team science collaborations to be genuinely equitable, HIC standards must evolve to acknowledge and address the realities of LMIC contexts.

Liminal researchers play a crucial role as intermediaries in this process, especially in grant writing and preparation, where LMIC researchers may not be familiar with the expectations and requirements set by external funding bodies. By leveraging their unique understanding of both HIC and LMIC environments, liminal researchers can offer valuable guidance on navigating the funding application landscape, ensuring that projects are more accessible to LMIC participants. Furthermore, funding institutions are encouraged to utilize the insights of liminal researchers to facilitate connections with LMIC scholars, enhancing the inclusivity and relevance of research initiatives. Success in these efforts requires a concerted push towards more equitable practices within HIC academic and research institutions, ensuring they listen to and actively incorporate LMIC researchers' recommendations into their operations.

Embracing these changes is imperative for building a research ecosystem that values diverse perspectives and shares the benefits and spoils of scientific inquiry equitably across global communities. The toolset of liminal researchers enriches open and big-team projects with deep cultural, linguistic, and historical knowledge, contextualizing research and building essential bridges. By leveraging these insights, open and big-team science networks can overcome barriers to collaboration, ensuring that research projects not only reach but resonate with diverse communities. This can be achieved with the recommendations above, always recognizing and crediting liminal researchers for their essential contributions in fostering a more inclusive and equitable research environment, enabling truly global scientific endeavors.

Strengthening research capacity for LMIC researchers

Strengthening research capacity refers to enhancing the ability of LMIC researchers to participate, conduct, manage, and disseminate high-quality scientific research in open and big-team science alongside HIC researchers. We argue this is a crucial step to address systemic disparities in global scientific output and ensure that LMIC researchers can contribute equitably to the international scientific community, highlighting diverse perspectives that enrich and broaden the scope of scientific inquiry. Hence, strengthening research capacity in LMICs is not just a matter of resource allocation but a fundamental shift towards equitable, culturally sensitive collaboration practices. Nchinda (2002) and Malekzadeh et al. (2020) underline the necessity of investing in local researchers and infrastructure to build sustainable research ecosystems in LMICs. Bowsher et al. (2019) add that such investments must be accompanied by sustainable funding, equitable partnerships, and mentorships to effectively bridge the gap between research and practice. The success of initiatives like the *GINGER* program in neuropsychiatric genetics (gingerprogram.org) but also *Open Life Sciences* (openlifesci.org), *The Turing Way* (the-turing-way.netlify.app), *Marginalia* (marginaliascience.com), the *Society for the Improvement of Psychological Science* (improvingpsych.org), FORRT (forrt.org), and other grassroots organizations exemplifies how targeted capacity building, rooted in the cultural and contextual realities of LMICs, can significantly advance global and inclusive research (Steltenpohl et al., 2021). However,

the diversification and inclusivity efforts in psychological sciences face systemic barriers, including historical biases, existing power structures, and a reluctance to fully embrace cultural complexities (see Syed, 2023). This resistance underscores a broader issue within the scientific community: a persistent undervaluation of LMIC contributions and a lack of genuine collaboration with HIC researchers. For open and big-team science to truly be inclusive, it must transcend mere tokenistic involvement of LMIC researchers, recognizing the indispensable value of culturally and contextually informed perspectives in enriching the scientific discourse and making meaningful advancements. Therefore, the path forward involves addressing logistical and financial challenges and dismantling the systemic biases that hinder the integration of diverse perspectives. Acknowledging and actively incorporating the expertise of LMIC researchers as equal partners in the scientific process is essential for developing a truly global, inclusive, and effective research community.

We propose bridge building as a possible first step to a solution where liminal scientists can act as bridge builders by translating knowledge to and from LMICs and HICs (see Corral-Frías et al., 2023). We envision a series of interconnected strategies aimed at fortifying DEIA principles in open and big-team science projects, including establishing transparent, fair, and equitable guidelines for collaboration, implementing mentorship programs, fostering bidirectional learning between LMIC and HIC researchers, enhancing funding access for LMIC researchers, cultivating inclusive research communities and networks, promoting local (often applied) research priorities, monitoring and reporting on equity measures, and advocating for policy changes toward greater adherence to DEIA principles. Central to our approach is the development and wide dissemination of guidelines that clearly articulate the principles of equitable collaboration. These explicit guidelines aim to address inclusivity by delineating fair practices regarding authorship, decision-making, and resource distribution, ensuring that researchers from LMICs are not just participants but integral contributors to the research processes. Similarly, to support the growth and visibility of LMIC researchers, especially early career scholars, a vital component is the establishment of mentorship programs. By pairing emerging talents from LMICs with experienced scholars across the spectrum of HICs and LMICs, open and big-team science projects can cultivate a generation of researchers who are skilled at both scientific pursuits and adept at navigating the complexities of the global research environments. These mentorship relationships can also boost opportunities for LMIC researchers in critical areas of such projects (leadership) and pass on strategies for successful publication, grant writing, and overall professional development.

Another recommendation is the promotion of bidirectional learning and cultural exchange via the organization of workshops, symposiums, and exchange programs to foster the democratization of knowledge, methodologies, and perspectives. This collaborative educational and research model has the potential to enhance the quality of the research output and deepen mutual understanding and respect within the scientific community. Equally important is the need to ensure LMIC researchers have more and better access to funding opportunities, which entails advocating for funding mechanisms that are

specifically designed to support projects led by or significantly involving LMIC researchers, enabling them to pursue research agendas that are directly relevant to their communities. By nurturing inclusive research networks, we aim to facilitate more accessible access to resources, foster robust international collaborations, and amplify the voices of those from underrepresented regions. By implementing these strategies, we not only leverage the unique insights and capabilities of liminal researchers but also pave the way for a more inclusive, equitable, and effective global research community.

> **Our standards for excellence in research must not only prioritize methodological rigor but also embrace inclusivity and diversity, ensuring that psychological science serves and reflects the entire global community.**

Indeed, we have attempted to move some of these initiatives forward through ABRIR and FORRT. Both academic groups were born out of a need for social justice in psychology academia. They seek to improve it through the democratization of resources, scientific outputs, and educational resources and pedagogies. FORRT and ABRIR were founded to empower scholars, especially those early in their careers and from LMIC, with either curated learning and teaching materials on the topic of open science or by creating greater access through talks, seminars, materials, and activities in various languages. Both organizations advocate for more integration of social justice principles in the practice of open science in both research and teaching and create opportunities aimed at progressively leveling the educational and research landscape. ABRIR and FORRT are partially governed by scientists who have experience in both LMIC and HIC. As liminal scientists and through our grassroots organizations, we can help those in LMIC with networking, access to literature, educational and teaching resources, insights about grant applications, and a community. ABRIR, for example, with less than three years since its founding, has already made great strides. ABRIR organized an international conference entitled 'Increased representation: a vision for inclusive big-team science.' Through this conference, academic resources (videos and written materials) were created with the aid of experts from all over the world (for example, the USA, France, Nigeria, India, Brazil, and the Netherlands) and were shared in different languages. These resources are available online on YouTube and our website (https://abrirpsy.org/events/increased-representation-2022). The FORRT community provides learning, educational and pedagogical resources (Azevedo et al., 2021; Parsons et al., 2022; Pownall et al., 2021, 2023) as well as guidance and advocacy towards greater adherence to DEIA principles (Azevedo et al., 2022; Elsherif et al., 2022; Gourdon-Kanhukamwe et al., 2023; Middleton et al., 2024).

Continued efforts by both organizations are underway as we develop and translate open and big-team science resources across LMIC's languages and offer free and open statistics and R tutorials. All these activities were planned based on the needs of the psychological

researchers from the LMIC with the aim of increasing their representation. In line with the objective of liminal scientists, through regional and global meetings we have plans to provide an even more detailed set of recommendations along with an analysis of culture-specific factors that influence research flow in LMIC and differences from HIC standards. These ideas, recommendations, and initiatives represent just the beginning of a much-needed transformation in the open and big-team science culture – a shift from viewing LMIC researchers merely as resources for advancing HIC projects to recognizing them as essential partners with valuable insights and contributions. Achieving this requires a fundamental change in the prevailing academic power dynamics and reevaluating what constitutes 'good science.' Our standards for excellence in research must not only prioritize methodological rigor but also embrace inclusivity and diversity, ensuring that psychological science serves and reflects the entire global community. This evolution towards a more equitable, diverse, accessible, and inclusive approach to science demands a collective commitment to change, challenging us to dismantle existing barriers and build a research environment where every voice is heard, valued, and integrated into the fabric of scientific discovery.

Conclusion

This paper highlighted the pivotal role of liminal researchers in bridging the representational gap within open and big-team science networks. Liminal researchers hold the key to fostering more inclusive and equitable scientific collaborations by acting as bridge builders, communicators, and facilitators. We delved into both the unique advantages they bring and the challenges they face, underscoring their critical function in connecting diverse academic worlds. Liminal researchers possess the unique ability to navigate and mediate between the varied linguistic, administrative, and cultural landscapes that often separate researchers in HICs from their counterparts in LMICs. Their role, crucial for initiating meaningful dialogue and partnerships, merits further exploration and amplification in the ongoing quest to democratize global scientific endeavors. Emphasizing the necessity to leverage their talents and insights, we advocate for a concerted effort to integrate and elevate the contributions of liminal scientists in future collaborations across the Global North and South, ultimately advancing the cause of truly global and inclusive open and big-team science.

References

Adetula, A., Forscher, P., Basnight-Brown, D.M., Azouaghe, S., & IJzerman, H. 2022. Psychology should generalize from – not just to – Africa. *Nature Reviews Psychology* 1(7), 370-1. https://doi.org/10.1038/s44159-022-00070-y

Arnett J.J. 2008. The neglected 95%: why American psychology needs to become less American. *American Psychologist* 63(7), 602-14. https://doi.org/10.1037/0003-066X.63.7.602

Azevedo, F., Liu, M., Pennington, C.R., Pownall, M., Evans, T.R., Parsons, S., ... & FORRT. 2021. Towards a culture of open scholarship: the role of pedagogical communities. *BMC Research Notes* 15(1), 1-5. https://doi.org/10.1186/s13104-022-05944-1

Azevedo, F., Middleton, S., Phan, J.M., Kapp, S., Gourdon-Kanhukamwe, A., Iley, B., ... & Shaw, J.J. 2022. Navigating academia as neurodivergent researchers: promoting neurodiversity within open scholarship. *APS Observer* 35.

Azevedo, F., Parsons, S., Micheli, L., Strand, J., Rinke, E., ... & FORRT. 2019. *Introducing a Framework for Open and Reproducible Research Training (FORRT)*. https://doi.org/10.31219/osf.io/bnh7p

Bigger, S. 2009. Victor Turner, liminality, and cultural performance. *Journal of Beliefs & Values* 30(2), 209-12. https://doi.org/10.1080/13617670903175238

Bowsher, G., Papamichail, A., El Achi, N., Ekzayez, A., Roberts, B., Sullivan, R., & Patel, P. 2019. A narrative review of health research capacity strengthening in low and middle-income countries: Lessons for conflict-affected areas. *Globalization and Health* 15(1). https://doi.org/10.1186/s12992-019-0465-y

Buchanan, E.M., Lewis, S.C., Paris, B., Forscher, P.S., Pavlacic, J.M., Beshears, J.E., ... & Vega, D. 2023. The Psychological Science Accelerator's COVID-19 rapid-response dataset. *Scientific Data* 10(1), 87. https://doi.org/10.1038/s41597-022-01811-7

Captari, L.E., Cowden, R.G., Sandage, S.J., Davis, E.B., Bechara, A.O., Joynt, S., & Counted, V. 2022. Religious/spiritual struggles and depression during COVID-19 pandemic lockdowns in the Global South: evidence of moderation by positive religious coping and hope. *Psychology of Religion and Spirituality* 14(3), 325-37. http://dx.doi.org/10.1037/rel0000474

Chu, J.Y., Chiu, S.Y., & Syu, Y.S. 2024. Returnee and local doctorates on research productivity: doctoral training, international mobility, and research collaboration. *Higher Education* 87, 963–89. https://doi.org/10.1007/s10734-023-01046-4

Coles, N.A., DeBruine, L.M., Azevedo, F., Baumgartner, H.A., & Frank, M.C. 2023. 'Big team' science challenges us to reconsider authorship. *Nature Human Behaviour* 7(5), 665-7. https://doi.org/10.1038/s41562-023-01572-2

Cologna, V., Mede, N.G., Berger, S., Besley, J.C., Brick, C., Joubert, M., ... & van der Linden, S. 2024. Trust in scientists and their role in society across 67 countries. *OSF Preprints*. January 25. https://doi.org/10.31219/osf.io/6ay7s.

Corral-Frías, N.S., Burin, D.I., Garduño-Realivazquez, K.A., García Estrada, L.I. & Jeftic, A. 2024. *Access to Open Science in Latin America behavioral science: successes and challenges of PSA Latin America Hub and ABRIR*. Talk presented at the Year of Open Science Culminating Conference, 21-22 March 2024. Retrieved from https://osf.io/nsgfj/

Corral-Frías, N.S., Castillo, E.L., Lucas, M.Y., Armenta, M.F., Rodriguez, Y.V., Dutra, N., ... & Azevedo, F. 2023. Latin American psychological science: will the Global North make room? *APS Observer* 36.

Crüwell, S., van Doorn, J., Etz, A., Makel, M.C., Moshontz, H., Niebaum, J.C., ... & Schulte-Mecklenbeck, M. 2019. Seven easy steps to open science. *Zeitschrift für Psychologie* 227(4), 237-48. https://doi.org/10.1027/2151-2604/a000387

Dorison, C.A., Lerner, J.S., Heller, B.H., Rothman, A. J., Kawachi, I.I., Wang, K., ... & Pantazi, M. 2022. In COVID-19 health messaging, loss framing increases anxiety with little-to-no concomitant benefits: experimental evidence from 84 countries. *Affective Science* 3(3), 577-602. https://doi.org/10.1007/s42761-022-00128-3

Elsherif, M.M., Middleton, S.L., Phan, J.M., Azevedo, F., Iley, B.J., Grose-Hodge, M., ... Dokovova, M. 2022. Bridging neurodiversity and open scholarship: how shared values

can guide best practices for research integrity, social justice, and principled education. https://doi.org/10.31222/osf.io/k7a9p

Forscher, P.S., Wagenmakers, E.J., Coles, N.A., Silan, M.A., Dutra, N., Basnight-Brown, D., & IJzerman, H. 2023. The benefits, barriers, and risks of big-team science. *Perspectives on Psychological Science* 18(3), 607-23. https://doi.org/10.1177/17456916221082970

Ghai, S. 2021. It's time to reimagine sample diversity and retire the WEIRD dichotomy. *Nature Human Behaviour* 5(8), 971-2. https://doi.org/10.1038/s41562-021-01175-9

Gourdon-Kanhukamwe, A., Kalandadze, T., Yeung, S., Azevedo, F., Iley, B.J., Phan, J.M., ... Elsherif, M.M. 2023. Opening up understanding of neurodiversity: a call for applying participatory and open scholarship practices. *The Cognitive Psychology Bulletin* 8, 23-7. https://doi.org/10.53841/bpscog.2023.1.8.23

Hedt-Gauthier, B.L., Jeufack, H.M., Neufeld, N.H., Alem, A., Sauer, S., Odhiambo, J., ... & Volmink, J. 2019. Stuck in the middle: a systematic review of authorship in collaborative health research in Africa, 2014-6. *BMJ Global Health* 4(5), e001853. https://doi.org/10.1136/bmjgh-2019-001853

Hoekstra, R.A., Girma, F., Tekola, B., & Yenus, Z. 2018. Nothing about us without us: the importance of local collaboration and engagement in the global study of autism. *BJPsych International* 15(2), 40-3. https://doi.org/10.1192/bji.2017.26

IJzerman, H., Dutra, N., Silan, M., Adetula, A., Basnight-Brown, D., & Forscher, P. 2021. Psychological science needs the entire globe, part 1. *Association for Psychological Science - APS*. https://www.psychologicalscience.org/observer/global-psych-science

Kathawalla, U.-K., Silverstein, P., & Syed, M. 2021. Easing into open science: a guide for graduate students and their advisors. *Collabra: Psychology* 7, 16684. https://doi.org/10.1525/collabra.18684

Kato, M. 2015. Continuity of international collaborations with the country of origin: the case of highly cited migrant researchers from developing countries. *Journal of International Development Studies* 24(2), 81-96. https://doi.org/10.32204/jids.24.2_81

Klein, O., Hardwicke, T.E., Aust, F., Breuer, J., Danielsson, H., Mohr, A.H., ... & Frank, M.C. 2018. A practical guide for transparency in psychological science. *Collabra: Psychology* 4(1), 20. https://doi.org/10.1525/collabra.158

Kijilian, C., Sharifian, M., Paris, B., Iyer, A., & Silan, M.A. 2022. *PSA study capacity report 2022-2023*. https://doi.org/10.31234/osf.io/7krhn

Korbmacher, M., Azevedo, F., Pennington, C.R., Hartmann, H., Pownall, M., Schmidt, K., ... & Evans, T. 2023. The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology* 1(1). https://doi.org/10.1038/s44271-023-00003-2

Liu, F., Rahwan, T., & AlShebli, B. 2023. Non-White scientists appear on fewer editorial boards, spend more time under review, and receive fewer citations. *Proceedings of the National Academy of Sciences* 120(13), e2215324120. https://doi.org/10.1073/pnas.2215324120

Malekzadeh, A., Michels, K., Wolfman, C., Anand, N., & Sturke, R. 2020. Strengthening research capacity in LMICs to address the global NCD burden. *Global Health Action* 13(1), 1846904. https://doi.org/10.1080/16549716.2020.1846904

Middleton, S., Iley, B., Sulik, J., Elsherif, M.E., & Azevedo, F. 2024. *The academic wheel of privilege: an equity-based tool for authorship order.* www.forrt.org/publications

Morrow, V., & Boyden, J. 2019. A view from the global south: commentary on the special

section. *Developmental Psychology* 55(3), 601. https://psycnet.apa.org/doi/10.1037/dev0000679

Munafò, M., Nosek, B.A., Bishop, D.V.M., Button, K.S., Chambers, C.D., Percie du Sert, N., ... & Ioannidis, J.P.A. 2017. A manifesto for reproducible science. *Nature Human Behaviour* 1, 0021. https://doi.org/10.1038/s41562-016-0021

Nchinda, T.C. (2002). Research capacity strengthening in the South. *Social Science & Medicine,* 54(11), 1699–711. https://doi.org/10.1016/s0277-9536(01)00338-0

Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S., Breckler, S., ... & DeHaven, A. 2015. Promoting an open research culture. *Science* 348(6242), 1422-5. https://doi.org/10.1126/science.aab2374

Nosek, B.A., & Bar-Anan, Y. 2012. Scientific utopia: I. Opening scientific communication. *Psychological Inquiry* 23(3), 217-43. https://doi.org/10.1080/1047840X.2012.692215

Nosek, B.A., Spies, J.R., & Motyl, M. 2012. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science* 7(6), 615-31. https://doi.org/10.1177/1745691612459058

Parsons, S., Azevedo, F., Elsherif, M.M., Guay, S., Shahim, O.N., Govaart, G.H., ... Aczel, B. 2022. A community-sourced glossary of open scholarship terms. *Nature Human Behaviour* 6(3), 312-8. https://doi.org/10.1038/s41562-021-01269-4.

Pownall, M., Azevedo, F., König, L.M., Slack, H.R., Evans, T.R., Flack, Z., ... & FORRT. 2023. Teaching open and reproducible scholarship: a critical review of the evidence base for current pedagogical methods and their outcomes. *Royal Society Open Science* 10(5), Article 221255. https://doi.org/10.1098/rsos.221255

Pownall, M., Azevedo, F., Aldoh, A., Elsherif, M.M., Vasilev, M.R., Pennington, C.R., ... Parsons, S. 2021. Embedding open and reproducible science into teaching: a bank of lesson plans and resources. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000307

Psychological Science Accelerator Self-Determination Theory Collaboration. 2021. A global experiment on motivating social distancing during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* 119(22). https://doi.org/10.1073/pnas.2111091119

Singh, L. 2022. Navigating equity and justice in international collaborations. *Nature Reviews Psychology* 1(7), 372-3. https://doi.org/10.1038/s44159-022-00077-5

Steltenpohl, C.N., Montilla Doble, L.J., Basnight-Brown, D.M., Dutra, N.B., Belaus, A., Kung, C. C., ... & Darda, K. 2021. Society for the Improvement of Psychological Science global engagement task force report. *Collabra: Psychology* 7(1), 22968. https://doi.org/10.31234/osf.io/4upqd

St John, G. (ed) 2008. *Victor Turner and contemporary cultural performance*. Berghahn Books.

Syed, M. 2023. The slow progress towards diversification in psychological research. *PsyArXiv*. https://doi.org/10.31234/osf.io/bqzs5

Thalmayer, A.G., Toscanelli, C., & Arnett, J.J. 2021. The neglected 95% revisited: is American psychology becoming less American? *American Psychologist* 76(1), 116-29. https://doi.org/10.1037/amp0000006

Turner, V.W. 2017. Liminality and communitas. In *Ritual*. Routledge, pp. 169-87

UKCDR. 2023. *Essence: seven principles for strengthening research capacity in low- and middle-income countries*. https://www.ukcdr.org.uk/resource/essence-seven-principles-for-strengthening-research-capacity-in-low-and-middle-income-countries-2014/

Van Bavel, J.J., Cichocka, A., Capraro, V., Sjåstad, H., Nezlek, J. B., Pavlović, T., ... & Jørgensen, F.J. 2022. National identity predicts public health support during a global pandemic. *Nature Communications* 13(1), 517. https://doi.org/10.1038/s41467-021-27668-9

Vlasceanu, M., Doell, K.C., Bak-Coleman, J.B., Todorova, B., Berkebile-Weinberg, M.M., Grayson, S.J., ... & Lutz, A.E. 2024. Addressing climate change with behavioral science: A global intervention tournament in 63 countries. *Science Advances* 10(6), eadj5778. https://doi.org/10.1126/sciadv.adj5778

Wang, K., Goldenberg, A., Dorison, C., Miller, J., Lerner, J., & Gross, J. 2021. A multi-country test of brief reappraisal interventions on emotions during the COVID-19 pandemic. *Nature Human Behaviour* 5, 1089-110.https://doi.org/10.1038/s41562-021-01173-x

Contextually grounded research in postgraduate research training in Africa: why and how

Aurelia Munene (Eider Africa)

Stepping into postgraduate research training in Kenya

The process of founding Eider Africa, an African-led research mentorship organization in Kenya in 2016, has been thought-provoking, healing, occasionally perplexing, and inspirational. I stepped into research mentorship and training with caution, imposter syndrome, and a strong sense of belief, even as a female researcher in the early stages of my career as a master's student. The main issue that motivated me to engage in this area was the numerous challenges post-graduate students face while working on their research projects. The post-graduate students I spoke with in meetings and informal conversations from 2015 onwards narrated their experiences of research project delays, with some spending eight years to complete a two-year master's project due to multiple factors that will be discussed later in this paper. Some had supervisors who were either absent or unresponsive. Others could not apply what they had learned in their research methods courses to developing their research projects.

Some post-graduate students could not find an appropriate non-predatory journal to publish their work, while others could not access journal articles and important scholarly materials to further their research projects because most of these journals were behind the publisher's paywall. Being self-sponsored students, most faced funding shortages that further delayed their studies. Their experiences are corroborated by various studies extensively documenting similar challenges (Mukhwana et al., 2016; Mbogo et al., 2020; Mugendi & Githae, 2021; Motseke, 2020; Desmennu & Owoaje, 2018).

As I began to think more deeply about how to contribute to responsive postgraduate research training, I started to explore and access the situated accounts of different stakeholders in the university and the research training ecosystem actors in general. Two main questions guided me: i) How have postgraduate research training challenges been explained and resolved previously in Kenya and Africa? ii) How can we understand these problems better to identify the root drivers and foster a way forward?

Research on postgraduate research training: gaps and opportunities

Several studies conducted on postgraduate education and research training in Kenya focused on what I would call symptoms of the problem. For example, a large study conducted on the

status of postgraduate research and training in Kenya, in its conceptualizing the problem and the recommendations, located postgraduate challenges within the existing institutional arrangements without questioning how these institutions and systems are created and maintained (Mukhwana et al., 2016). For example, the study proposes more policies, robust codes of conduct, supervision, and improved management strategies, which is essential. However, as 'systems of established and embedded social rules/ norms that structure social interactions' (Hodgson, 2006), institutions and the social rules and norms embedded in them structure interactions and actions of universities, post-graduate students, faculty, and the market as well as the way these interact with the research ecosystem more generally. Examining these norms is important, especially in understanding where they come from, their purposes, and what happens when they interact with the contextual realities and knowledge of the people and spaces to which they are applied.

Becoming increasingly aware of the limitations in these studies, which mainly arise from the fact that these studies do not adequately examine how ideology and the politics of knowledge inform the current postgraduate research training and the research landscape, this paper will discuss dominant ideologies that are persistent in our education system and make the linkage between the current gaps in postgraduate training and the politics of knowledge. As such, the roots of these struggles are deeper and linked to the existing and ongoing call from notable African scholars and others on the centrality of embracing norms that foster contextually grounded research in Africa. This means being conscious and willing to remedy, as a researcher, the historical and persistent knowledge production inequalities in the continent, such as the marginalization and invisibility of African scholars' contributions to research, including the misrepresentation of African ways of knowing and communities in research (Ntaragwi & Okwany, 2020). It also means not doing research for the sake of it but to collaboratively develop research goals with communities so that research becomes meaningful to those it intends to serve. This type of research also means adapting more inclusive methodologies, research approaches, theories, knowledge, questions and methods of data collection, and analytical categories that resonate with the diversities of the communities we intend to work with (Okwany & Hasina, 2015). To enable this, we need to meaningfully conceptualize, design, and analyze research projects with communities from the beginning to the end, even at the postgraduate level. This means that funding research norms and systems may need to become more flexible and open to funding these diverse methodologies.

Based on the preceding, I argue in this paper that the challenges of postgraduate studies and research training in Kenya can, in part, be traced to the complex interactions between a troubling colonial history of education in Africa and neoliberal ideologies imported from outside Africa into the higher education sector and how they intersect with the diverse and dynamic knowledge and social-cultural context in the continent. In the context of postgraduate research training, this means shifting from the current teaching approaches that are based on what Paul Freire in the *Pedagogy of the Oppressed* calls the 'banking concept'2

(2005: 72), which turns students into bankers of knowledge and rewards the same rather than evoking their critical consciousness and their transformation. The banking concept approach alienates students and trainers from their world and reality, serving the interest of what Freire calls the 'oppressors' (ibid.: 73-4). This is illustrated by how some researchers treat communities as objects of constant scrutiny through research rather than social actors. I then showcase how, considering these complexities, the organization I founded, Eider Africa, continually examines these challenges and offers expanded spaces for African scholars to reflect and move towards articulating contextually grounded research.

If asked to describe my approach at Eider Africa and my orientations as a researcher, I would quote Maya Angelou's *Our Grandmother's Poem* and say, 'I go forth along and stand as ten thousand.' I draw inspiration, resilience, support, and strength from the collective heritage of African voices in scholarship and communities. In addition, my scholarly foundations have been influenced by knowledge activism, decolonial work, and the intellectual output of great African scholars and scholars from other contexts who call for African voices to be at the center of telling African stories. I was a laureate of the Children and Youth Institute at the *Council for the Development of Social Science Research in Africa* (CODESRIA) 2015. This is one of the central institutions in Africa that was set up to promote African-led scholarship. While there, I was exposed to a wealth of contributions from African scholars in social research that I did not know existed. This is the reality for most postgraduate students who cannot access this work quickly and are exposed to Western scholarship in their training such that they unintentionally marginalize African contributions, which explains the glaring absence of African scholars' contributions in their thesis' bibliographies (Ntaragwi & Okwany, 2020).

Locating postgraduate research training challenges

To understand the challenges that affect postgraduate research training, we need to highlight the foundations of education and research in Kenya and Africa in general. A troubling colonial history has influenced scholarship in Africa, and I agree with prominent African scholar Oddora Hoppers, who states that it is necessary to 'acknowledge that the education systems inherited from the colonial period must be challenged and transformed, which means redefining the goals, content, structures, methods, approaches, and values of education, as part of a mold-breaking strategy' (2000). As a result of this colonial history, the development of education systems in Africa has been founded on an erasure of African contextualized knowledge and systems of knowing in favor of western-based knowledges. This is evidenced in the dominance of Western language in research, publication spaces, ideology, theories, research agenda, and the elevation of Western scholars over others in respect. This erasure has translated into distorted representations of African communities in research, where instead of being framed as speaking subjects and active participants in producing and using their knowledge, Africans have become mere objects of study (Okwany & Hasina, 2015). I, for instance, currently work in the space of development consultancy. I witness problematically how the research agenda, research theories, questions, methods,

and dissemination mechanisms of specific development projects like water, climate, finance, education, and health are set from the donors' worldview (often western-based). We, as consultants, have to work within these frameworks with sometimes little wiggle room. Although we push for more inclusive research approaches, the dominant narrative is still enduring, and we keep going, nonetheless.

Therefore, I argue that meaningful deliberations should be made to disentangle the complexities that challenge postgraduate research training in Kenya and the continent. We need to constantly reflect on and address historical legacies. This can be done by examining the norms it has produced and how they shape the systems developed for postgraduate research training. These systems have interacted with changing economic, political, socio-cultural, and technological systems in Africa today and have created enduring challenges and opportunities in the higher education space. They have influenced how education, research, and their purpose, is conceptualized in institutions of higher learning and impacted how the goals of postgraduate research training are understood, thereby helping to create unequal systems of knowledge production that have been insensitive and marginalizing to the local context.

> **We want to build a critical mass of African scholars embracing contextually grounded research practices. We recognize that for this to happen, it must go beyond the individual and extend our work to build research movements of African storytellers (researchers) together with our diverse communities.**

For example, neoliberalism is one of the most dominant economic models interacting with our colonial history and shaping the purpose and value of education (Okune & Ulrich, 2021; Nyamnjoh, 2019; Njoya, 2019). The neoliberal ideology promotes reduced government intervention in terms of regulation reduction in government spending, resulting in limited social safety nets in favor of the privatization of the education sector that is left to the mercy of market forces. The effects of a market-driven higher education have been the following: introduction of tuition fees, which has driven unequal access to higher education; increased privatization of universities; reduced government spending on research and public institutions; a market-driven curriculum, which has led to a push for education for employment and the market. (Kigotho, 2018) A growing pool of part-time academic staff who are involved in consultancies to supplement their income limits their engagement in research and research training (Nyamnjoh, 2019).

I have seen how these ideologies shape postgraduate students' aspirations in conversations with them. For example, I would ask them why they are doing this master's or PhD; the majority would answer that they chose it because it is marketable, while others were looking forward to a job promotion after completion. When I asked what led them to choose this or that research topic, the responses were varied. While some were deeply passionate about

the issue they were studying, others had chosen a topic that suited their current time and financial capacities or one that fulfilled the graduation criteria. All these responses are valid; however, I agree with the African scholar Wandia Njoya, who calls for the delinking of university education with employment and states that '[D]'egrees are academic qualifications for an education that improves the human experience through a raised consciousness, broad knowledge, and mastery of skill' (Njoya, 2019). I believe education has a broader, self-transformational, relational, and collective goal than just to create employment or learn skills. This view is particularly important in a continent that has faced erasure and silencing of its knowledge and people's experiences through research.

Initiatives to actualize contextually grounded research in Africa – The example of Eider Africa

Eider Africa's role is to create collaborative spaces and processes that center the voice of African postgraduate scholars and revive the African research ecosystem by complementing the role of universities. We aim to ensure that African scholars are responsible for producing and using knowledge on the continent. We envision nurturing African researchers or professionals (we recognize not all postgraduate students have to be researchers) who transform and become competent, critical, reflexive, collaborative, and ethical. We want to build a critical mass of African scholars embracing contextually grounded research practices. We recognize that for this to happen, it must go beyond the individual and extend our work to build research movements of African storytellers (researchers) together with our diverse communities. I highlight some of the work we are doing with like-minded partners at Eider Africa to build these research movements and advance contextually grounded research below:

> *Establishing communities of postgraduate students*. We are leveraging accessible technology like WhatsApp to grow communities of African scholars through what we call the *Africana Journal Club*[1], a multi-disciplinary, multi-university online space we have created to bring together over six hundred and growing African postgraduate students and accompany them throughout their research journeys for as long as they want to remain in the group. Doctoral students are often isolated from their universities and each other; they work in silos and interact with students from their discipline and university only. To build research movements, we must unite African scholars and promote non-hierarchical communities that enable scholars to support each other. For example, we have co-designed activities in the club that members participate in at no cost, and we motivate them to peer review each other's work and learn from and celebrate each other. Our pool of mentors is mainly made up of African researchers and lecturers. Our activities are facilitated through what I would call *African Philanthropy*, whereby the scholars and mentors

1 Eider Africa Africana Journal Clubs, https://eiderafricaltd.org/programmes/africana-journal- or club/.

voluntarily share their knowledge, resources, and support. And I, as the founder, provide both labor and financial resources to cover the costs of our activities.

Transforming the learning and teaching of research in our universities. We recognize that the approaches and methodologies currently used to teach in our universities cannot promote practices that embrace contextually grounded research because they essentially promote the above-mentioned banking approach to teaching and learning. They do not provide an environment where postgraduate students can develop their voices and worldviews as African researchers and learn diverse methodologies to comprehensively and responsibly tell African stories. Therefore, Eider Africa joined the *Association for Faculty Enrichment in Learning and Teaching* (AFELT), an organization trying to improve learning and teaching in universities by embracing transformative pedagogy[2]. Together, we are developing a transformative research learning and teaching course for lecturers. This work is essential because it is part of building the research movement on contextually grounded research among African scholars.

Contributing to the Research Data Share, a scholarly-owned and -run digital platform that attempts to explore, through practice, the challenges and opportunities of sharing qualitative data. The *Research Data Share* platform draws together Kenya-based researchers to reflect on our experiences and practices and to create an archive of data and meta-data that grows the knowledge commons in Kenya[3]. The aim of this collective is to advance more open, inclusive, responsible research practices and non-commercial infrastructures that embrace diverse collaborative ways of knowing, examining data, and increasing research transparency by posting our meta-data publicly for further debate and analysis.

Training on open peer review in Africa in the context of scholarly publication and pre-prints. Open peer review is critical for African scholars because opaque peer review processes have hindered and prevented many postgraduate students from getting published. The course invites participants to reflect on their biases as they conduct peer review; it exposes them to different ways the peer review processes can be open, promotes collective peer review practices, and provides technical guides on conducting a peer review. Eider Africa was part of a partnership that included *PREreview*, *elife*, *Training Center in Communication-Africa*, and *AfricArXiv*. This partnership aimed to deliver training for African Scholars and develop open-source resources for anyone who wants to conduct this training. After the training, Eider Africa has been conducting training with university-based journal editors and lecturers on open peer review. The most important outcome of these trainings is

2 Association of Faculty Enrichment in Learning and Teaching, https://www.afelt.org/home/.
3 Research Data Share platform, https://www.researchdatashare.org/

lecturers reflecting on how they provide postgraduate students feedback on their thesis, including manuscripts, and how they can do better at giving clear, actionable, and constructive feedback.

In conclusion, as the movement for meta-research is gaining momentum, contextually grounded research practices and approaches could be woven in as a mechanism to re-distribute power in research economies, particularly in Africa, to reposition communities and African scholars from the margins to the center so that they are enabled to lead and tell African stories.

References

Desmennu, A.T., & Owoaje, E.T. 2018. Challenges of research conduct among postgraduate research students in an African university. *Educational Research and Reviews* 13(9), 336-42. https://doi.org/10.5897/ERR2017.3359

Freire, P. 2005. *Pedagogy of the oppressed*. Continuum.

Hodgson, G.M. 2006. What are institutions? *Journal of Economic Issues* 40(1), 1-25. https://doi.org/10.1080/00213624.2006.11506879

Hoppers, C.O. 2000. Turning the monster on its head: lifelong learning societies for all. In M. Jain & S. Jain (eds), *Unfolding learning societies: challenges and opportunities*. Shikshantar: The People's Institute for Rethinking Education and Development.

Kigotho, W. 2018. *The dangerous rise of neo-liberal universities*. University World News 9. https://www.universityworldnews.com/post.php?story=20181108130628468

Mbogo, R.W., Ndiao, E., Wambua, J.M., Ireri, N.W., & Ngala, F.W. 2020. Supervision challenges and delays in completion of PhD programmes in public and private universities: Experiences of supervisors and graduate students in selected universities in Nairobi, Kenya. *European Journal of Education Studies* 6(11), 261-78.

Motseke, M. 2016. Reasons for the slow completion of masters and doctoral degrees by adult learners in a South African township. *Australian Journal of Adult Learning* 56(3), 424-41.

Mugendi, S.M., & Githae, E.N. 2021. Prevalence of non-completion among postgraduate students in selected public universities in Kenya. *European Journal of Education Studies* 8(12), 143-59. http://dx.doi.org/10.46827/ejes.v8i12.4017

Mukhwana, E., Oure, S., Too, J. & Some, D.K. 2016. *State of postgraduate research training in Kenya*. Commission for University Education. Discussion Paper 02, Nairobi.

Njoya, W. 2019 (March 29). Let's delink degrees from employment. Wandia Njoya Blog. http://www.wandianjoya.com/blog/lets-delink-degrees-from-employment

Ntarangwi, M., & Okwany, A. 2020. *Contextually grounded scholarly writing*. Partnership for African Social and Governance Research (PASGR), Nairobi.

Nyamnjoh, F.B. 2019. Decolonizing the university in Africa. *Oxford research encyclopedia of politics*. https://doi.org/10.1093/acrefore/9780190228637.013.717

Okune, A. & Ulrich, K. 2021. Building collaborative habits, establishing sustaining relations: what is the role of a scholarly society today? *Platypus, The Castac Blog*. March 23, 2021, https://blog.castac.org/2021/03/building-collaborative-habits-establishing-sustaining-relations-what-is-the-role-of-a-scholarly-society-today/

Okwany, A., & Ebrahim, H. 2015. Rethinking epistemology and methodology in early childhood research in Africa. In A. Farrell, S.L. Kagan, & E.M. Tisdall (eds), *The SAGE Handbook of Early Childhood Research*. SAGE, pp. 432-48.

How inclusive and equitable is research in clinical psychology that focuses on the Global South?

Helen Niemeyer (Freie Universität Berlin) and Louis Schiekiera (Freie Universität Berlin)

This chapter focuses on how research in clinical psychology is conducted in low- and middle-income countries. Currently, Western researchers and funders determine the approaches and research agendas in clinical psychology. Approaches that assume universal realities and cross-cultural invariance of concepts and theories constructed in the Global North and applied to the Global South are called etic perspectives (Cheung, 2012). Psychotherapeutic interventions are generally developed and investigated mainly in WEIRD (Western, Educated, Industrialized, Rich, and Democratic) samples but are assumed to work universally with some minor cultural adaptations. The content of assessment materials for inclusion criteria or outcome measures is often not modified in studies in the Global South. The studies refer either to diagnostic criteria of international classification systems of disorders mostly assumed to be universal (APA, 2013; WHO, 2019) or to psychological constructs for which the original and translated assessment instruments are considered comparable and valid if the items are not modified. If evidence-based interventions, such as cognitive behavioral therapy (CBT), are culturally adapted, this typically involves replacing Western-specific examples in psychoeducation or materials, such as positive activity lists, with examples specific to the Global South to ensure adherence and fidelity.

Positive activity lists, for example, comprise various alternatives from which patients can choose. The activities should be integrated more often into the patients´ daily routine to stabilize and improve their mood. Cultural adaptation means that activities that require specific resources or material equipment or are more common in western countries, such as driving around by car just for distraction or a leisure trip, going to the cinema, doing yoga, or playing golf, are replaced by more easily accessible activities that do not require specific material resources and do not resemble western habits, such as community activities, meeting friends, preparing tea, and manual activities or handicrafts. Specific examples mean that if exemplary persons are described in the psychoeducative material to inform patients about disorder-specific symptoms and behaviors, their names and living situations are adapted. Exemplary persons for studies in Arab countries, for example, are given Arab names and are described as living in family settings that are common in this culture so that patients can more easily identify with them.

Beyond such surface adaptations, clinical studies are rarely informed by local knowledge. Evidence is lacking on whether it is sufficient that interventions are slightly culturally adapted, as described, or whether further adaptations to local contexts might be necessary, as well as how such adaptations should be informed or if specific interventions are not appropriate at all in some cultural contexts. The documentation of adaptations in clinical studies and reporting on the collaboration with partners in the Global South is often insufficient, even though documentation scales and recommendations have been published (Heim et al., 2021). We consider this strong influence of the Global North as problematic. As opposed to transferring interventions that are effective in WEIRD samples to the Global South, an alternative approach would be to design interventions to address mental health issues in the Global South based on local expertise and in non-hierarchical collaborations with local researchers. To our knowledge, no study has been conducted in which a locally informed approach was compared to Western CBT.

How psychotherapy is embedded in this context is also important. For mental health improvements in the Global South, interventions against poverty might be relevant in the first place (Ohrnberger et al., 2020; Wollburg et al., 2023). Poverty is hypothesized to reduce executive functioning and induce negative affective states such as anxiety and depression (Haushofer & Fehr, 2014). It is recommended that the effects of poverty reduction measures and mental health interventions (particularly, focusing on psychological mechanisms relevant to economic behavior) and their interconnections be investigated (ibid.). A recent study (Haushofer et al., 2020) examined the impact of an unconditional cash transfer and a five-week psychotherapy program on 5,756 individuals in rural Kenya. The results revealed that one year post-intervention, cash transfer recipients exhibited higher levels of economic well-being (increased consumption, asset holdings, and revenue) and psychological well-being compared to control groups. Psychological well-being was assessed by a subjective well-being index comprising questions about general health, perceived stress, happiness, and life satisfaction. In contrast, the psychotherapy program alone did not significantly affect psychological or economic outcomes. The effects of the combined treatment were similar to those of the cash transfer alone.

Blattman, Jamison, and Sheridan (2017) investigated the effects of CBT and cash grants on crime and violence reduction among criminally engaged men in Liberia. They found that CBT and cash grants independently reduced crime and violence initially, but these effects diminished over time. However, when cash grants followed CBT, there was a significant and sustained reduction in crime and violence for at least a year, suggesting a synergistic effect. The evaluation period was extended to ten years in a follow-up study (Blattman et al., 2023). Men who received therapy or therapy with cash were significantly less likely to engage in antisocial behaviors, including robbery, drug selling, and street fights, compared to a control group. The most robust impacts were observed in those receiving therapy and cash, especially among the highest-risk individuals. Taken together, these results indicate that

psychotherapy might be an essential component for mental health and behavior changes in some contexts if it is combined with measures targeting poverty. However, this may not always be the case, given that needs are complex and multifaceted.

Beyond the design of intervention studies in the Global South to best serve the intended beneficiaries, the diversity of the research teams and equality between researchers in collaboration and authorship are important. Clinical psychological studies are mainly conducted and published by researchers from the Global North, despite the importance of building research communities and fostering new researchers in the Global South.

First authorship

We conducted a systematic literature search on randomized controlled trials (RCTs) in psychotherapy research and investigated the authors' affiliations. Beyond the most recent body of knowledge, we also analyzed time trends. To investigate the regional representation of the first authors, we analyzed 17,095 authorship affiliations of psychotherapy RCTs published between 1990 and 2022. Our data collection process involved a PubMed search that yielded 20,862 abstracts. In our preprocessing phase, we excluded 275 hits lacking an abstract, 93 non-English abstracts, and 61 abstracts missing publication years. We then gathered further information on the authors' affiliations using OpenAlex. For 17,095 studies, we found information regarding the ISO 3166-1 alpha-2 country code of the academic institution the first author was affiliated with.

In the dataset, an overrepresentation of specific regional affiliations of first authors was observed, with 8,076 studies from North America, 5,682 from Europe, 1,972 from Asia, 1,031 from Oceania[1], 226 from Latin America and the Caribbean, and 108 from Africa. Figure 1 shows the changes in the relative representation of world regions in the psychotherapy RCT literature over time. In North America, there was a decline in relative representation, from 65 % in the 1990s to 36 % in the early 2020s. On the other hand, Europe showed a rise and then a slight decline in representation, from 27 % in the 1990s to 32 % by 2022. An upward trend was observed in Asia, increasing from 3 % in the 1990s to 23 % in the 2020s.

However, regions like Latin America, the Caribbean, Africa, and Oceania consistently had lower representation. For instance, Latin America and the Caribbean only slightly increased from 1% in the 1990s to 2 % in the 2020s. Africa's presence remained marginal, rising from 1 % in the 1990s to 2 % in the 2020s, and Oceania's contribution remained stable at around 5-6 % throughout the period. These patterns highlight a shift towards more globally diverse but still highly uneven first authorships, with growth in the psychotherapy literature in Asian countries contrasted by persistent underrepresentation in African and Latin

1 All papers from the 'Oceania' region originate from either Australia or New Zealand, both of which are considered Global North economies.

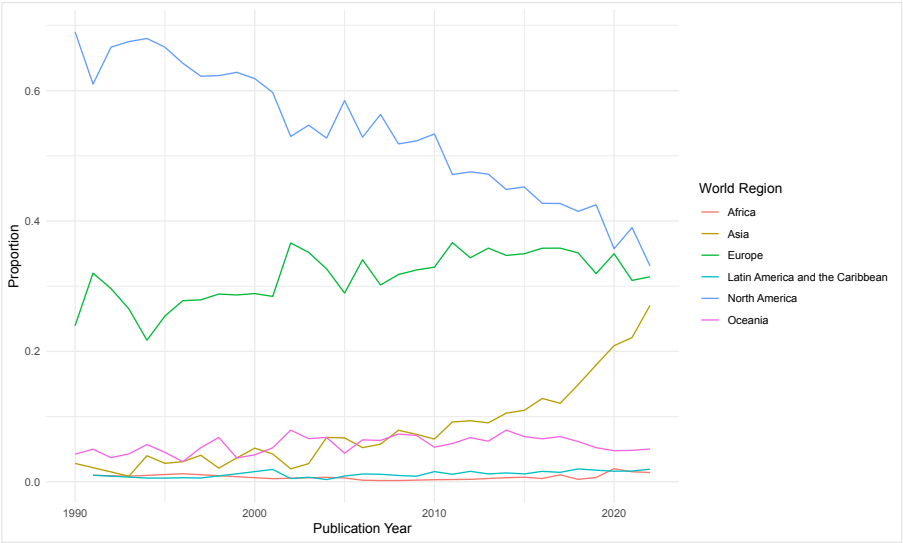

FIGURE 1. Proportion of academic affiliation in psychotherapy RCTs by world region over time. Note: n = 17,085 first authorships from 1990 to 2022

American countries. Two notable examples of Asian countries experiencing rapid growth in the relative representation of first-authorship affiliations are China and Iran, with relative representations of 8 and 5 %, respectively, in 2022, compared to no publications for either country before 1994.

Co-authorship

Moreover, we investigated the composition of research teams by analyzing the information about the affiliation of all authors. In the analyzed dataset, a predominance was observed in the authorship distribution in the Global North, with 14,965 studies featuring exclusively Global North affiliations, compared to 1,242 studies of sole Global South authorship and 888 studies reflecting a collaborative mix of Global North and South affiliations. Notably, a subset of the data, specifically 2,462 studies with otherwise exclusively Global North affiliations and 326 otherwise mixed-affiliation studies, lacked complete information on all authors' affiliations. Interestingly, all papers with sole Global South authorship affiliations reported complete details on the authors' affiliations.

Between 1990 and 2022, there was also a notable change in the composition of research teams. Figure 2 shows the changes in the relative representation of Global North, Global South, and mixed Global South/North teams in the psychotherapy RCT literature over time. In the 1990s, teams predominantly consisted of authors from the Global North (98 %). However, by 2020-2022, this number decreased to 75 %, indicating a reduction in the

dominance of Global North representation in the psychotherapy RCT literature. Conversely, the representation of teams entirely from the Global South showed an upward trend. In the 1990s, such teams were nearly absent in the literature (1 %), but by 2020-2022, their presence increased to 16 %.

Furthermore, the proportion of mixed teams, with authors from the Global South and North, increased over time. Initially, in the 1990s, mixed teams were scarce (2%). By the last observed period (2020-2022), their representation increased by nearly fivefold to 9 %, highlighting an increasing trend towards international collaboration between the Global North and South in this field. However, this analysis focuses solely on the quantitative aspects of collaborations. It is important to note that this methodology does not evaluate the quality of these collaborations or delve into the dynamics of hierarchy within them.

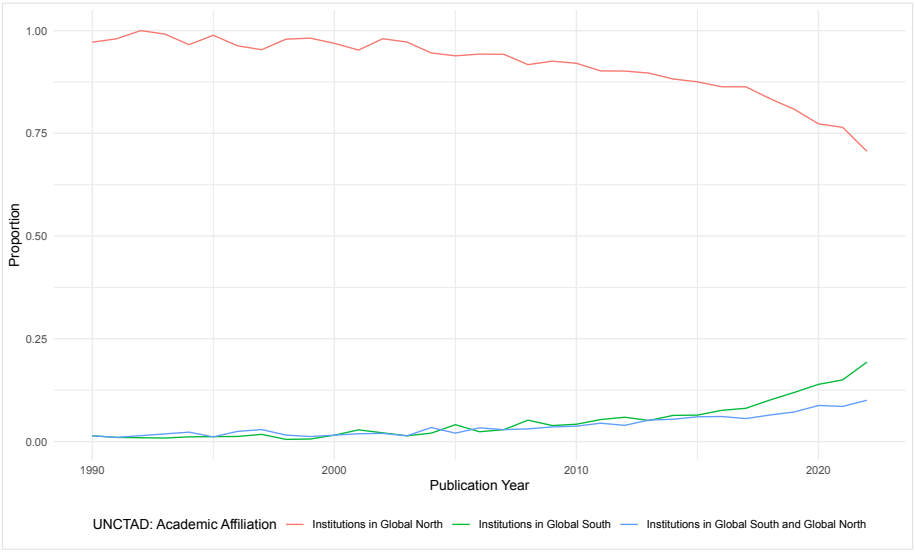

FIGURE 2. Proportion of authorship in psychotherapy RCT literature (1990-2022) by UNCTAD-region[2]. Note: n = 17,095 first authorships from 1990 to 2022.

Practical implementation

How are clinical psychological studies implemented in practice in the Global South? We conducted three exemplary interviews with German researchers focusing on traumatic experiences, forced displacement (internally and externally displaced persons), violence

2 The United Nations Conference on Trade and Development (UNCTAD) defines the Global South to broadly comprise Africa, Latin America and the Caribbean, Asia without Israel, Japan, and the Republic of Korea, and Oceania without Australia and New Zealand, while Global North economies broadly comprise Northern America and Europe, Israel, Japan, the Republic of Korea, Australia, and New Zealand.

and maltreatment, and the efficacy of prevention and intervention programs in the Global South. The interview partners were chosen based on an unsystematic literature review of Global South studies initiated or co-authored by German researchers. Due to the small number of interviews, these implementations should only be seen as examples.

One of the interviewees always collaborated with researchers from local universities. The scientific design and responsibility in these projects were initially located in the Global North and, in most cases, gradually shared with and handed over to local research partners under supervision and supported by education, if necessary. Local partners were otherwise non-profit organizations and lay persons or clinicians. The projects involved training interviewers, translators, and clinicians or lay persons.

The form and extent of the cultural adaptation and collaboration with local stakeholders varied. Two interview partners reported that the study designs and interventions were slightly culturally adapted with the help of locals to fit the needs and environment of the recipients. It comprised the surface adaptation of metaphors and images and of activities that are positively assessed and realistically feasible in the respective countries (as also described in the introduction of this book chapter). The third interview partner reported that each project starts with a survey and a feasibility study, in which local researchers are closely involved, to assess local needs and customs and adapt the study design and prevention programs.

One interview partner suggested that Western researchers should ideally live in the respective country for about one year before concretely planning a project there to understand the local habits and to experience the environment and culture. In contrast, another interviewee did not consider that necessary if good cooperation with local researchers and partners had already been established before a project started and if supervision and support, for example, via web conferences, is possible.

Implementing studies in the Global South was sometimes perceived as challenging. One researcher found that local participants or stakeholders did not well understand certain design aspects. An example was the implementation of a waitlist in intervention RCTs, which local partners interpreted as the withholding of support. However, this example is also found when working with non-profit organizations in the Global North and might affect the partner organizations. This points to the general dilemma that methodological priorities in research settings can impede comprehensive supply for the target population in the short term.

Moreover, the support of large Western agencies that can afford and provide the necessary IT infrastructure was considered indispensable for internet-based interventions. Even though open-source software can be passed on, non-profit organizations in the Global South often do not have the IT infrastructure on-site. This currently results in a central infrastructure in the Global North and decentralized teams in the Global North and South. Of note is that the

interviewees sometimes felt that partners in the Global South perceived it as an advantage if the infrastructure or funding were in a high-income country that is considered reliable and trustworthy.

Differences concerning desired cooperation within mixed Global South/North teams were reported. In a feasibility study, the partners on-site were determined and did not want influence from the Global North, and they clearly demanded authorship. In contrast, in another country, the attitude of local partners was rather subordinate ('we can learn from you, but there is nothing you can learn from us').

Evaluation

Studies in clinical psychology are evaluated when the results are published. The quality of the study and the reporting are assessed when a manuscript is submitted to a peer-reviewed journal. Reporting the successes and failures of research implementation and cultural adaptation in the Global South is optional. There is also no general evaluation of the collaboration with partners in the Global South concerning inclusivity, equality, and sustainability. Non-profit organizations sometimes evaluate the implementation of measures in their evaluation units. It is a challenge to bring together both worlds and evaluation systems of research and practice and to achieve optimal collaboration.

Two interviewees reported that many funders from the non-profit sector hardly evaluate their projects externally. One interview partner reported that an external evaluator's statistical evaluation of projects in the Global South seems rare, and internal evaluations often do not satisfy research standards. The validity can also be questioned because dependencies might bias the answers: If the jobs of local partners depend on the projects, and there are no alternative offers for participants, respondents might not want projects to be dropped due to negative evaluations.

Another interviewee described implementing a specific local evaluation system in all the projects: while partners from local non-profit organizations conducted the prevention programs in the field and collected the data, local research partners supervised and evaluated the implementation and screened the data concerning irregularities. Critical cases were discussed in the mixed Global South/North research team.

Two interviewees reported that many projects were not continued after the funding from Western agencies stopped, even if the infrastructure had been built up. An interviewee emphasized the priority of implementing long-term structures and the long-term promotion of local scientists in the projects, even if this requires considerable resources.

Close supervision was often perceived as necessary to ensure that local partners worked well in the structures that the Global North determined and that were new and not culturally familiar to them, especially when working solely with non-profit organizations. However, one interviewee noted that this seems to be an existing difficulty between research and practice

(also in the Global North). Taken together, we assume that the problems we identified in the three exemplary interviews are also prevalent in other studies in this field. We would like to encourage more research on implementing studies in the Global South. Moreover, we recommend that the publications of studies conducted in the Global South always include sections in which the collaboration with the partners in the Global South and the cultural adaptation of the intervention are described in detail. In addition, if sections on the cultural context and the political situation were always required for publication, this would be a bold step beyond changes in the research collaborations.

Conclusion

Our contribution critically analyzed the global organization of research within clinical psychology, focusing on research conducted in low- and middle-income countries. However, it is important to note that our analysis also emerges from a meta-perspective rooted in the Global North, an aspect we openly acknowledge. For future studies in clinical psychology, it is crucial to consider fundamentally culturally adapted strategies to address mental health challenges and living conditions in the Global South and to integrate local partners into the research process in equal and inclusive teams. Approaches that are locally co-developed are currently rare in clinical psychology but encouraging examples do exist. The global research community must ensure that research is equal, diverse, and serves the intended beneficiaries. Therefore, building local research skills might be necessary, including through practical training and mentorship (Adetula et al., 2022). We consider integrating local researchers in big team science and treating them as full partners in the research process an essential aim for clinical psychology. If research in the Global South does not merely take locals as participants but is constructed as teaming up with local researchers in a collaborative partnership for co-creating research agendas, this would change the current practice in which Western approaches and research instruments are merely transferred.

References

Adetula, A., Forscher, P., Basnight-Brown, D.M., Azouaghe, S., & IJzerman, H. 2022. Psychology should generalize from – not just to – Africa. *Nature Reviews Psychology* 1(7), 370-1. https://doi.org/10.1038/s44159-022-00070-y

American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders: Dsm-5* (5th ed.). Washington, DC: American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596

Blattman, C., Chaskel, S., Jamison, J.C., & Sheridan, M. 2023. Cognitive behavior therapy reduces crime and violence over 10 years: experimental evidence. *American Economic Review: Insights* 5(4), 527-45. https://doi.org/10.1257/aeri.20220427

Blattman, C., Jamison, J.C., & Sheridan, M. 2017. Reducing crime and violence: experimental evidence from cognitive behavioral therapy in Liberia. *American Economic Review* 107(4), 1165-206. https://doi.org/10.1257/aer.20150503

Cheung, F.M. 2012. Mainstreaming culture in psychology. *American Psychologist* 67(8), 721-30. https://doi.org/10.1037/a0029876

Haushofer, J., & Fehr, E. 2014. On the psychology of poverty. *Science* 344, 862-7. https://doi.org/10.1126/science.1232491

Haushofer, J., Mudida, R., & Shapiro, J.P. 2020. *The comparative impact of cash transfers and a psychotherapy program on psychological and economic well-being*. National Bureau of Economic Research, Working paper 28106. https://www.nber.org/papers/w28106

Heim, E., Mewes, R., Abi Ramia, J., Glaesmer, H., Hall, B., Harper Shehadeh, M., ... & Knaevelsrud, C. 2021. Reporting Cultural Adaptation in Psychological Trials – The RECAPT criteria. *Clinical Psychology in Europe* 3, 1-25. https://doi.org/10.32872/cpe.6351

Ohrnberger, J., Fichera, E., Sutton, M., & Anselmi, L. 2020. The worse the better? Quantile treatment effects of a conditional cash transfer programme on mental health. *Health Policy and Planning* 35(9), 1137-49. https://doi.org/10.1093/heapol/czaa079

Wollburg, C., Steinert, J.I., Reeves, A., & Nye, E. 2023. Do cash transfers alleviate common mental disorders in low- and middle-income countries? A systematic review and meta-analysis. *PLoS ONE* 18(2): e0281283. https://doi.org/10.1371/journal.pone.0281283

World Health Organization. 2019. *International statistical classification of diseases and related health problems* (11th ed.).

Collaborating to support a more inclusive evidence ecosystem

Maya Ranganath (Center for Effective Global Action)

Global threats such as climate change, violent conflicts, food insecurity, and inequality continue to disproportionately burden people living in poverty, especially in low- and middle-income countries (LMICs). Local scholars are particularly well-suited to study these threats and inform innovative and effective solutions. However, serious barriers – including access to training, networks, infrastructure, and resources – constrain their ability to participate in and lead global development research fully. African scholars, in particular, face significant structural barriers to generating research outputs. Currently, only two percent of the world's research output is produced by Africans, and less than eight percent of publications on Africa in top economics journals include an African co-author (UNESCO 2019). While this has been the topic of many studies, questions remain: What drives inclusive development research? What are the best ways to strengthen the research capacity of African institutions? How can equitable collaborations between researchers of different backgrounds best be encouraged? How can various actors in the space (for instance, think tanks, research centers, universities, and funders) fit together to support solutions to this complex problem? What incentivizes various actors to conduct, support, and utilize global development research? Who is funding these activities, and what are the power dynamics involved?

These questions, and many others, motivated *The Center for Effective Global Action* (CEGA) and the *Network of Impact Evaluation Researchers in Africa* (NIERA) to establish the *Collaboration for Inclusive Research* (CIDR) in 2022. CIDR aims to better understand the inclusion of African scholars in global development research along various stages of the evidence to impact the pipeline, create actionable guidelines for multiple actors, and ultimately drive change in development research toward a more diverse, equitable, and inclusive ecosystem. To date, CIDR has conducted an initial convening of over fifty stakeholders (African scholars, administrators and funders of capacity strengthening initiatives, journal editors, and others) to provide feedback on our research agenda, completed a literature review of relevant articles, held focus group discussions and qualitative interviews, and launched an extensive online survey which reached over five hundred stakeholders. Below, I share a description of the research process as well as emerging findings from the ongoing research.

Description of survey and demographics of respondents

CIDR's team leads (consisting of nine African academics and US faculty members affiliated with CEGA/NIERA) created the survey, which launched in December 2023. It was organized into four sections: higher education; training, and mentorship; publication and co-authorship; and policy impact. Various modules were displayed based on participants' chosen identities (they were asked to select up to two). Of the respondents, 25% were students, 25.2% were university faculty members, and 24% were research professionals (see table 1 and table 2 for further information).

VARIABLES	FREQ.	PERCENT
Professional role (N=512)		
Student	128	25
University faculty member	129	25.2
University administrator	6	1.2
Research professional	122	23.8
Development practitioner	67	13.1
Government official in Africa	44	8.6
Representative of a funding agency	16	3.1
Working as publisher or editor of an academic journal (N=512)		
No	365	71.3
Yes	147	28.7
Working as an administrator/coordinator of a training and/ or mentorship program for African scholars (N=512)		
No	398	77.7
Yes	114	22.3
Gender (N=497)		
Man	337	67.8
Woman	157	31.6
Non-binary/ third gender	1	0.2
Prefer not to say	2	0.4
Majority of education through secondary school in an African country (N=497)		
No	59	11.9
Yes	438	88.1

TABLE 1. *Demographic characteristics*

VARIABLES	FREQ.	PERCENT
Highest level of completed education (N=497)		
Secondary school	3	0.6
Bachelors	71	14.3
Masters	255	51.3
PhD	164	33
Other	4	0.8
Field of study for final degree (N=488)		
Economics	244	50
Statistics	25	5.1
Public health	20	4.1
Public policy	17	3.5
Development studies	55	11.3
Other	127	26
Highest degree institution based in Africa (N=487)		
No	149	30.6
Yes	338	69.4
Currently studying in Africa (N=188)		
No	52	27.7
Yes	136	72.3
Ever studied outside of Africa (N=429)		
No	282	65.7
Yes	147	34.3
Period of schooling attended outside of Africa (N=147)		
Primary school	1	0.7
Secondary school	1	0.7
Undergraduate	18	12.2
Masters	94	64
Doctoral	67	45.6
Other	26	17.7

TABLE 2. *Education in Africa vs. outside Africa*

Discussion of preliminary findings from survey and qualitative research

While our analysis is still ongoing, we share some of the perspectives of our respondents on higher education, training and mentorship, publication issues and authorship, and policy impact below.

Higher education

Our early findings show that African scholars' decision of where to pursue their Master's or PhD (in Africa or the US) is influenced by a lack of information about available opportunities and scholarships, the cost of application, job possibilities that students can pursue with an advanced degree, and the financial trade-offs of deferred earning potential. Notably, there are familial expectations about the cost of continued university education. Many families encourage students to get a job to pay them back for the investment in their education.

It has been suggested that organizing presentations at universities, both face-to-face and virtually, during which successful researchers are asked to describe their career paths and share essential facts with students, such as the information that moving directly from undergraduate to PhD-level education is typical in the US and that it is possible to attain a Master's degree along the way, can expand students' horizons.

Training and mentorship

Several organizations are running their own training and mentorship programs for scholars (apart from the mentorship students receive from advisors as part of their degree program), many of which we have collated into the CIDR Opportunity Hub (CO-Hub) for African Scholars, an online resource we have distributed to the community (see for the full list, https://airtable.com/appR4Y2kRkqjy5GVB/shr1VIz3hwN7ITTlt). Many offer short-term training, opportunities for research collaborations with established scholars, and longer-term training and have research funding earmarked for African scholars. Notably, African institutions offer many programs – a signal of an emerging training capacity. Of the African scholars we surveyed, 71% had applied for an external training or mentorship program, and 82% were accepted. Of those who did not apply, 73% cited not receiving information about the programs, demonstrating again that the lack of information is a crucial barrier. Our results showed that the men in our sample applied to these programs more frequently than women (72% vs. 66%). Scholars have indicated qualitatively that these programs are vital steps for students aiming to become prominent development researchers, given that their graduate programs do not focus on evidence-informed policy (and mostly cover macroeconomic research) and that their faculty supervisors do not have time to provide them with adequate mentorship on these topics because of their heavy teaching burden. Without their supervisors exposing them to relevant literature and resources, African students can find engaging in the development research space harder, making external mentorship programs valuable and necessary.

NGOs can play a role in this. For example, *Eider Africa* strives to decrease the isolation researchers experience, especially among recent graduates trying to publish their research (see https://eiderafricaltd.org). *Eider Africa* offers mentorship on essential research and writing skills and supports enhancing curricula at African universities. CEGA also offers the *East Africa Social Science Translation Collaborative* (EASST), a multi-pronged initiative that provides funding, mentorship, and in-depth individual and institutional capacity strengthening. The importance of identifying resources, such as the Afrobarometer, the

Harvard Dataverse, or datasets of published papers and training students to use them was also highlighted.

Publication and co-authorship

Addressing the publication gap mentioned above is challenging and requires a multi-pronged approach. Participants in one of our focus group discussions mentioned that a lack of information is behind why some African scholars find it challenging to publish in 'top' development journals, particularly the relevant knowledge about which journals to publish in, and what they are looking for. Other scholars mentioned that their universities do not value their publications or use this as a basis for promotion. A lack of funding to do high-quality research was cited as an additional obstacle. Female scholars face additional barriers to publishing, even at the highest levels of academia. For instance, Jeanine Condo, an EASST Fellow and the current Executive Director of Rwanda's *Centre for Impact, Innovation and Capacity Building for Health Information Systems and Nutrition*, recently published a paper on gender inequities in publishing. She found that of a sample of documents from sub-Saharan Africa (SSA), men comprised 61% of first authors. A recent survey of over two hundred alumni of SSA STEM PhD programs confirmed that women obtain less university and external funding for graduate studies than their male counterparts (Fisher et al., 2020). Other female researchers have stated that they had to push back against senior leadership at their institutions; often, they are expected to achieve the same results as their male counterparts but are constrained by additional, invisibilized care labor.

'Money is the problem and the solution,' CEGA affiliate and former EASST fellow Constantine Manda shared in a panel at the 2023 Africa Evidence Summit. Funds can be leveraged to support and nurture African researchers so that journals will be eager to publish their work. Jeanine Condo recently wrote a grant to research the effects of incentivizing researchers to publish instead of accepting small consultancies, where their contributions were less likely to be recognized by name. Preliminary results suggest that this is a compelling incentive for SSA researchers. It is also important to have HIC journals take strides to improve geographic equity and invest in the quality of African journals. For example, PLOS recently announced a policy that authors conducting research outside of their country of origin will be asked to complete a questionnaire that details the ethical, cultural, and scientific considerations taken to uphold inclusivity in their research, including if local authors are included in the authorship list (PLOS 2021). Making journals open-access was another solution proposed by the panel. Some have also urged well-cited HIC scholars to publish in African journals, which can be an important signal of their quality, while also taking strides to invest in and elevate them. It is also important to encourage other forms of research outputs, such as policy briefs and white papers, rather than solely striving for peer-reviewed publications.

Policy impact

The pathways for research to impact policy (for both foreign and local researchers) remain nebulous. In our early investigations, the medium to convey research findings to

policymakers is vital. One researcher, for instance, mentioned, 'When I have a new paper, I turn it into an infographic that I send directly to the Minister for Health's office in Rwanda via WhatsApp. This is the only way it will be seen'. Additionally, training policymakers on using rigorous evidence in their decision-making can be promising but challenging. While it can open up new avenues for research collaborations, it can also be tricky to involve the right person in the training; those who have the time to attend are often more junior and, therefore, have less decision-making power. Policymakers have urged researchers to study their strategic plans and to make space for research and methods according to their needs because decisions can often not wait until a randomized controlled trial is completed.

Overall, it remains to be seen whether local or foreign researchers ultimately have a more significant impact on policymaking. In some instances, local researchers can have deep relationships necessary to be trusted, and foreign researchers advising African governments can be perceived as colonial. In others, policymakers mention that the 'name brand', and lack of perceived political party affiliation, of those from top US and European universities can lead them to trust the findings more. Some have stated that if the project is a collaboration and includes researchers of different backgrounds, it can be strategic to have both present in the room.

Conclusion: towards true decoloniality?

The movement for inclusive scholarship has grown, underscored by the numerous articles published and the establishment of initiatives over the past decade. While many open questions on how to best center African scholars' leadership in development research remain unanswered, this important pursuit should continue and expand. For the research community to move to true decoloniality, the balance of power in global knowledge production must shift, especially when researchers study low and middle-income countries. We look forward to expanding this evidence base in the coming months as our CIDR research continues. Ultimately, while there remain open questions on how best to support African scholars, this important pursuit cannot (and should not) be contested. The power dynamics in global knowledge systems must be addressed to move towards true decoloniality in partnership with African institutions. The full results of the CIDR study will inform this. We will share our results in the summer of 2024, proposing recommendations for various stakeholders: universities in LMICs and HICs, funders, scholars, journals, and civil society organizations, and actively ensuring they are used.

References

Fisher, M., Nyaboro, V., Mendum, R., & Osiru, M. 2020. Making it to the PhD: gender and student performance in sub-Saharan Africa. *Plos One*. https://doi.org/10.1371/journal.pone.0241915

PLOS. 2021. Announcing a new PLOS policy on inclusion in global research. *The Official PLOS Blog*. https://theplosblog.plos.org/2021/09/announcing-a-new-plos-policy-on-inclusion-in-global-research/

UNESCO. 2019. *Recommendation on open educational resources (OER)*. https://unesdoc.unesco.org/ark:/48223/pf0000235406

PART TWO

How to improve the evidence-to-policy pipeline

Behavioral public policy for global challenges

Sanchayan Banerjee (Vrije Universiteit Amsterdam & London School of Economics and Political Science) and Matteo M. Galizzi (London School of Economics and Political Science)

Behavioral insights have been increasingly used to inform policy making over the last fifteen years (Oliver, 2018). Since the opening of the Behavioural Insights Team (the so-called 'nudge unit') within the UK Cabinet Office, a growing number of countries have set up over 250 behavioral units within governments, ministries and departments, national regulatory agencies, and other public bodies (OECD, 2023; Hunt & Adams, 2023). This trend has occurred all around the world, from western developed countries such as Canada, Ireland, and the US to middle eastern countries such as Lebanon, and from Latin American countries such as Peru to countries in the Global South such as South Africa and India (see Sanders et al., 2023). These behavioral units have then successfully informed public decision-making in several key behavioral public policy (BPP) areas, such as tax paying (Chadborn et al., 2023), public health (Ruggeri et al., 2024), labor programs, as well as pensions and savings (Thaler & Benartzi, 2004).

Despite this tremendous growth, it is fair to note that these national behavioral units have worked separately so far, employing a single-country perspective. This is natural as the various units report to national governments and authorities and reflects the inherent richness and diversity of national cultures and values, policy priorities and agendas, political programs and orientations, and electoral cycles. However, this also means that BPPs have been discussed and implemented within a single-country framework even for challenges requiring a much broader perspective. Increasingly, we face many global challenges, the damaging effects of which can unroll beyond national borders and last us a lifetime or beyond. These challenges pose manifold risks for humanity and the planet.

First and foremost, we face existential threats from climate change (IPCC, 2022). There are also other related, catastrophic risks, such as pandemic outbreaks, wars and conflicts; accelerating natural disasters, food, water, and energy shortages and insecurity, civil unrest; mass migrations caused or accelerated by all the previous reasons. All these can significantly impede human development too because stark geographical and income inequalities amplify them.

Tackling and hopefully solving these major global challenges requires modifying human behaviors collectively, systematically, and globally. Only with the foundation of behavioral units within major cross-country organizations has a broader, multi-country perspective been brought to design and implement BPPs. Multi-country behavioral insights units have been set up by leading non-profit organizations active in the Global South, such as Busara (Jang & Singh, 2023), but also, increasingly, within international organizations such as the OECD (OECD, 2021), the World Bank (Munoz Boudet et al., 2023), the European Union, the United Nations (MacLennan & Martin, 2023), the FAO, the World Health Organisation, as well as within the WHO Regional Office for Europe (Habersaat et al., 2023). This exciting and promising development in BPP marks the opportunity to further step up its potential to solve more complex challenges.

The future of human behavior in most policy domains, in fact, urgently requires a collective, coordinated, multi-disciplinary, and multi-sector effort (Box-Steffensmeier et al., 2022). Yet how we can do so in tractable, pragmatic, and feasible ways is still unclear. As we argue below, addressing these challenges in the future, first and foremost, requires a radical shift in our thinking – more attention is needed towards systemic changes and global structural issues. To realize this proposition, we suggest three ways to make BPPs better suited for tackling global challenges. Our first recommendation advances ongoing debates on expanding the toolkit of BPP and using its conceptual richness to embrace the methodological pluralism inherent in its makeup. More specifically, we suggest policymakers should use a broader BPP toolkit that goes beyond simple behavioral nudges and combines different tools for synergy. Following this, our second recommendation relates to the need to assess heterogeneity in the treatment effects of BPPs. This can enable the design of very specific and tailored BPPs that are more effective and legitimate. Finally, our third recommendation builds on the need for evidence-informed policymaking using systematic, reproducible, and transparent multi-country experimentations.

Adopt a broader toolkit of behavioral interventions

The first challenge shadowing contemporary BPP relates to the over-reliance on 'nudges,' which rely on systematically presenting choices to people to minimize unnecessary frictions that hinder the adoption of welfare-improving behaviors. To count as a nudge, an intervention must meet specific criteria, such as modifying the decision environment, or 'choice architecture,' without altering individual freedom, the number of available options, the relevant information, and the economic incentives (Thaler & Sunstein, 2009; Banerjee & John, 2023). Given the popularity of nudges, these original conceptual boundaries were often blurred or neglected, leading to fuzzy and broad definitions. More substantially, nudges are just one of the many possible policies in the broad BPP toolbox. This is well illustrated by the behavioral policy cube (Oliver, 2018; Banerjee, 2021), a collection of behaviorally informed public policies organized along three main dimensions, namely: i) regulatory vs. liberty-preserving policies; ii) appealing to rational vs. behavioral decision-making; and iii) tackling externalities vs. internalities (Oliver, 2018). Nudges are just one example of the

behaviorally informed public policies within the behavioral policy cube, other policies being bans ('shoves') or regulations of the supply side of markets ('budges') (ibid.).

Another notable example is strategies aimed at enhancing human agency and rationality (Banerjee et al., 2024), such as 'nudge+,' that is, nudging interventions accompanied by elements of reflective deliberation (Banerjee & John, 2023; 2024) and 'boosts,' that is, educational interventions aiming at enhancing informed decision-making (Hertwig & Grüne-Yanoff, 2017). Essentially, both nudge+ and boosting interventions work by creating empowered citizens. Additionally, behaviorally informed regulatory policies like shoves and sin taxes also exist, serving as disincentives intended to moderate the utilization of specific behavioral tactics or the consumption of welfare-reducing products or services.

However, despite this richness in the behavioral policy cube, BPP has often relied primarily on nudging, thereby failing to leverage the potential of alternative behavioral interventions (BIs). This is a barrier to tackling many of the global challenges listed above. For example, engaging in climate action requires lasting behavior change with autonomous decision-making. Nudges alone are unlikely to fully deliver on this challenge, as they are simple choice architectural modifications and thus light-touch interventions. While they can help close intention-behavior gaps for some individuals, in most cases they will fail to lead to sustained behavior change as people do not internalize psychological cues. In the long term, educational, agency-enhancing strategies may be better. Similarly, if we must decarbonize, more stringent measures, like carbon taxes, will be necessary too.

Recent tests have suggested that agency-enhancing toolkits like nudge+ are more effective than classic nudges (Banerjee et al., 2023a; Thamer et al., 2024). The lack of scaling-up or persistence in effect sizes of nudges indicates the need to harness synergies between different interventions. There is growing evidence that policy sequencing – which refers to a staggered implementation of policies in order of their stringency – can effectively achieve policy outcomes. For example, Gravert and Shreedhar (2022) argue that green nudging can help overcome behavioral biases, which otherwise hinder the acceptance of carbon taxes, thereby posting such a combination to be more effective – a point which is now garnering empirical support (see Faccioli et al., 2022; Alt et al., 2024). Multiple combinations of nudges have also been posited to be necessary for sustaining behavioral change, and they are usually more effective than single standalone nudges (Barbosa & Bermundez-Rey, 2024).
Our first suggestion thus hinges on the need to openly adopt the broadest toolkit of BIs and BPPs, not limited to nudges but also including information policies, boosts, nudge+, incentives, taxes, subsidies, regulation, and bans, and its policy combinations and sequencing therein. To overcome major global challenges, pluralism in the toolkit of the behavioral policy-maker and synergies between different BPPs should be leveraged better. We acknowledge that most real-life policies are not shaped in silos and that different policies often interact with each other in practice. However, this is yet to be common in academic scholarship (Beshears & Kosowsky, 2020). Therefore, we call for researchers to embrace

this reality when developing and testing policies in more controlled environments.

Assess heterogeneity, sustained effects, and behavioral spillovers

Most contemporary BPP studies over-rely on nudges and typically focus exclusively on assessing whether, on average, such nudges can change the outcome of interest (that is, on estimating the 'average treatment effect' on the dependent variable). To effectively tackle global challenges, the next wave of BPP studies must innovate on the current evidence in at least three respects.

First, they will need to track the sustained impact of the BIs over time to assess their carryover longer-term effects (Dolan & Galizzi, 2015; Thamer et al., 2024). Adding this longitudinal perspective to BPP studies is becoming increasingly feasible using a 'behavioral data linking' approach, that is, linking data from behavioral experiments to sources of longitudinal and 'smart' data such as biomarkers banks, administrative registers, electronic records, panel surveys, mobile and wearable devices, apps, smart cards, scan data, and geo-coded data (Thomas et al., 2024).

Second, BPP studies need to systematically map not only the effects of the interventions on the primarily targeted outcomes of interest but also their 'behavioral spillover' and anticipatory effects, that is, whether and how BIs affect other, non-targeted behaviors (Dolan & Galizzi, 2015; Galizzi & Whitmarsh, 2019; Picard & Banerjee, 2023). Understanding carryover and spillover effects of BIs is crucially important when it comes to complex, systemic patterns of behaviors, mainly because these 'ripple' effects are amplified by increasingly complex interconnections between our 'online' and in-person 'selves': to have a durable impact on human wellbeing, the planet, and global societal welfare, the change in behaviors needs to endure over time and across multiple contexts and decisions, not being just a one-off change.

Third, future BPP studies will have to look at the heterogeneity of the effects of BIs systematically (Ruggeri et al., 2024). There is growing acknowledgment in behavioral science that we fully account for human heterogeneity along several dimensions. People have heterogeneous attitudes, beliefs, perceptions, and preferences. There is growing evidence on how heterogeneous individual beliefs and perceptions drive, mediate, and moderate behavioral change as a response to different interventions and BPPs (Galizzi et al., 2022; Banerjee et al., 2023a, 2023b). For instance, an extensive body of behavioral economics evidence has documented high heterogeneity of individual preferences (for risk preferences, for example, see Camerer, 1989; Hey & Orme, 1994; Ballinger & Wilcox, 1997). The role of heterogeneous preferences is even more critical given the growing literature exploring the ability of economic preferences to predict real-world outcomes (Barsky et al., 1997; Chabris et al., 2008; Tanaka et al., 2010; Sutter et al., 2013, Galizzi & Navarro-Martinez, 2019; Campos-Mercade et al., 2021, Epper et al., 2022). Similarly, people are inherently heterogeneous in that they have heterogeneous beliefs and perceptions.

This opens up the intriguing and promising possibility of linking broad ranges of BPPs and BIs to a preceding 'measurement' stage measuring a host of relevant preferences, attitudes, and perceptions at an individual level, such as time preferences, time perception, risk aversion, probability weighting, loss aversion, regret aversion, information avoidance preferences, altruism, fairness, trust, cooperation, positive and negative reciprocity, other social preferences, personality traits, psychological reactance, as well as other key attitudes, psychological traits, and beliefs (Steinert et al., 2022; Banerjee et al., 2023b; Thomas et al., 2024).

Linking the measurement and the intervention stages will potentially allow BPP researchers to look at the heterogeneity of the treatment effects across different 'behavioral phenotypes,' such as underlying preferences. It will also enable BPP studies to better understand latent variables' etiology and to uncover the complex mechanisms that can mediate or moderate the effectiveness of different BIs. By doing so, it will also be possible to isolate the ultimate behavioral motive of an observed spillover or carryover effect from alternative or concurrent explanations. BPP researchers will, in addition, be able to better inform the design of customized and personalized nudges and other BIs and to map and measure their asymmetric, distributional, and welfare effects (Mills, 2022; Sunstein, 2022).

Furthermore, people not only have heterogeneous preferences, attitudes, and beliefs but also make heterogeneous decisions even when facing similar situations and constraints. As a result, we should not only employ a complete and diverse spectrum of BPPs and BIs, but we should also expect that individual responses to such BIs will be heterogeneous (Steinert et al., 2022; Campos-Mercade et al., 2021; Milkman et al., 2022). As it is unclear upfront what BPPs will work for whom, we should thus engage more in more systematic and transparent experimentation to inform BPPs to tackle global challenges. This naturally leads us to the final suggestion.

Experiment more and include systematic, transparent, and reproducible comparisons and cross-country studies

A related challenge is the lack of systematic and transparently comparable findings across BPP studies. Typically, BPP studies are conducted in one-off settings, which makes it challenging to generalize findings across other samples and contexts (OECD, 2021). Studies also often limit themselves to policy evaluation of singular BPPs, such as nudges, which further limits any comparative analysis of evidence across the broader behavioral toolkit. One natural response to this challenge has been to undertake large-scale replications or extend tests of BPPs to other settings. Such an exercise is potentially helpful and promises to improve the overall replicability, external validity, and generalisability of BPP studies (Munafo' et al., 2017).

However, replicating existing BPP studies in different settings may still suffer from three main drawbacks that can limit their suitability in tackling global challenges effectively. First,

pure replication does not necessarily ensure the transparency and reproducibility of BPP studies. The next wave of BPP studies needs to fully embrace the whole set of best practices to ensure transparency and reproducibility, such as Registered Reports, pre-registrations, and pre-analysis plans for hypotheses, experimental designs, data coding, and analyses; ex-ante pre-registered sample size calculations, and sufficiently powered sample sizes; pre-registered rigorous statistical analysis of experimental data; transparent, pre-registered, and adequately powered sub-group analysis and analysis of heterogeneous treatment effects; statistical corrections for multiple hypotheses testing; statistical equivalence tests, including two one-sided tests (TOST) procedure; data and code sharing, also via open science and publicly available reproducibility packages; engagement, collation, and sharing of data for systematic reviews and meta-analyses (Mertens et al., 2022).

Second, pure replication does not necessarily allow a direct comparison of different BPPs targeting the same behavior. Studies have recently been suggested to undertake systematic tests of different BPPs in the same experimental setup. This is important, especially because different evaluation frames can influence behavioral outcomes underlying different policies, as demonstrated by Davidai and Shafir (2019). Bradt (2022) compared two different BIs, a nudge and a boost, in improving flood insurance demand. Galizzi et al. (2022) systematically compared the effects of different social norms and messages about the proportions of vaccinated people in a community on the intention to get a seasonal flu vaccine, finding both band-wagoning and free-riding effects. Banerjee et al. (2023a) substantially extended this recent line of research by systematically testing ten interventions across four broad behavioral toolkits, namely nudges, boosts, thinks, and nudge+ interventions, in the context of sustainable diets.

> **Designing and running comprehensive experiments generalizable to different contexts, samples, and toolkits enables behavioral policymakers to compare and contrast evidence about what works and what does not. Doing so is critical to solving many of the global challenges which will have heterogeneous impacts on communities across the globe.**

Third, pure replication does not necessarily add greater comprehensibility since the different studies – even on the same topic or BPP – might be temporally separated and, in many cases, contextually different (Feest, 2019). Given the greater need to assess numerous BIs simultaneously under the same conditions, a large-scale version of systematic testing called 'mega-study' has been recently proposed. Mega studies are 'massive field experiment(s) in which many different treatments are tested synchronously in one large sample using a common, objectively measured outcome' (Duckworth & Milkman, 2022, p. 214). Mega-studies in behavioral sciences are fast-growing (Milkman et al., 2021a), with examples spanning different fields of application such as, among others, personal and public health (Milkman et al., 2022; Koenig et al., 2024), misinformation (Arechar et al., 2023),

environment and climate change (Vlasceanu et al., 2024).

Mega studies are thus more systematic, comprehensive, and transparent than standard one-off experiments. Importantly, they also foster ex-ante collaboration of large teams around cooperative and comparative research, potentially steering the norms and incentives of the behavioral community towards large-scale collaborative efforts rather than competing, often duplicating, small-scale projects (OECD, 2021). However, they also have limitations, including focusing on selective samples, thereby generating only partial evidence. There are at least two potential solutions to this limitation. One is to conduct systematic and coordinated mega-studies across different countries and settings. Steinert et al. (2022), for example, systematically compare the effects of BIs revolving around social norms, health literacy, and messages about societal and individual benefits (for example, vaccine passports) on the intention to get vaccinated against COVID-19 in Bulgaria, Italy, France, Germany, Poland, Spain, Sweden, and the UK. In a similar spirit, Banerjee et al. (2024) compares a hypothetical default nudge and nudge+ policy using 24,303 people across the G-7. Another solution is to run integrative experiments, as proposed by Almaatouq et al. (2022). Integrative experiments apply the science of experimentation to experiments themselves, first by mapping all the possible experiments that could underlie any given problem and then subsampling and testing a subset of them to infer evidence more generally about the whole population of experiments. Mega studies are integrative experiments that take a convenience sampling approach to this large-scale testing. Ghai and Banerjee (2024) propose extending this integrative experimental approach by explicitly accounting for sample diversity in the design of these multiple experiments to increase the generalizability of the evidence.

Designing and running comprehensive experiments generalizable to different contexts, samples, and toolkits enables behavioral policymakers to compare and contrast evidence about what works and what does not. Doing so is critical to solving many of the global challenges which will have heterogeneous impacts on communities across the globe. Not doing so and limiting tests to one-off settings runs the risks of overgeneralizing and extrapolating evidence to design policy where it might not be applicable – such a misfit can worsen behavioral outcomes and create bottlenecks that can be avoided. We thus recommend more experimentation and more systematic and transparent cross-country evaluations of behavioral toolkits spanning different samples and settings globally.

Conclusion

BPP should next address global challenges such as climate change, pandemics, conflicts, food security, and mass migration. In order to succeed in such a major step-up in its mission, BPP should fully embrace a global and systemic approach. Greater attention to systemic changes and global structural issues is needed. The rigorous analysis and systematic comparison of multiple BPP interventions is imperative to shed light on their relative effectiveness. Identifying heterogeneity in behavioral changes is a fundamental endeavor

that can unlock the isolation of the ultimate drivers and mechanisms behind responses to different BPP interventions. These mechanisms and underlying motives can, in turn, trigger reinforcing or compensatory feedback effects that can outlast the initial response to BPP actions and reverberate across individuals, over time, and across different behaviors. Finally, systematic, transparent, and reproducible cross-country studies spanning different samples and settings in the global world and comparing multiple BPP interventions are essential to scale up evidence from single-country small-scale insights to a proper global perspective. If BPP scholars and practitioners want to stand a chance to genuinely change the world, a global and systemic point of view is long due.

References

Almaatouq, A., Griffiths, T.L., Suchow, J.W., Whiting, M.E., Evans, J., & Watts, D.J. 2022. Beyond playing 20 questions with nature: integrative experiment design in the social and behavioral sciences. *Behavioral and Brain Sciences* 47, 1-55. https://doi.org/10.1017/S0140525X22002874

Alt, M., Bruns, H., DellaValle, N., & Murauskaite-Bull, I. 2024. Synergies of interventions to promote pro-environmental behaviors. A meta-analysis of experimental studies. *Global Environmental Change* 84, 102776. https://doi.org/10.1016/j.gloenvcha.2023.102776

Arechar, A.A., Allen, J., Berinsky, A.J., Cole, R., Epstein, Z., Garimella, K., ... & Rand, D. G. 2023. Understanding and combatting misinformation across 16 countries on six continents. *Nature Human Behavior* 7(9), 1502-13. https://doi.org/10.1038/s41562-023-01641-6

Ballinger, T.P., & Wilcox, N.T. 1997. Decisions, error and heterogeneity. *The Economic Journal* 107(443), 1090-105.

Banerjee S., Grüne-Yanoff, T., John, P., & Moseley, A. 2024. It's time we put agency into behavioral public policy. *Behavioural Public Policy*. https://doi.org/10.1017/bpp.2024.6

Banerjee, S. 2021. Rethinking the origin of the behavioral policy cube with nudge plus. In V. Mihaila (ed.), *Behavioral-based interventions for improving public policies*. IGI Global, pp. 1-16.

Banerjee, S., & John, P. 2023. Nudge+: putting citizens at the heart of behavioral public policy. In C. Sunstein & L. Reisch (eds), *Research handbook on nudges and society*. Edward Elgar Publishing, pp. 227-41.

Banerjee, S., & John, P. 2023. Nudge and nudging in public policy. In M. Van Gerven, C. Rothmayr Allison & K. Schubert (eds), *Encyclopedia of public policy*. Springer.

Banerjee, S., & John, P. 2024. Nudge plus: incorporating reflection into behavioral public policy. *Behavioural Public Policy* 8(1), 69-84. http://doi.org/10.1017/bpp.2021.6

Banerjee, S., Galizzi, M.M., John, P., & Mourato, S. 2023a. Sustainable dietary choices improved by reflection before a nudge in an online experiment. *Nature Sustainability* 6(12), 1632-42. https://doi.org/10.1038/s41893-023-01235-0

Banerjee, S., Galizzi, M.M., John, P., & Mourato, S. 2023b. Immediate backfire? Nudging sustainable food choices and psychological reactance. *Food Quality and Preference* 109, 104923. https://doi.org/10.1016/j.foodqual.2023.104923

Banerjee, S., John, P., Nyhan, B., Hunter, A., Koenig, R., Lee-Whiting, B., ... & Savani, M. 2024. Thinking about default enrollment lowers vaccination intentions and public

support in G7 countries. *PNAS Nexus* 3(4), pgae093. https://doi.org/10.1093/pnasnexus/pgae093

Barbosa, S., & Bermudez-Rey, J.P. 2024. The whole is larger than the sum of its parts: additive effects of SMS nudge bundles. *Frontiers in Communication* 9, 1298607. https://doi.org/10.3389/fcomm.2024.1298607

Barsky, R.B., Juster, F.T., Kimball, M.S., & Shapiro, M.D. 1997. Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study. *The Quarterly Journal of Economics* 112(2), 537-79. https://doi.org/10.1162/003355397555280

Beshears, J., & Kosowsky, H. 2020. Nudging: progress to date and future directions. *Organizational Behavior and Human Decision Processes* 161, 3-19. https://doi.org/10.1016/j.obhdp.2020.09.001

Box-Steffensmeier, J.M., Burgess, J., Corbetta, M., Crawford, K., Duflo, E., Fogarty, L., ... & Wagner, C. 2022. The future of human behaviour research. *Nature Human Behaviour* 6(1), 15-24. https://doi.org/10.1038/s41562-021-01275-6

Bradt, J. 2022. Comparing the effects of behaviorally informed interventions on flood insurance demand: an experimental analysis of 'boosts' and 'nudges'. *Behavioural Public Policy* 6(3), 485-515. https://doi.org/10.1017/bpp.2019.31

Camerer, C.F. 1989. An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty* 2, 61-104. https://doi.org/10.1007/BF00055711

Campos-Mercade, P., Meier, A.N., Schneider, F.H., & Wengström, E. 2021. Prosociality predicts health behaviors during the COVID-19 pandemic. *Journal of Public Economics* 195, 104367. https://doi.org/10.1016/j.jpubeco.2021.104367

Chabris, C. F., Laibson, D., Morris, C.L., Schuldt, J.P., & Taubinsky, D. 2008. Individual laboratory-measured discount rates predict field behavior. *Journal of Risk and Uncertainty* 37, 237-69. https://doi.org/10.1007/s11166-008-9053-x

Chadborn, T., Sallis, A., & Porter, L. 2023. The world's first government behavioral insights team dedicated to public health: ten lessons learned over a decade of experience. In M. Sanders, S. Bhanot & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 97-109.

Davidai, S., & Shafir, E. 2020. Are 'nudges' getting a fair shot? Joint versus separate evaluation. *Behavioural Public Policy* 4(3), 273-91. https://doi.org/10.1017/bpp.2018.9

Dolan P., &, Galizzi M.M. 2015. Like ripples on a pond: behavioural spillovers and their consequences for research and policy. *Journal of Economic Psychology* 47, 1-16. https://doi.org/10.1016/j.joep.2014.12.003

Duckworth, A.L., & Milkman, K.L. 2022. A guide to megastudies. *PNAS Nexus* 1(5), pgac214. https://doi.org/10.1093/pnasnexus/pgac214

Epper, T., Fehr, E., Hvidberg, K.B., Kreiner, C.T., Leth-Petersen, S., & Nytoft Rasmussen, G. 2022. Preferences predict who commits crime among young men. *Proceedings of the National Academy of Sciences* 119(6), e2112645119. https://doi.org/10.1073/pnas.2112645119

Faccioli, M., Law, C., Caine, C.A., Berger, N., Yan, X., Weninger, F., ... & Bateman, I.J. 2022. Combined carbon and health taxes outperform single-purpose information or fiscal measures in designing sustainable food policies. *Nature Food* 3(5), 331-40. https://doi.org/10.1038/s43016-022-00482-2

Feest, U. 2019. Why replication is overrated. *Philosophy of Science* 86(5), 895-905. https://

doi.org/10.1086/705451
Galizzi, M.M., & Whitmarsh, L.E. 2019. How to measure behavioural spillovers? A methodological review and checklist. *Frontiers in Psychology* 10, 0342. https://doi.org/10.3389/fpsyg.2019.00342
Galizzi, M.M., Lau, K.W., Miraldo, M., & Hauck, K. 2022. Bandwagoning, free- riding and heterogeneity in influenza vaccine decisions: an online experiment. *Health Economics* 31(4), 614-46. https://doi.org/10.1002/hec.4467
Galizzi, M.M., & Navarro-Martinez, D. 2019. On the external validity of social preference games: a systematic lab-field study. *Management Science* 65(3), 976-1002. https://psycnet.apa.org/doi/10.1287/mnsc.2017.2908
Ghai, S., & Banerjee, S. 2024. The future of experimental design: integrative, but is the sample diverse enough? *Behavioral and Brain Sciences* 47, e42. https://doi.org/10.1017/S0140525X23002212
Gravert, C., & Shreedhar, G. 2022. Effective carbon taxes need green nudges. *Nature Climate Change* 12(12), 1073-4. https://doi.org/10.1038/s41558-022-01515-1
Habersaat, K.B., Fietje, N., Koylyu, A., Scherzer, M., & Likki, T. 2023. Behavioral and cultural insights for better health: the BCI unit at WHO regional office for Europe. In M. Sanders, S. Bhanot & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 317-27.
Hertwig, R., & Grüne-Yanoff, T. 2017. Nudging and boosting: steering or empowering good decisions. *Perspectives on Psychological Science* 12(6), 973-86. https://doi.org/10.1177/1745691617702496
Hey, J.D., & Orme, C. 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica: Journal of the Econometric Society* 62(6), 1291-326.
Hunt, S., & Adams, P. 2023. Building a multidisciplinary team: lessons from financial regulation. In M. Sanders, S. Bhanot, & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 277-94.
IPCC. 2022. *Climate change 2022: impacts, adaptation, and vulnerability. Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, doi:10.1017/9781009325844.
Jang, C., & Singh, A. 2023. The Busara Center: letters to our past. In M. Sanders, S. Bhanot, & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 361-78.
Koenig, R., Savani, M.M., Lee-Whiting, B., McAndrews, J., Banerjee, S., Hunter, A., ... & Nyhan, B. 2024. Public support for more stringent vaccine policies increases with vaccine effectiveness. *Scientific Reports* 14(1), 1748. https://doi.org/10.1038/s41598-024-51654-y
MacLennan, M., & Martin, L. 2023. Embedding behavioral science into the work of the United Nations. In M. Sanders, S. Bhanot, & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 329-42.
Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. 2022. The effectiveness of nudging: a meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences* 119(1), e2107346118. https://doi.org/10.1073/pnas.2107346118
Milkman, K.L., Gandhi, L., Patel, M. S., Graci, H.N., Gromet, D.M., Ho, H., ... & Duckworth, A.L. 2022. A 680,000-person megastudy of nudges to encourage vaccination in

pharmacies. *Proceedings of the National Academy of Sciences* 119(6), e2115126119. https://doi.org/10.1073/pnas.2115126119

Milkman, K.L., Gromet, D., Ho, H., Kay, J.S., Lee, T.W., Pandiloski, P., ... & Duckworth, A.L. 2021. Megastudies improve the impact of applied behavioural science. *Nature* 600(7889), 478-83. https://doi.org/10.1038/s41586-021-04128-4

Mills, S. 2022. Personalized nudging. *Behavioural Public Policy* 6(1), 150-9. https://doi.org/10.1017/bpp.2020.7

Munafò, M., et al. 2017. A manifesto for reproducible science. *Nature Human Behaviour* 1, 0021. https://doi.org/10.1038/s41562-016-0021

Munoz Boudet, A.M., Scheunemann, L., Farfan, G., & Castaneda, J.L. 2023. eMBeDding behavioral sciences in international development. In M. Sanders, S. Bhanot, & S. O'Flaherty (eds), *Behavioral public policy in a global context*. Palgrave Macmillan, pp. 295-314.

OECD. 2021. *Achieving cross border government innovation: surfacing insights and experimenting across borders*. OECD OPSI Report. Available Here: 2021-08-09-OECD-Report2-Print-V3.pdf (oecd-opsi.org)

OECD. 2023. *Behavioural insights and public policy: institutions applying behavioural science to public policy around the world*. Available here: https://oecd-opsi.org/blog/mapping-behavioural-insights/

Oliver, A. 2018. Nudges, shoves and budges: behavioural economic policy frameworks. *The International Journal of Health Planning and Management* 33(1), 272-5. https://doi.org/10.1002/hpm.2419

Picard, J., & Banerjee, S. 2023. *Behavioural spillovers unpacked: estimating the side effects of social norm nudges*. Centre for Climate Change Economics and Policy Working Paper 427/Grantham Research Institute on Climate Change and the Environment Working Paper 402. London School of Economics and Political Science.

Ruggeri, K., Stock, F., Haslam, S.A., Capraro, V., Boggio, P., Ellemers, N., ... & Willer, R. 2024. A synthesis of evidence for policy from behavioural science during COVID-19. *Nature* 625, 134-47. https://doi.org/10.1038/s41586-023-06840-9

Sanders, M., Bhanot S., & O'Flaherty S. (eds) 2023. *Behavioural public policy in a global context*. Palgrave MacMillan.

Steinert, J. I., Sternberg, H., Prince, H., Fasolo, B., Galizzi, M.M., Büthe, T., & Veltri, G.A. 2022. COVID-19 vaccine hesitancy in eight European countries: Prevalence, determinants, and heterogeneity. *Science Advances* 8(17), eabm9825. https://doi.org/10.1126/sciadv.abm9825

Sunstein, C.R. 2022. The distributional effects of nudges. *Nature Human Behaviour* 6(1), 9-10. https://doi.org/10.1038/s41562-021-01236-z

Sutter, M., Zoller, C., & Glätzle-Rützler, D. 2019. Economic behavior of children and adolescents: a first survey of experimental economics results. *European Economic Review* 111, 98-121. https://doi.org/10.1016/j.euroecorev.2018.09.004

Tanaka, T., Camerer, C.F., & Nguyen, Q. 2010. Risk and time preferences: linking experimental and household survey data from Vietnam. *American Economic Review* 100(1), 557-71. https://doi.org/10.1257/aer.100.1.557

Thaler, R.H., & Benartzi, S. 2004. Save more tomorrow™: using behavioral economics to increase employee saving. *Journal of Political Economy* 112(S1), S164-87. https://doi.

org/10.1086/380085

Thaler, R.H., & Sunstein, C.R. 2009. *Nudge: improving decisions about health, wealth, and happiness*. Penguin.

Thamer, P., Banerjee, S., & John, P. 2024. Pledging after nudging improves uptake of plant-based diets: A field experiment in a German university cafeteria. *Environmental Research Communications*. https://doi.org/10.1088/2515-7620/ad2625

Thomas, R., Galizzi, M.M., Moorhouse, L., Nyamukapa, C., & Hallett, T.B. 2024. Do risk, time, and prosocial preferences predict risky sexual behaviour of youths in a low-income, high-risk setting? *Journal of Health Economics* 93, 102845. https://doi.org/10.1016/j.jhealeco.2023.102845

Vlasceanu, M., et al. 2024. Addressing climate change with behavioral science: A global intervention tournament in 63 countries. *Science Advances* 10, eadj5778. https://doi.org/10.1126/sciadv.adj5778

Building research capacity in Africa via big team science: challenges and lessons learned from the ManyLabs Africa initiative

Adeyemi Adetula (Alex Ekwueme Federal University Ndufu-Alike & Université Grenoble Alpes)

Despite constituting 79% of the world population, the Global South constitutes less than 6% of the samples used in papers in popular psychology journals (Thalmayer et al., 2021). In economics, even among those articles that focus on Africa, only 25% have at least one Africa-based co-author (Chelwa, 2021). In light of these facts, scholars have argued that making room for Global South research contributions will maximize their impact on human and social development (Forscher et al., 2021). Unfortunately, while a few Global South countries, such as China and Singapore, have the economic resources to meet their research and development needs, most countries from the Global South – especially African ones – are low-income economies grappling with limited research resources, facilities, and capacities to participate in global research (Kigotho, 2021). The cumulative effort to diversify and expand research investigation to countries in the Global South has fallen short and typically failed to capture local priorities (Liverpool, 2021). This is especially true of African countries.

One research approach to potentially address these problems is big team science. Big team science is an approach to pool resources, infrastructure, and expertise across geographically dispersed teams to accomplish projects larger than any individual team could individually (Forscher et al., 2022). Big team science is a product of the credibility revolution, a movement for greater rigor, openness, transparency, and more credible claims that emerged after psychology's replication crisis (Korbmacher et al., 2023). Big team science commits to inclusive and robust investigations for more generalizable and applicable science at a low cost for partners and with greater benefit to all. When implemented well, big team science offers all team members – including those located in African countries – extensive access to resources, infrastructure, and expertise, as well as credibility-enhancing practices. These resources can position African research partners to expand their personal and professional networks, improve their research skills, and acquire the right tools and resources to work globally. Yet, this approach remains unsuccessful at engaging and integrating Africans (Adetula et al., 2022).

Big team science originated in the Global North and is, as of yet, unpopular among African researchers. Even the few African researchers involved in big team science have limited expertise and struggle with poor infrastructure (such as unreliable power, intermittent and expensive internet, a lack of a proper workspace, and inadequate computing) to contribute and coordinate these studies locally. Furthermore, most research questions that are currently investigated in big team science research Western-centric with little or no relevance to Africans. Big team science, therefore, runs the risk of supercharging Western dominance and missing out on crucial variations, research questions, and methods in African behavioral sciences. By laying the groundwork for participating in such initiatives as big team science, improving research capacities in Africa will optimize research for African development and promote synergistic cooperation between African and non-African countries.

In this chapter, I argue that building research capacity will improve African research productivity. I make my argument by (1) highlighting big team science's huge potential but limited impact on research capacity building, and (2) describing the efforts of myself and my collaborators to build research capacity through our own ManyLabs Africa initiative. I conclude by (3) describing lessons from the ManyLabs Africa initiative and recommendations on navigating big team science in Africa.

Big team science can advance research capacity

Many African research capacity-building initiatives target human development in Africa, some of which are covered in part 1 of this edited volume. While such initiatives are laudable, the research sector can also be improved by engaging researchers in big-team science replication-type studies for on-the-job training on cutting-edge research skills. These studies can engage African researchers while testing replicability and generalizability across and beyond their populations. Big team science distributes tasks that require partners' expertise and experience. These tasks include project management for central and local coordination, manuscript writing and review, project conceptualization and design, material preparation such as study selection, ethics review, instrument cultural adaptation and translation, tool development and translation implementation for online and laboratory studies, data collection, and funding which entails grant sourcing and the writing of applications. Such big team science initiatives can facilitate the sharing and synergizing of thoughts, expertise, and resources on a large scale. As long as project roles are distributed equitably, these studies can also minimize the power imbalances between the partners with more resources and Africans with fewer resources, allowing the latter to negotiate barriers and center their priorities. This approach would also enable differences in psychology and behavior to emerge within studies between Africans and non-Africans and different African sub-populations.

Despite this potential for capacity building and career growth, few Africans have taken to big-team science, as evidenced by Africans' gross underrepresentation in popular, mostly Western-centric big-team science studies (Adetula et al., 2022). When African researchers

do participate, they are typically relegated to data collection roles. For Africans to embrace big team science and maximize its benefits, there is a need to develop and nurture the required expertise to do the job and include African thoughts and priorities. These are the central objectives of the ManyLabs Africa initiative.

Building research capacity via the ManyLabs Africa initiative

The ManyLabs Africa initiative consists of two big team science studies, which aim to address the limitations of previous big team science studies. The first study in the ManyLabs Africa initiative is a preparatory study that familiarizes African collaborators with the structure and aims of a big team science replication study, involves them in explicit capacity-building activities, and enables the team to identify the pitfalls and challenges of such a study on the African continent. This first preparatory study paved the way for the second study in the initiative, a larger, more extensive cross-cultural big team science test of African-origin claims. Together, these studies 1) build the capacity of African collaborators to use available credibility-enhancing research tools and the intellectual and digital infrastructures freely, 2) center more strongly on the issues and priorities of Africans, and 3) explore the richness of African behavioral science, thereby attempting to improve the participation of Africans.

> **[M]ost research questions that are currently investigated in big team science research Western-centric with little or no relevance to Africans. Big team science, therefore, runs the risk of supercharging Western dominance and missing out on crucial variations, research questions, and methods in African behavioral sciences. By laying the groundwork for participating in such initiatives as big team science, improving research capacities in Africa will optimize research for African development and promote synergistic cooperation between African and non-African countries.**

Our preparatory study was termed CREP Africa after our partner in the study, the Collaborative Replication and Education Project (CREP, https://osf.io/wfc6u/). CREP is an open-science-driven big team science collaborative designed to train users in credibility-enhancing tools and practices. CREP projects rely heavily on experiential learning, whereby learners come to understand credibility-enhancing tools and practices by using them in an actual research project. CREP Africa uses the CREP collaborative structure to conduct a so-called 'Registered Replication Report,' an article format hosted by some journals (the complete list of which is at https://cos.io/rr) where authors submit a protocol for a replication study before collecting data. At Stage 1, the protocol is peer reviewed and revised if necessary, and if the authors satisfy the reviewers that their protocol is sound, accepted in principle for in-principle publication. In Stage 2, after the authors have collected, analyzed, and interpreted the data, the finished article is peer-reviewed again to ensure the authors have adhered to their protocol. If it has, the article is accepted for publication – regardless

of whether the finished article reports statistically significant results. In CREP Africa, we leveraged this format to conduct a training-focused, Africa-wide replication study to test the replicability and generalizability to Africa of a non-African claim of interest to our African collaborators (our Stage 1 reviewed manuscript is available at https://osf.io/hdrf6).

To prepare to conduct this study, we consulted with African collaborators on available resources and infrastructure to assess their readiness for the studies. We found that although African collaborators are interested and reported that the CREP study is feasible, most labs lack the funds and facilities to conduct the study. We also developed an extensive bank of training materials on the use of open science practices (available for free at https://osf.io/8akz5/ and https://osf.io/b2pz6/). We provided direct training on these practices to twenty-three of our collaborators. To find possible studies to replicate, we adopted a 'top-down' study selection process where we, the study leads, found candidate studies that met a list of pre-set criteria. To incorporate the preferences of our African collaborators, we surveyed them as to their preferences for each of the studies on our list. Studies that our collaborators judged to be personally interesting, feasible to conduct, applicable and adaptable to their settings, and ethically acceptable were eligible for selection. After conducting our search, we selected Rottman's & Young's (2019) studies on moral transgression. We replicated these studies in twelve sites across five African countries, specifically Burkina Faso, Kenya, Morocco, Nigeria, and Tanzania. CREP Africa is being finalized for Stage 2 submission. Overall, CREP Africa has served as a training and preparatory project that familiarized African collaborators with big team science and a wide variety of open science practices and allowed us to identify pitfalls to prepare collaborators for the second study, the ManyLabs Africa study.

The ManyLabs Africa study is a transnational multisite online preregistered replication of three effects discovered initially in Africa to determine whether these effects replicate and generalize to European, North American, and African populations. By focusing on African-discovered claims, we can test the generalizability of these claims beyond – not to, as is almost always the case outside of our project – Africa, as well as provide a reference for why and how behavioral sciences should expand African claims via replication. To ensure that the selected studies presented claims that were Africa-discovered, reproducible, feasible, applicable, relevant, ethically acceptable, and of interest to researchers and targeted communities, we first consulted with Global South researchers (these consultations are described in full at https://osf.io/9qjw3). Subsequently, we adopted a four-step bottom-up study selection approach that heavily relied on our African colleagues to nominate actual studies or determine a study focus, as opposed to the top-down approach used in our preparatory CREP Africa study. The study selection procedure consisted of the following four steps (see https://osf.io/v6j28 for more details):

Step 1: African collaborators nominated research subjects and specific studies.
Step 2: We searched, evaluated, and shortlisted five studies.
Step 3: Both African and non-African collaborators assessed the shortlisted studies.
Step 4: We selected the final three studies from the shortlist based on feasibility and other considerations, strongly prioritizing African-preferred studies.

The claims selected for replication were the following:

1. Mgbokwere et al. (2015) claim that Nigerian parents have mixed attitudes towards teenage pregnancy.
2. Teye-Kwadjo et al. (2018) claim that male Ghanaian adolescents held more positive attitudes towards, greater feelings of control about, and a higher level of self-reported condom use than their female counterparts.
3. Vera Cruz (2018) claims that Mozambican women were more willing to forgive a husband involved in a less emotionally involved extramarital affair than a more involved affair and when regret expressed by a husband is high (versus low).

Presently, we are collecting data in forty-six sites across Africa (Burkina Faso, Gabon, Morocco, Nigeria, Tanzania, Zambia), Europe (Bulgaria, Denmark, Germany, Ireland, Italy, Netherlands, Poland, Portugal, Serbia, Slovakia, Slovenia, Turkey, United Kingdom), and North America (Canada, United States of America). For the full report on the ManyLabs Africa study protocol, see the study preregistration at https://osf.io/9m5qw.

How our African collaborators benefitted from the ManyLabs Africa initiative

Across the CREP African and ManyLabs Africa studies, our collaborators have attended open science workshops, nominated studies and articles as well as assessed the shortlisted studies for the final selection, culturally adapted and translated the measure from English into five African languages (Arabic, Chichewa, French, Swahili, and Portuguese), obtained local ethics clearances, coordinated their local labs to collect data, read and reviewed the studies' manuscripts, and managed a page on the Open Science Framework website that curates their site-specific materials. We believe these activities have built capacity among our collaborators and equipped them with appropriate expertise and networks for further contributions. Our local coordinators are now familiar with big team science study demands and operations and have perhaps grown their visibility as big team scientists in coauthoring these studies. They are more confident using the open science framework and working with the CREP platform, which may inspire more CREP projects in Africa. Our African collaborators are better equipped to prepare study materials by selecting a study for replication, reporting preregistered studies and closely related Registered Reports, translating and culturally adapting the measures, managing ethics reviews, and conducting online data collection. Nevertheless, leading a big team science project would require consistent participation, improved capacity and resources, and a more comprehensive network and multiple consultations.

However, throughout conducting these studies, we encountered substantial barriers, including a lot of collaborators dropping out of the project, lengthy time spent on completing tasks as well as uncompleted tasks, such as measure translation, ethics clearance and data collection, inadequate facilities, and lack of funds to pay participants and purchase necessary tools. In the remaining sections, I describe those barriers, our attempts to overcome them, and the lessons learned from these attempts. I conclude with general recommendations about how best to conduct big team science in Africa.

What to know when conducting big team science in Africa

CREP-Africa was the first rollout of a study from the CREP collaborative in Africa, which provided our African collaborators with the opportunity to improve their research capacity in open science practices. ManyLabs Africa is one of the only studies replicating African claims across Africa, Europe, and North America on the scale of big team science. These studies entailed some characteristics, such as the bottom-up study selection for a replication study that could help improve research workflows. Owing to the challenges and lessons learned from these studies, we recommend a few practical solutions to questions on building research capacity, understanding the variations and methods in African research, and improving African contribution globally.

Know how to navigate African literature

Finding African studies to replicate that met all our search criteria was challenging (see a full report on this at https://osf.io/v6j28). Some notable issues that limited our study search and nomination include limited access to African literature due to the poor visibility of African journals and the fact that most African researchers' works are not published in mainstream international journals. Furthermore, most African studies used qualitative methods, such as focus groups and interviews, which did not fit our predetermined criteria for suitable design and parameters for a large-scale online quantitative replication study. Another problem we encountered was that many African studies used Western-developed measures. Although these measures may apply to African settings, relevance cannot be assumed without evidence. We, therefore, avoided studies that used Western-developed measures, resulting in a smaller pool of studies that used African-developed measures or self-developed designs to consider for nomination. Lastly, non-English publication languages are standard practice in Africa. About thirty African countries communicate their scientific publication research to a large extent in Arabic or French, languages that we have not mastered.

To deal with these limitations, African researchers can take advantage of the open-access revolution to improve their visibility and access to their research via preprint servers and open-access journals. Researchers interested in African studies can search journals and repositories (for example, the *African Journals Online*, https://www.ajol.info/index.php/ajol; and *African Digital Research Repositories*, https://www.internationalafricaninstitute.org/repositories) and may find it easier to focus on qualitative research.

Mitigate poor online data collection rate

The contributing sites in Africa needed help in administering the studies and meeting the targeted sample sizes. It took about six months to attain the minimum sample size per site for the CREP Africa study, and only 1320 out of roughly 4700 respondents had above 85% completion rate to allow us to apply our exclusion criteria. To avoid these poor response rates in ManyLabs Africa, we selected studies with relatively simple procedures and few measures so that participants were required to spend less time completing the survey.

Also, we required at least a 40% completion rate to include a response. Even so, only 1169 (66%) out of 1785 African participants had a 40% completion rate compared to 4872 (87%) out of 5592 European and North American participants. Some African participants complained about the lengthy survey and the need for an internet subscription to complete the study. The African local leads observed that online data collection barriers could lead to an extended data collection period and a high rate of uncompleted responses. These barriers include intermittent power supply, lack of workspaces, the collection of data without compensation for time spent by participants and local coordinators, and lastly, the internet is unaffordable, and the internet quality is poor.

Moreover, internet penetration in most countries (Tanzania, 31.9%; Zambia, 31.2%; and Burkina Faso, 19.9%) where our African sites are located is low and even below the African average of 43% (Galal, 2024). Researchers sampling African populations via online survey should thus check for data quality by, for example, excluding participants who spent little time completing the study or failed attention tasks. It is also good practice to recruit a large sample of participants to accommodate exclusion. We recommend including an additional 35% of the target sample size. The tradeoff of excluding so many participants is that the sample could become unrepresentative of the target population, an issue that must be handled with care (Kennedy et al., 2020). Lastly, I recommend providing sufficient compensation to participants, especially for a lengthy study.

Know the realities of your African collaborators for effective collaboration

The expenses of a typical large-scale online replication study include participant compensation, internet access subscription, research assistants, ethics review fees, material translation, and the cost of instruments and tools. Only a few of our African collaborators and labs had the resources and facilities required to shoulder these costs in our studies (see the full report at https://osf.io/gds7b). This meant that we could not move forward with our two studies with some labs. Our experience with the costs for the ethics review is a case in point here. While our African collaborators paid review fees of $15, $50, and $100 for Nigerian, Kenyan, and Malawian ethics reviews, respectively, our collaborators were unable to afford the costs of the Tanzanian ($350) and the South African review fees ($917). Although ethics reviews in a few African countries, such as Egypt, do not require an application fee, it is ironic that, in some African countries, researchers who are poorly funded are also burdened with high ethics review costs. Thus, I recommend that African

research-regulating institutions, such as review boards, provide services at little or no cost to African researchers. Governments, NGOs, scientific research funders, wealthier scientific societies, and foreign partners can pay for these subsidized services.

There is also a need for better infrastructure and facilities, such as a computer and access to a stable and affordable internet connection, the lack of which frustrates our African collaborators' efforts, blocks their access to cutting-edge research tools, and reproduces poor working conditions. In addition to a lack of facilities, some of our African collaborators struggled with heavy teaching and research project supervision workloads (see also Naidoo-Chetty & du Plessis, 2021). We observed that African researchers take comparatively longer to complete some tasks, such as ethics application, data collection, and translation, compared to their European and North American counterparts. For instance, while most sites from the United States and Europe met their data collection target within two weeks, it took many African sites two months. We think that this is partly due to the poor infrastructure and tools available to our African collaborators. Hence we recommend that wealthier partners and funders can provide small grants to purchase needed materials, such as internet subscriptions, pay participant compensation, as well as the project lead or a designated person or support unit for the project to assist African collaborators who might require tools or help to complete a task.

In addition, despite primarily focusing on African populations and giving them a generous six-month period for data collection, we only managed to recruit collaborators from five out of fifty-four African countries for CREP Africa and six out of fifty-four for ManyLabs Africa. Moreover, due to both a lack of funding and the need to access the internet to deliver our study, we were unable to recruit large subsections of the countries we recruited from, such as people from rural areas. Thus, while the ManyLabs Africa initiative substantially improved upon the dismal African participation in much of psychology and big team science (Buchanan et al., 2023; Thalmayer et al., 2021), the heterogeneity among our African samples was low. For big team science interested in African populations, it is important to set feasible target sample sizes and realistic time frames in light of local realities and to provide the resources (for example, for participant compensation) to recruit diverse samples and increase the participation rates of Africans.

Lastly, managing a big team science study can be challenging and insufficient for an African researcher's career growth, a barrier not unique to Africans (Forscher et al., 2022). Some researchers we contacted and who were interested in the studies, for instance, told us that they would instead conduct single-authored small-scale studies or contribute to studies co-authored by a few scientists instead of contributing to multi-authored studies that potentially have a less positive impact on their careers. Some African research-regulating agencies, such as ethics review boards, as well as potential collaborators, questioned the effective management, relevance, and benefits of these large-scale online replication investigations. Unfortunately, big team science as it is currently constituted implicitly imposes these costs

and more work on African research labs when they have little or no funding and limited work time to spare. As individual researchers, we can do little about funding and providing facilities except to make aware of the dire conditions we face on the ground. However, understanding these realities helps us understand better how to collaborate with African researchers. Big-team scientists from the Global North who work with African collaborators should communicate well-specified and feasible tasks and contributions while being acutely aware of and considering the resources available to African labs and researchers.

Conclusion: what actions and changes are needed for big team science in Africa?

The ManyLabs Africa initiative provides an example of engaging Africans on their terms while enhancing credibility, improving participation, and building research capacity on the continent. Our experience revealed the need for all stakeholders to integrate and foster big team science (replication) studies. Researchers should familiarize themselves with big team science to build their capacity and network, access freely available tools and expertise, and advance their theories and impact globally. Beyond data collection typical of African labs' contribution, collaborators should support African collaborators in contributing to the development of measures, the adaptation and translation of materials, the preparation of applications to ethical review boards, and the leadership of big team science collaborations. As the research methods and processes continue to evolve to enhance credibility, African professors should review research methods curricula to include courses and subjects on open science practice and credibility-enhancing tools to nurture African scholars for the future. Because big team science studies are resource-intensive, it is crucial for scientific societies, publishers, and funders to review their policy and legal frameworks to make big team science-enabling structural adjustments for required support and credits to Africans involved in big team science. If at least some of these changes are made, we can continue to equip African researchers with resources and skills to produce quality research for local development and improve global participation.

References

Adetula, A., Forscher, P., Basnight-Brown, D.M., Azouaghe, S., & IJzerman, H. 2022. Psychology should generalize from – not just to – Africa. *Nature Reviews Psychology* 1(7), 370-1. https://doi.org/10.1038/s44159-022-00070-y

Buchanan, E.M., Lewis, S., Paris, B., Forscher, P.S., Pavlacic, J.M., Beshears, J., ... & Primbs, M. 2023. PSACR: The Psychological Science Accelerator's COVID-19 rapid-response dataset. *Scientific Data* 10(1). https://doi.org/10.1038/s41597-022-01811-7

Chelwa, G. 2021. Does economics have an 'Africa problem'? *Economy and Society* 50(1), 78-99. https://doi.org/10.1080/03085147.2021.1841933

Forscher, P. S., Basnight-Brown, D. M., Dutra, N., Adetula, A., Silan, M., & IJzerman, H. 2021 (December 29). Psychological science needs the entire globe, part 3: Does psychological science deserve brilliant researchers from outside North America and Europe? *APS Observer* 35(1). https://www.psychologicalscience.org/observer/entire-globe-part-3

Forscher, P.S., Wagenmakers, E.-J., Coles, N.A., Silan, M.A., Dutra, N., Basnight-Brown, D., & IJzerman, H. 2022. The benefits, barriers, and risks of Big-Team Science. *Perspectives on Psychological Science* 18(3), 607-23. https://doi.org/10.1177/17456916221082970

Galal, S. 2024 (March 14). Internet penetration in Africa January 2024, by country. *Statista*. https://www.statista.com/statistics/1124283/internet-penetration-in-africa-by-country/

Kennedy, C. Hatley, N., Lau, A., Mercer, A., Keeter, S., Ferno, J., & Asare-Marfo, D. 2020 (February 8). *Assessing the risks to online polls from bogus respondents*. Pew Research Center. https://www.pewresearch.org/methods/2020/02/18/assessing-the-risks-to-online-polls-from-bogus-respondents/

Kigotho, W. 2021 (June 17). Countries spend less than 1% of GDP on research. *University World News: Africa Edition*. https://www.universityworldnews.com/post.php?story=20210616151534847

Korbmacher, M., Azevedo, F., Pennington, C.R., Hartmann, H., Pownall, M., Schmidt, K., ... & Evans, T. 2023. The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology* 1. https://doi.org/10.1038/s44271-023-00003-2

Liverpool, L. 2021. Researchers from global south under-represented in development research. *Nature*. https://www.nature.com/articles/d41586-021-02549-9

Mgbokwere, D.O., Ekpoanwan, E.E., & Uyana, D.A. 2015. Perception and attitudes of parents towards teenage pregnancy in a rural community of Cross River State, Nigeria. *Global Journal of Pure and Applied Sciences* 21, 181-90- http://dx.doi.org/10.4314/gjpas.v21i2.10

Naidoo-Chetty, M., & du Plessis, M. 2021. Job demands and job resources of academics in higher education. *Frontiers in Psychology* 12. doi: 10.3389/fpsyg.2021.631171

Teye-Kwadjo, E., Kagee, A., & Swart, H. 2018. Condom use negotiation among high school adolescents in Ghana: the role of gender. *South African Journal of Psychology* 48(4), 465-75. https://doi.org/10.1177/0081246317730282

Thalmayer, A.G., Toscanelli, C., & Arnett, J.J. 2021. The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist* 76(1), 116-29. https://doi.org/10.1037/amp0000622

Vera Cruz, G. 2018. Would Mozambican women really 'tolerate' their husbands' extramarital sexual relationships as socially recommended? *Sexuality & Culture* 22, 1263-78. https://doi.org/10.1007/s12119-018-9520-8

The gender data gap in development policy research

Tommie Yeo Thompson (Charter Cities Institute), Winnie Mughogho (Queen Mary University of London) and Anisha Singh (London School of Economics and Political Science)

Thus, humanity is male, and man defines woman not in herself, but as relative to him.
Simone de Beauvoir, *The Second Sex*

The masculine research process

Women are persistently underrepresented in empirical research. In a world increasingly defined by evidence-based decision-making, the gender imbalance in research – the so-called 'gender data gap' (Perez, 2019) – manifests as policies and products optimized for men and hostile to women. Such disparities and consequences are seen in almost every research domain. In biomedical research, for example, the lack of gender representation has led to elevated risks of adverse treatment effects and the underdiagnosis of diseases in women (Martin et al., 1998; Dusenbery, 2018). Consumer products designed for men often disadvantage and even harm women. For instance, vehicle safety research reliant on male crash data has resulted in cars optimized to protect men. Consequently, women in the US are 17% more likely to die in a car crash (Perez, 2019; National Highway Traffic Safety Administration, 2013).

The lack of gender parity is not just a product of policy implementation and gendered barriers to access. Rather, the gender data gap is also a consequence of the research process: the questions we ask, the problems we prioritize, and the methods we practice. Feminist scholars have long highlighted the inherent difficulty of conducting objective social research free from the biases of its researchers and the power dynamics of the social context. In most cases, this means research that takes on a masculine default. That is, the presumption that masculine norms, behaviors, and bodies are neutral and universal (Cheryan & Markus, 2020; Holdcroft, 2007; Harding & Norberg, 2005).

The gender data gap is of particular concern for social and behavioral research in the Global South. Many countries exhibit patriarchal gender norms that might interact with research protocols, intervention design, and policy implementation. To varying degrees, these norms

confine women to the domestic sphere and impose time constraints, which makes them less accessible to researchers. This inadvertently creates an incentive to overrepresent men in development research and perpetuates a masculine model of human behavior. Standard social and behavioral research tools are also predominantly designed for WEIRD (Western, highly Educated, Industrialized, Rich, and Developed) contexts (Arnett, 2008). While this creates problems for generalizing policy prescriptions, it also means the methodological assumptions of standard tools do not adequately account for gender dynamics in diverse local contexts. This is not only true for generalizing Western-derived research tools to the Global South, but also for generalizing within the Global South. As a result, behavioral and social research may misrepresent and mismeasure female perspectives.

Tackling gendered research in the Global South

Gendered research practices in development reflect many factors without any 'silver bullet' solutions. As a starting point, the development community has focused on addressing gender imbalances in project funding, implementation, and outcomes. Grant organizations like the US Agency for International Development and Millennium Challenge Corporation have set funding requirements to include gender-based analysis in their programs, and research organizations have developed tools and guides to increase female participation in programs.

However, we have paid relatively less attention to the day-to-day research process itself, and we lack meaningful and actionable guidance on how to approach gender-inclusive research in even the most common research environments in the Global South. Solutions focused on gender equity at implementation will not address the implicit drivers of gender bias prevalent in the 'creative' and unstructured stages of research. To make research approaches more inclusive of women's perspectives, development researchers and behavioral scientists must introspectively interrogate their methods to identify what drives their bias. We identify two primary drivers of gendered research in the research process:

1. Sampling Bias: how do we ensure women can be accessed as research participants and that their voices are accurately included in all stages, from research to policy?
2. Measurement Bias: how do we design research tools (for example, surveys, experiments, psychosocial instruments, etc.) that are better at capturing the behaviors and beliefs of women?

Sampling bias

Gendered sampling bias arises when research protocols do not access women participants. This can come from using inappropriate data collection tools or an inattention to gender norms. For instance, women are 21% less likely to own a cellphone and 25% less likely to have access to the internet than men globally (Buvinic et al., 2014). Hersh et al. (2021) similarly found that male 'control' over household cell phones makes it challenging to administer surveys to women. Based on metadata from household surveys in India, men

picked up the phone as much as 71% of the time but only passed it to women between 7.5 and 11% of the time.

It can also be more costly for women to participate in studies than men. For example, women in the Global South face a 'time tax,' in which they spend more combined time on paid and unpaid labor than men (Res & Swaminathan, 2006). Unpaid labor, such as childcare, is also typically unbounded and unpredictable. These barriers create scheduling conflicts and cognitive costs to in-person research participation. Likewise, women may also require approval from their families to participate in studies, raising the personal risks.
We faced these recruitment problems firsthand during our time at Busara – a behavioral research firm. Relative to Kenya, where Busara is headquartered, women in India are less likely to receive permission to travel to research study sites. To accommodate this difference, we built our India practice around a set of 'mobile labs' that could reduce the distance women had to travel. In Lagos, we found through qualitative interviews that women may face relatively higher time poverty than those in Nairobi, which demands different culturally sensitive lab accommodations (Shipow & Singh, 2020).

Unfortunately, while it is widely acknowledged that sampling protocols can perpetuate gender imbalances, our understanding of these dynamics and their solutions is far from comprehensive. Instead, they come primarily from professional experience. They are typically disseminated within the research community through informal channels (e.g., Arandara & Gunasekera, 2019) rather than from intentional and formal research insights.

> **Many countries exhibit patriarchal gender norms that might interact with research protocols, intervention design, and policy implementation. To varying degrees, these norms confine women to the domestic sphere and impose time constraints, which makes them less accessible to researchers. This inadvertently creates an incentive to overrepresent men in development research and perpetuates a masculine model of human behavior.**

Measurement bias

Empirical social and behavioral research aims to accurately measure an underlying theoretical construct. These constructs, however, are often difficult to operationalize into concrete data collection tools. Designing these tools is unavoidably subjective and gendered, which can lead to measurement error and the reinforcements of male perspectives even when women are well-represented in study samples. For instance, research into women's autonomy has traditionally assumed that household members of any gender can communicate an objective snapshot of intra-household power dynamics. Status quo research approaches measure women's empowerment by asking a set of fixed questions to a single respondent,

usually the male head of the household. However, there is evidence that intra-household roles are highly contested between genders (World Bank, 2021) and that the direction of disagreement is not straightforward. Becker et al. (2006) found that women in Guatemala perceived themselves as having less household power than their husbands reported. This contrasts with findings from Twyman et al. (2015) and Ambler et al. (2021), which show that women report having more power than their husbands believed.

These disagreements can have deep consequences for our interpretation of empirical relationships. For example, based on surveys in five Asian countries, Ghuman et al. (2006) observed that child mortality seemingly declined when women perceived themselves as having more autonomy, but increased when men in the household perceived women as having more autonomy. They explained this as men ascribing more responsibility to their wives for a child's death rather than as a direct relationship between female empowerment and child mortality.

Measurement bias can also come from decisions as innocuous as question ordering. Ambler, Herskowitz, and Maredia (2021) documented how survey ordering can bias data towards men. In many household roster surveys, heads of household are asked to list all the members of their household then to answer questions about each of them in the same order (e.g,. if they mention their son first, then they will answer questions about their son first). However, when taking lengthy surveys, respondents learn they can manipulate their responses to reduce their interview time. Ambler, Herskowitz, and Maredia found that respondents typically listed female members later than men (for instance, they discussed their daughters after their sons). Since respondents are more likely to manipulate responses the further they are in the survey, this bias led to more significant data loss for women household members. In their survey, which asked respondents about the livelihood activities of each household member, women were reported to engage in 3.1% fewer livelihood activities per position in the roster than we would expect, compared to only 1.4% for male members. In other words, standard methods to measure household labor activities lead to more than two times more data loss for women than men. As a fix, they recommend randomizing the household roster before asking member-specific questions.

While little attention has been given to gendered measurement bias in lab settings, we speculate that similar issues are likely present. Evidence shows that language can be interpreted differently across genders (Jakiela & Ozier, 2019; Cheryan & Markus, 2020). This dynamic could affect experiments in which an assumed neutral framing is in fact gendered. Similarly, many experimental economic games (e.g., dictator game) use male silhouettes to represent players anonymously. This 'masculine default' may impact how women participants behave. For example, Niederle and Vesterlund (2011) found that women are more willing to compete against other women than men. If experimental tools prime women to think they are in a masculine environment, that may interfere with their competitive behavior.

Case study: measuring domestic violence

Much of social science research relies on self-reported data. However, the reliability of this data comes into question when individuals have compelling reasons to conceal the truth. The subjects that captivate researchers and policymakers sometimes align with deeply personal or sensitive matters, leaving survey participants grappling with the decision to divulge or withhold crucial information. Fears of potential repercussions may cause respondents to hesitate, for instance, when faced with admitting to illicit activities such as drug use or acknowledging experiences of domestic violence. Of course, these decisions often overlap with gender dynamics. Consider the dilemma faced by someone who asked, 'Have you ever been hit by your partner?' A truthful response could bear consequences for the other person, the relationship with the partner, or personal image. For instance, a woman financially dependent on her husband may fear that confirming such an experience could put him in trouble and jeopardize her source of livelihood. Or perhaps she would fear being punished by her husband were he to find out what she reported, regardless of how much the researcher guarantees anonymity.

Navigating these challenges becomes paramount in accurately gauging the prevalence of sensitive issues. The ramifications of underestimating the extent of problems like intimate partner violence (IPV) can be profound, leading authorities to misallocate resources and underestimate the severity of threats. We, therefore, set out to explore innovative approaches that balance participant concerns for their privacy with the need for precise measurements. For a study on intimate partner violence in the low-income areas of Kibera and Kawangware in Nairobi, we sought to evaluate whether an indirect measurement method, specifically the 'List Experiment', could capture more reliable estimates of the rate of intimate partner violence than direct questioning. We also examined variations across gender, education, and marital status to understand how different groups respond to various questioning methods. We surveyed 977 people, which included 502 women and 475 men.

Contrary to the literature, which often suggests that indirect measurements yield significantly larger estimates than direct methods, we did not find statistically significant differences between the two questioning methods for the entire sample. Surprisingly, the prevalence estimates of intimate partner violence were similar whether we used the List Experiment or direct questioning. In this context, direct questioning[1] emerges as the preferred method, providing more precise estimates for understanding intimate partner violence.

As we delved into subgroups, women stood out for reporting a higher incidence of intimate

1 A list experiment is a survey technique designed to measure sensitive or socially undesirable opinions by presenting respondents with a list of items, only some of which they are asked to admit to supporting. By comparing the average responses of those exposed to the sensitive item with those who were not, researchers can estimate the prevalence of the sensitive opinion in a population while maintaining respondent privacy.

partner violence when directly questioned – 34% compared to the 13% reported by men. While it is known that women are more likely to be victims of intimate partner violence, we questioned whether societal expectations may influence the size of the gaps. For instance, women may be more fearful of directly implicating their partners in violence, and men may feel more reluctant to disclose their experiences directly because of a fear of being perceived as unmasculine. The List Experiment added a layer to our analysis, revealing that men estimated a higher rate of physical intimate partner violence (23%) when questioned indirectly (compared to 13% in direct questioning). Surprisingly, we did not find a statistical difference for women.

These findings underscore the critical role of measurement as an integral, not secondary, aspect of any project. Our initial assumption that directly inquiring about intimate partner violence is inherently problematic was challenged by the context-specific results of our study. Likewise, whereas we might expect women to be more hesitant to report IPV directly, we, in fact, found that men were more reluctant. Researchers must diligently assess the strengths and limitations of employing diverse measurement techniques, validating them across various contexts. Through this rigorous evaluation, policies grounded in reliable evidence can be formulated.

From awareness to action: prioritizing gender-inclusive research in development

The gender data gap in development policy research underscores the urgent need for a comprehensive reevaluation of research practices. The masculine default inherent in the research process perpetuates biased outcomes and hinders the creation of policies inclusive of diverse perspectives, particularly in the Global South. To close this gap, we must go beyond addressing gender imbalances at the implementation stage and focus on introspective interrogation of research methods. Researchers and policymakers must actively refine sampling protocols and measurement tools, acknowledging the unique challenges faced by women in diverse contexts.

The call to action is clear: prioritize gender-inclusive research practices. This involves ensuring women's accessibility as research participants, exploring innovative recruitment methods, and developing tools that accurately capture women's behaviors and beliefs. Collaboration within the development community is crucial to refining and expanding our understanding of gender-inclusive research, paving the way for evidence-based policies that truly reflect the diversity of human experiences. The time for change is now, and it starts with a collective commitment to more inclusive and equitable research processes.

References

Ambler, K., Doss, C., Kieran, C., & Passarelli, S. 2021. He says, she says: spousal disagreement in survey measures of bargaining power. *Economic Development and Cultural Change* 69(2), 765-88. https://doi.org/10.1086/703082.

Ambler, K., Herskowitz, S., & Maredia, M.K. 2021. Are we done yet? Response fatigue and rural livelihoods. *Journal of Development Economics* 102736, 102736. https://doi.org/10.1016/j.jdeveco.2021.102736.

Arandara, R., & Gunasekara, S. 2019. Ensuring women's voices are heard in financial inclusion. *World Bank Blog*. https://blogs.worldbank.org/psd/ensuring-womens-voices-are-heard-financial-inclusion.

Arnett, J. J. 2008. The neglected 95%: why American psychology needs to become less American. *American Psychologist* 63(7), 602-14. https://doi.org/10.1037/0003-066x.63.7.602

Becker, S., Fonseca-Becker, F., & Schenck-Yglesias, C. 2006. Husbands' and wives' reports of women's decision-making power in western Guatemala and their effects on preventive health behaviors. *Social Science & Medicine* 62(9), 2313-26. https://doi.org/10.1016/j.socscimed.2005.10.006.

Buvinic, M., Furst-Nichols, R., & Koolwal, G. 2014. *Mapping gender data gaps*. Data2x. https://data2x.org/resource-center/mapping-gender-data-gaps/.

Cheryan, S., & Markus, H.R. 2020. Masculine defaults: identifying and mitigating hidden cultural biases. *Psychological Review* 127(6), 1022-52. http://dx.doi.org/10.1037/rev0000209.

Dusenbery, M. 2018. *Doing harm: the truth about how bad medicine and lazy science leave women dismissed, misdiagnosed, and sick*. HarperOne.

Ghuman, S J., Lee, H. J., & Smith, H. L. 2006. Measurement of women's autonomy according to women and their husbands: Results from five Asian countries. *Social Science Research* 35(1), 1-28. https://doi.org/10.1016/j.ssresearch.2004.06.001.

Harding, S., & Norberg, K. 2005. New feminist approaches to social science methodologies: an introduction. *Signs* 30(4), 2009-15. https://doi.org/10.1086/428420.

Hersh, S., Nair, D., Komaragiri, P.B., & Adlakha, R.K. 2021. Patchy signals: capturing women's voices in mobile phone surveys of rural India. *BMK Global Health* 6(Suppl. 5), e005411. https://doi.org/10.1136/bmjgh-2021-005411.

Holdcroft, A. 2007. Gender bias in research: how does it affect evidence based medicine? *Journal of the Royal Society of Medicine* 100(1), 2-3. https://doi.org/10.1258%2Fjrsm.100.1.2.

Jakiela, P., & Ozier, O. 2019. Gendered language. *Center for Growth Development Working Paper, Nr. 500*. https://www.cgdev.org/sites/default/files/gendered-language.pdf.

Martin, R.M., Biswas, P.N., Fremantle, S.N., Pearce, G.L., & Mann, R.D. 1998. Age and sex distribution of suspected adverse drug reactions to newly marketed drugs in general practice in England: analysis of 48 cohort studies. *British Journal of Clinical Pharmacology* 46(5), 505-11. https://doi.org/10.1046%2Fj.1365-2125.1998.00817.x.

National Highway Traffic Safety Administration. 2013. *Injury vulnerability and effectiveness of occupant protection technologies for older occupants and women* (DOT HS 811 766). US Department of Transportation.

Niederle, M., & Vesterlund, L. 2011. Gender and competition. *Annual Review of Economics* 3,

601-30. https://doi.org/10.1146/annurev-economics-111809-125122.
Perez, C.C. 2019. *Invisible women: data bias in a world designed for men*. Abrams Press.
Res, A., & Swaminathan, H. 2006. Gender and time poverty in sub-Saharan Africa. In C. M. Blackden & Q. Woden (eds), *Gender, time-use, and poverty in sub-Saharan Africa*. World Bank, pp. 13-26. https://doi.org/10.1596/978-0-8213-6561-8.
Shipow, A., & Singh, A. 2020. Is your data inclusive? Optimizing results by eliminating the hidden costs of research participation. *The Busara Blog*. https://medium.com/busara-center-blog/is-your-data-inclusive-ddd59933f108.
Twyman, J., Useche, P., & Deere, C.D. 2015. Gendered perceptions of land ownership and agricultural decision-making in Ecuador: who are the farm managers? *Land Economics* 91(1), 479-500. https://doi.org/10.3368/le.91.3.479.
World Bank. 2021. *How to better capture women's agency through decision-making measures: five tips from research in the Philippines*. World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/799341619600129275/how-to-better-capture-women-s-agency-through-decision-making-measures-five-tips-from-research-in-the-philippines.

Assessing deliberative polling methodology for decolonizing and indigenizing research

Dennis Chirawurah (University for Development Studies, Tamale) and Niagia F. Santuah (Millar Institute for Transdisciplinary and Development Studies, Bolgatanga)

Research intends to uncover hidden truths, generate new knowledge as well as the evidence needed to clarify principles for further research (Stroupe, 2015; Blaikie, 2007), and provide the basis for sound public policy decision-making (Castro Laszlo & Laszlo, 2002). Over time, research has been dominated by scientists of the conceptual western worldview – which is secular, individualistic, mechanical, and leaning towards the natural sciences – to the extent that the outcomes of research largely reflect a particular perception of reality that is far from being universal (Millar, Santuah & Tengolzor, 2023). The conceptual western worldview gives its adherents the absolute certainty that their way of knowing is the only way to access the truth. For them, 'there is only one superior way of knowing. That is rational and scientific knowledge, considered universally applicable because it is based on rational theoretical concepts and robust research methods' (Haverkort & Reijntjes, 2010, p. 12).

Western science thus became aggressively expansive and intolerant of non-western epistemologies. Besides, researchers have been instruments of colonialism and continue to be instruments of neo-colonialism and dependency. And the reasons for this are not far-fetched. The functionalist researchers in the early twentieth century in Africa were funded directly by the colonial administration to look at tribal politics, the interrelationships between social institutions, and how they function because they needed to know how these institutions could be used to maintain colonial control within a framework of local custom. These self-styled researchers developed their models of cognate structures and politics, which guided the policy of Indirect Rule in British West and East Africa (Staniland, 2008). Today, foreign governments, universities, and multinational corporations continue to sponsor researchers and development interventionists whose work does not only disempower and alienate the local community but also fails to address the latter's felt needs. Thus, philosophical positions and personal opinions, masquerading as scientific evidence, continue to guide development thinking and programming. Over the last half-century, alternative voices and views have emerged that challenge the hegemonic ownership of research by scientists of the conceptual western worldview. The concepts of 'inclusion,'

'involvement,' and 'participation' are attempts by these 'outsiders' to increase their understanding of the main features and dynamics of indigenous knowledge and resilience systems and of 'how opportunities are identified, problems are defined and solved' (Millar, Haverkort, Shankar & Apusigah, 2012, p. 69).

Indeed, it has long been established that an understanding of the centrality of open, transparent and participatory processes in sustainable development helps us to design policies, strategies, and processes, that are more likely to lead to long-term economic growth and that reinforce the strengths of the processes themselves (Stiglitz, 2002, p. 164). The World Bank and other key institutions concur that 'development goals and strategies should be "owned" by the country, based on citizen participation in shaping them' (World Bank, 2003, p. xviii). Whereas the meta-discourse is to decolonize research and indigenize the philosophies that guide research, policymakers, program implementers, and development experts agree on the importance of involving the citizenry in the decisions that affect them but are grappling with how best to do this (Abelson et al., 2003). Several experts support the view that an active, engaged citizen – rather than the passive recipient of information – is the prescription of the day (Brunet-Jailly & Martin, 2010; Smith, 2009; Phillips & Orsini, 2002). Thus, the most urgent issue is to improve the methodological approaches that do not generate compelling evidence for public policy decision-making. These are considered the lowest hanging fruits. New techniques that emphasize two-way interaction between decision makers and the public, as well as deliberation among participants, have a better appeal (Abelson et al., 2003). This is where Deliberative Polling (DP) provides proof of concept for improving the quality of local-level decision-making using an inclusive and participatory research methodology.

> **Western science thus became aggressively expansive and intolerant of non-western epistemologies. Besides, researchers have been instruments of colonialism and continue to be instruments of neo-colonialism and dependency.**

Deliberative Polling concept

The premise of DP is that when policy options are important for a community, then public consultations about them should be representative of the population and thoughtfully based on the best information available (Fishkin, 2021). There is, therefore, a strong case for bringing together a good random sample through transparent and good conditions for considering the issues and arguments for and against various policy options. This is justified as most citizens, most of the time in most countries worldwide, expend little effort considering public policy questions in depth before reaching a decision (Chirawurah et al., 2019).

DP methodology

DP methodology offers certain advantages over other research methods, such as public consultation. Some people argue that the expert community, not the public, should make inputs into complex questions of policy, especially value-laden questions that pose trade-offs on issues of general concern. However, the distinctive issue of policy choices posing trade-offs is that the weight of the values and value-laden goals must be tested in context. Otherwise the public has to live with whatever values the experts happen to bring (Fishkin, 2021). Standard public opinion polls could be effective if conducted well, but the public may not be well informed about the complex issues at stake. This may lead them to answer questions randomly, probably because they do not wish to admit that they do not know (Converse, 1964). Thus, the outcomes may not reflect what the public would think if they were engaged in thinking about the issues before deciding. Survey experiments are also not appropriate, as experience shows that the weighing process is greatly improved by discussion. In-depth interviews only provide rich, individualized perspectives that do not represent the general population's opinion. Even 'focus groups cannot be used to represent opinion because they are too small to be statistically meaningful' (Chirawurah et al., 2019, p. 6).

> **Discussions under the baobab tree to reach a consensus on issues of collective importance have been the hallmark of how indigenous people conducted their affairs. Indigenous people are increasingly recognized as having tacit knowledge that holds sustainable answers to contemporary development challenges, which is the first rung in decolonizing and indigenizing research and its outcomes.**

DP application

The first DP in Africa was conducted in Uganda and replicated in Ghana and Senegal. Like in many other countries, the policy process does not adequately involve the communities in these three countries. Governments often use subjective assessments of situations to craft policies toward addressing the needs of the people. The challenge has always been how to consult the communities and seek their informed opinion in an unbiased way using an adequately representative sample. DP provides a method by which a representative community sample can be consulted in-depth on critical issues. It offers representative and informed opinion data, both quantitative and qualitative, about the public's views once they have had the chance to review different and competing options.

In Ghana, a scientific random sample of the Tamale Metropolitan Area was gathered for two days of face-to-face deliberation. The samples were recruited through random selection of households and random selection of participants within the households. A total of 243 persons were interviewed, and only two persons selected declined to take the initial survey. In all, 208 persons participated in the two days of face-to-face deliberations. This represented a response rate of over 85%, a very high level by world standards for surveys

and even more remarkable considering that the deliberations lasted two days. The sample was 48% male and 52% female, with an average age of 33.7 years, of which 27.9% had never been to school, and only 3.9% were first-degree holders.

A broad thematic area and key topics for the deliberation had previously been identified through an extensive literature review triangulated with results of focus groups and key informant interviews. The thematic area was how to deal with environmental disasters and the population pressures that challenge life in vulnerable communities of the Tamale Metropolitan Area. The two emerging topics were 1) Livelihood and food security, and 2) Water, sanitation and hygiene. A group of subject area specialists, NGO leaders, academic experts, and government officials then developed and vetted information kits around these two topics for the information of participants (deliberators). Participants were randomly assigned to groups of between eight and twelve participants to meet at a designated venue for face-to-face discussions under the guidance of a moderator. On the first day, participants discussed livelihood and food security, and on the second day, they debated water, sanitation, and hygiene. Given the low literacy rate of the population, a fifteen-minute video version of the information kits was produced and shown at the beginning of each day of deliberation.

Opinion changes

All of the policy proposals were rated before and after deliberation on a zero to ten scale, where zero is 'extremely unimportant,' ten is 'extremely important,' and five is 'undecided', following Fishkin's advice that '[W]ithout pre- and post-event opinion measurements at the individual level, there is only anecdotal evidence about whether or how opinions changed' (Fishkin, 2021, p. S22). The proposals were all rated highly before and after deliberation. All of them stayed on the 'important' side of the scale. In all, 28 of the 39 policy proposals (71.8%) showed statistically significant changes after deliberation. This is not surprising since they all focused on basic health and food security issues for a population facing severe challenges in both areas. Yet some of the proposals posed some hard choices. The briefings made clear that the public latrines and the areas for gardening were currently very much in the same places. On the map, they appeared to be in nearly identical locations. Given the scarcity of water and the fragility of food security, it is unsurprising that many people survive by raising food in gardens using wastewater. Hence, focusing on food security would have its cost in spreading diseases, especially cholera. However, a focus on health would require sacrifices in food security. This tradeoff was explored in the question: 'Some people think that vegetable farms should produce as much as possible, even if they have to use the wastewater from toilets (at point 0). Other people think that vegetables should only be produced with clean water, even if fewer vegetables are produced (at point ten).' Before deliberation, the support was already firmly on the clean water side of the trade-off (at a mean of 9.04 on the ten-point scale). After deliberation, it moved even further to 9.53 on the scale, a gain of nearly half a point and a very significant change ($p = .0004$). Participants were significantly more willing to emphasize clean water to avoid disease, even at the cost of food security.

This tradeoff is also reflected in the policy option 'Ban the use of untreated wastewater for gardening,' which increased from 8.53 on the ten-point scale to 9.09, an increase of more than half a point and a strongly significant change (p= .0004). This option shows the willingness to require a ban on the practice, so it is not merely a prescription but a proposal for a legal requirement.

Efficacy and expectations

Participants were asked before and after whether or not they thought anyone would make use of these results. When asked, 'How serious or not serious do you think the government will take into account your views and suggestions provided in this event?' participants' views on government seriousness increased significantly from 6.90 to 7.93, more than a full point on the scale (p= .000). And about their own opinions, on whether 'I have opinions about my community worth listening to' there was a significant increase from 8.30 to 8.76 (p=.023). These increases in external efficacy (whether the government or community will pay attention or respond to their opinions) and internal efficacy (whether their views are worth listening to) are impressive dialogue results. The deliberations seem to have considerably affected their sense of efficacy.

Event evaluation

When asked whether or not the DP event was a valuable use of their time, 99.5% of participants reported that the DP was valuable, with 88% saying that it was extremely valuable. Further, all participants felt the information kits were valuable; with 83% indicating that the briefing materials were extremely valuable. Therefore, it is highly unlikely that participants acted in 'anticipatory obedience,' hoping to continue relationships with the research team. In terms of participants' views of fellow participants, 99% agreed that they learned a lot about people who are different from them. According to Chirawurah et al. (2019), the Tamale Metropolitan Area DP results are among the strongest evaluations seen in DPs worldwide and show the high level with which the model was implemented from the standpoint of the participants. To a large extent, the DP method creates, maintains, reproduces, and successfully mediates hierarchies between researchers and research participants.

Experts have, however, raised concerns about 'anticipatory obedience' and 'politically correct' answers given by participants who may have 'performed' according to what they expect that the hierarchically superior 'researcher' wants to hear. However, working on the ground and observing how the research-researched relationship was structured enabled the DP process to avoid the pitfalls associated with conventional research approaches. Participants were first made to understand that their frank opinions were more valuable than giving 'politically correct' answers and that their responses would form the basis for initiating practical actions to address community resilience challenges around the discussion topics. Participants indicated that the researchers were consistent and that they had clarified the processes and followed them through. A man came on day two to stand in for his wife who

was said to have been taken ill after attending the first-day event, but the DP methodology did not allow him to be admitted. All this points to the robustness of the DP methodology. Nonetheless, participants were specifically asked to assess the deliberation process, the information kits used, how the moderators fared in moderating the deliberation, and whether the video documentary hindered or enabled the deliberation. The results were highly satisfactory.

Implications of decolonizing research and indigenizing methodologies

Discussions under the baobab tree to reach a consensus on issues of collective importance have been the hallmark of how indigenous people conducted their affairs. Indigenous people are increasingly recognized as having tacit knowledge that holds sustainable answers to contemporary development challenges, which is the first rung in decolonizing and indigenizing research and its outcomes. Though there is clearly the need to raise the number of indigenous experts to increase the momentum of the decolonization and indigenization process, we cannot afford to leave the task of doing the fieldwork to indigenous people holding PhDs from western universities.

In the process of acquiring a PhD from, say, Stanford, Cambridge, or any of the Ivy League universities, a research scientist is likely to be so indoctrinated in the traditions of the discipline that he or she emerges perceiving the world in the manner of its leading academics. The Tamale Metropolitan Area DP project also offers a good test case for determining whether quality research can be effectively conducted in a development context with low literacy levels. The results suggest a positive answer. Therefore, despite its imperfections and limitations, mainly of cost and convenience, the DP methodology, and evaluation of the research outcomes, provide a superior template for decolonizing research philosophies and indigenizing research methodologies for maximum impact.

References

Abelson, J., Forest, P.G., Eyles, J., Smith, P., Martin, E., & Gauvin, F.P. 2003. Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science & Medicine* 57(2), 239-51. https://doi.org/10.1016/s0277-9536(02)00343-x

Blaikie, N. 2007. *Approaches to social enquiry: advancing knowledge*. Polity.

Brunet-Jailly, E., & Martin, J.F. (eds). 2010. *Local government in a global world: Australia and Canada in comparative perspective*. University of Toronto Press.

Castro Laszlo, K. & Laszlo, A. 2002. Evolving knowledge for development: the role of knowledge management in a changing world. *Journal of Knowledge Management* 6(4), 400-12. https://doi.org/10.1108/13673270210440893

Chirawurah, D., Fishkin, J., Santuah, N., Siu, A., Bawah, A., Kranjac-Berisavljevic, G., & Giles, K. 2019. Deliberation for development: Ghana's first deliberative poll. *Journal of Public Deliberation* 15(1), Article 3. https://doi.org/10.16997/jdd.314

Converse, P. 1964. The nature of belief systems in mass publics. In D.E. Apter (ed) *Ideology*

and discontent. University of Michigan Press, pp. 206-61.
Fishkin, J.S. 2021. Deliberative public consultation via deliberative polling: criteria and methods. *Hastings Center Report* 51, S19-S24. https://doi.org/10.1002/hast.1316
Haverkort, B., & Reijntjes, C. 2010. Diversities of knowledge communities, their worldviews and sciences: on the challenges of their co-evolution. In S.M. Subramanian & B. Pisupati (eds), *Traditional knowledge in policy and practice: approaches to development and human well-being*, United Nations University Press, pp. 12-30.
Millar, D., Santuah, N.F., & Tengolzor, M.B. 2023. Cosmovisions and critical sustainability sciences: an African ontology of 'Vurr' (an energy) among the Dagara of Southwest Burkina Faso and Northwest Ghana. In S. Rist, P. Bottazzi & J. Jacobi (eds), *Critical sustainability sciences: intercultural and emancipatory perspectives* (1st ed.). Taylor & Francis, pp. 75-94.
Millar, D., Haverkort, B., Shankar, D., & Apusigah, A.A. 2012. *Learning together: developing inclusive knowledges and sciences towards operational methods for endogenous research, education and development*. Guidebook for CAPTURED. Vol. 2.
Phillips, S.D., & Orsini, M. 2002. *Mapping the links: citizen involvement in policy processes*. CPRN Discussion Paper.
Smith, G. 2009. *Democratic innovations: designing institutions for citizen participation*. Cambridge University Press.
Staniland M. 2008. *The lions of Dagbon: political change in Northern Ghana*. Cambridge University Press.
Stiglitz, J.E. 2002. Participation and development: perspectives from the comprehensive development paradigm. *Review of Development Economics* 6, 163-82. https://doi.org/10.1111/1467-9361.00148
Stroupe, D. 2015. Describing 'science practice' in learning settings. *Science Education* 99(6), 1033-40. https://doi.org/10.1002/sce.21191
World Bank. 2003. *Toward country-led development: a multi-partner evaluation of the comprehensive development framework--synthesis report*. Retrieved from https://openknowledge.worldbank.org/handle/10986/15080.

Evaluating interventions: a practical primer for specifying the smallest effect size of interest

Hannah K. Peetz (Radboud University, Nijmegen), Maximilian A. Primbs (Radboud University, Nijmegen), Leonie A. Dudda (University Medical Center Utrecht & Erasmus University Rotterdam), Pia K. Andresen (Utrecht University), Charlotte R. Pennington (Aston University, Birmingham), Samuel J. Westwood (King's College London) and Erin M. Buchanan (Harrisburg University of Science and Technology)

Effect sizes measure the magnitude of an experimental or intervention effect and thus represent one of the most important outcomes of quantitative empirical studies. However, many researchers experience difficulties in deciding what effect sizes are meaningful and how to interpret them. This chapter offers a much-needed practical primer on how effect sizes can be interpreted. Specifically, we focus on the smallest effect size of interest (SESOI) approach to determine whether the effects observed in studies are meaningful and provide real-world examples to demonstrate how this approach can be implemented.

Experimental research plays a crucial role in shaping policy decisions by providing valuable insights into the efficacy or effectiveness of various interventions, strategies, and initiatives. However, interpreting findings from these studies can be complex. You might, for example, want to test whether chocolate consumption increases self-reported happiness in the workplace. You, therefore, conduct a randomized controlled trial (RCT) in which employees are randomly assigned to receive free chocolate (experimental group) or no free chocolate (control group). To test whether your intervention significantly increases happiness in the experimental group compared to the control, you perform a statistical test (for example, a t-test), which provides a range of statistical results that require careful interpretation.

When using Null Hypothesis Significance Testing (NHST), your first go-to may be the *p*-value of the test that you have conducted, which provides information on the probability of observing the sample data or more extreme data, assuming that there is no effect (that is that the null hypothesis is true; Lakens, 2021). In our example above, you would use a chosen threshold, most commonly $p < .05$ in psychological science, to conclude that there is a statistically significant effect: the chocolate intervention significantly increased employee happiness. The *p*-value, however, does not tell you anything about the strength of the effect (even if your *p*-value is very small) and instead, to make conclusions about the strength of

your effect, you need to interpret the effect size. Effect sizes indicate the magnitude of an effect and, unlike *p*-values, can be used to interpret the practical significance of your results. To assess the magnitude of the chocolate intervention, you can calculate an effect size of group differences, known as Cohen's d, and there are many different kinds of effect sizes (for example, correlation coefficient r, odds ratio, etc.) dependent on the statistical test that you chose. Let us imagine that the effect size for the chocolate intervention was Cohen's $d = 0.15$. Now, consider how you would interpret this effect size. Is your chocolate intervention effective? Is the effect size meaningful?

A popular yet misguided option: benchmarks

One approach to interpreting effect sizes that is often applied is the categorization of effect sizes as small (e.g., $d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$). These absolute numerical cut-off values are based on Cohen's guidelines (1988) and are currently the most widely used approach to interpreting effect sizes in psychological science (cf, for example, Funder & Ozer, 2019; Mesquida et al., 2023). However, Cohen expressed that such benchmarks were 'recommended for use only when no better basis for estimating the index is available' (Cohen, 1988, p. 25). Indeed, scholars have explained that underlying such an approach is an unrealistic assumption that a single set of benchmarks can adequately classify effect sizes regardless of populations, interventions, time frames, or fields of study (see Anvari et al., 2023).

For example, our chocolate intervention's effect size is $d = 0.15$. Using Cohen's guidelines, you may conclude that this is a small effect on happiness, but what does small actually mean in this context? Does it mean that it is not worth eating some chocolate? The judgment of whether this effect size is meaningful and how it can be interpreted depends on the context. For example, if you are feeling down, a slight increase in happiness might already be meaningful for you. If you are already very happy, you might not notice this slight difference and would not want to risk gaining weight by eating chocolate. Based on such conventional benchmarks, the same 'small' effect size can therefore be interpreted as meaningful in one context and not meaningful in another. One size does not fit all, and effect sizes thus need to be interpreted in the wider context in which they are produced.

Alternative approaches: the smallest effect size of interest

Given the problems with Cohen's benchmarks, researchers have developed alternative approaches to effect size interpretation that are field and context-specific. Here, we wish to highlight one useful type of benchmark that is of particular importance to applied researchers: the smallest effect size of interest (SESOI). The SESOI is the smallest effect size that a researcher considers meaningful (see Lakens, 2017; Lakens et al., 2018). Therefore, anything smaller than this is not of interest, regardless of the statistical significance of the finding. If the observed effect is larger than the SESOI, then you can interpret the effect as being practically meaningful. The SESOI should be considered at the design stage of your study, thus determining the a-priori before testing your hypothesis. For example, you

could set your SESOI as Cohen's $d = 0.2$ before you test your intervention, so if chocolate improves happiness by this much or more, you would deem it practically meaningful. However, determining and justifying an SESOI, in this case, justifying how much of an increase in happiness is meaningful, can be complex. You may choose an SESOI based upon a quantifiable theoretical model and, ideally, such a theory would specify how the intervention works, for whom, when, where, to what degree, and so on. Unfortunately, theories in the social sciences are often not well-defined. Predictions based on these theories may indicate whether an effect exists, but rarely, if ever, do they indicate the size of the effect. Therefore, researchers often use effect sizes found in previous research or meta-analyses to guide the selection of an SESOI. Especially when testing an intervention, one could identify the effect size of existing interventions and gold standards to determine an SESOI. While this may tell you whether your intervention is more or less effective than previous interventions, interpreting the effectiveness of your intervention may still present a challenge, as effect sizes are commonly expressed as arbitrary statistical units.

To ease the interpretation of effect sizes, we present three alternative ways of specifying an SESOI in more useful units to applied researchers. These alternatives do not require much statistical expertise and allow applied researchers and laypeople alike to evaluate the effectiveness of interventions more readily. These methods are not mutually exclusive and can be combined. We want to emphasize that the choice of method very strongly depends on the research question and the effect the researchers are interested in.

1. ***Person as effect size.*** In most traditional approaches, mean scores are compared, such as in a t-test where we compare a group that consumes chocolate to a group that does not, as in the example above. The interpretation could then be that the intervention works 'on average.' However, the average will include some people who feel less happy than before eating chocolate, people who aren't affected at all, and some who experience an increase in happiness. The disadvantage of this approach is that you usually do not know how many people actually experience increased or decreased happiness. The 'person as effect size' approach does exactly this (Grice et al., 2020). This approach asks the simple question of how many people show an effect consistent with theoretical expectations. For instance, one could be interested in the number of people who benefit from a medical treatment. If less than 25 % show an increase in their health, this treatment might not be worth pursuing. Hence, the person-as-effect size approach gives researchers an easily calculable proportion of persons or units that show an effect in the expected or desired direction (see ibid. for details). The person-as-effect size is, therefore, particularly useful in specifying SESOIs if researchers are interested in questions of intervention effectiveness. For example, Sayette and colleagues (2022) applied the person-as-effect size approach to show that a commonly used method to induce craving in smokers indeed works for the majority of smokers.

2. ***Cost-benefit analysis.*** As the name suggests, a cost-benefit analysis approach weighs

an intervention's potential benefits against the potential costs (Riegg Cellini & Kee, 2015). For everything you spend – be it money, time, effort, or other resources – how much do you gain in return? For example, consider a grant program that aims to facilitate the growth of small businesses in Africa. You decide to give 100 businesses a starting grant of 10,000$, thereby spending 10,000,00$ in total. You might deem your program successful if the businesses make at least 10,000,000$ more profit. Often, cost-benefit analyses are described in terms of monetary value, but other types of resources can often be converted into an expected monetary equivalent. For instance, in many disciplines, such as economics, even the value of life is expressed as a monetary value. By using cost-benefit analyses, you pre-specify an SESOI and evaluate the effectiveness of your intervention in terms of the value produced relative to the associated costs – which, for many economic policies and grant or aid programs, is a more useful unit of measurement than abstract effect sizes. Using cost-benefit analysis to specify an SESOI is especially useful in settings with limited resources, as it allows us to evaluate the effectiveness of these limited resources. For example, Dopp and colleagues (2018) used a cost-benefit analysis to show the effectiveness of a large-scale multisystem therapy for juvenile offenders in New Mexico, concluding that each dollar spent resulted in approximately three dollars saved in future costs.

3. ***Minimal important differences.*** Finally, the minimal important difference approach estimates 'the smallest change in an outcome measure that individuals consider to be meaningful enough in their subjective experience such that they are willing to rate themselves as feeling different' (Anvari & Lakens, 2021, p. 3). This approach shifts the focus towards the subjective experience of the participants, the people whose life you want to change for the better. If you conduct a weight loss intervention, you can determine how much bodyweight reduction is necessary for patients to notice a change. This way, you consider participants' subjective experience when determining your SESOI. For example, Mao and colleagues (2021) used data from an intervention study on postnatal depression to determine a minimal clinically relevant difference for changes on the Edinburgh Postnatal Depression Scale (EPDS). Improvement of 4 points and worsening of 3 points on the EPDS were found to be clinically relevant, meaning that patients report a noticeable change in postnatal depression due to the intervention.

A practical example: migraine treatment

We now apply these approaches to a real-life study to determine the SESOI for migraine treatment to aid comprehension. Migraine is a common headache disorder that affects 10% or more of individuals worldwide (Walter, 2022). There are a large number of therapeutic options for patients with migraine, yet no clear diagnostic criteria for what constitutes a positive patient response to treatment. As migraine disorder affects patients' physical, work, social, and cognitive factors (Buchanan et al., 2023), it is important to define what level of change is meaningful to patients for clinical research to develop individualized treatment

plans. Given that migraine is defined by frequent head pain, it is tempting to define a positive response as a reduction in the number of migraine attacks and/or headaches. However, research has shown that patients desire both relief from head pain and improvement in other life aspects (Buchanan et al., 2023; Smith et al., 2019).

To assess the wide range of patient concerns, patients were given the Functional Assessment of Migraine Scale in a recent set of validation studies (FAMS; Buchanan et al., 2023; Buchanan et al., in prep). The FAMS is an assessment survey given to patients to gauge their perceived response to a migraine treatment plan. The survey was given at the start of treatment and used throughout treatment to determine changes within patient response across several concerns, including pain frequency, number of attacks, effects on social life and self-worth, and normal daily functioning.

Person as effect size. We can define improvement as a change in FAMS scores between measurement time points of at least one point. We may determine that at least 25% of patients need to show improvement in their assessment of treatment scores as our SESOI, given a known baseline in responding to different treatment options. Any percent over this value would be considered an important and practical change, thus, supporting the treatment protocol.

Cost-benefit analysis. For each patient who showed a positive change in FAMS scores, we can calculate the cost of their new migraine treatment and subtract the cost of their old migraine treatment. If the new treatment costs $10 more a year or less, then we would consider the treatment practical to benefit patients while keeping costs approximately the same or less.

Minimal important differences. In the previous two examples, we defined positive change as a one-point difference on our scale. However, defining change from the patient's perspective may be useful. To determine the smallest detectable effect from patients themselves, each patient was also given a simple question: 'In relation to your migraine management or treatment, how much has each of the following areas changed in the last month? Total: overall migraine treatment or management in the last month' with response options of 1 = much worse, 2 = little worse, 3 = same, 4 = little better, and 5 = much better. In Figure 1, patient's responses to this question are shown on the y-axis, and the change in their overall FAMS scores across time is shown on the x-axis. The results suggest that patients who self-report that they are much worse than during their previous measurement time also show a decline in their FAMS score compared to those who indicate they feel the same.

By taking the average patient-perception scores, we can determine our smallest negative and positive change of interest on the FAMS. In general, patients show a small negative change (i.e., a decrease in their response to treatment) when their scores on the FAMS decrease by two points but feel an improvement in their treatment when they increase their scores by five points. These cut-off scores can then be used in calculating person as effect sizes,

cost-benefit analysis, and clinical practice in discussions with a patient's provider or within a clinical trial to determine change from the current gold standard treatment – providing the researcher with a SESOI.

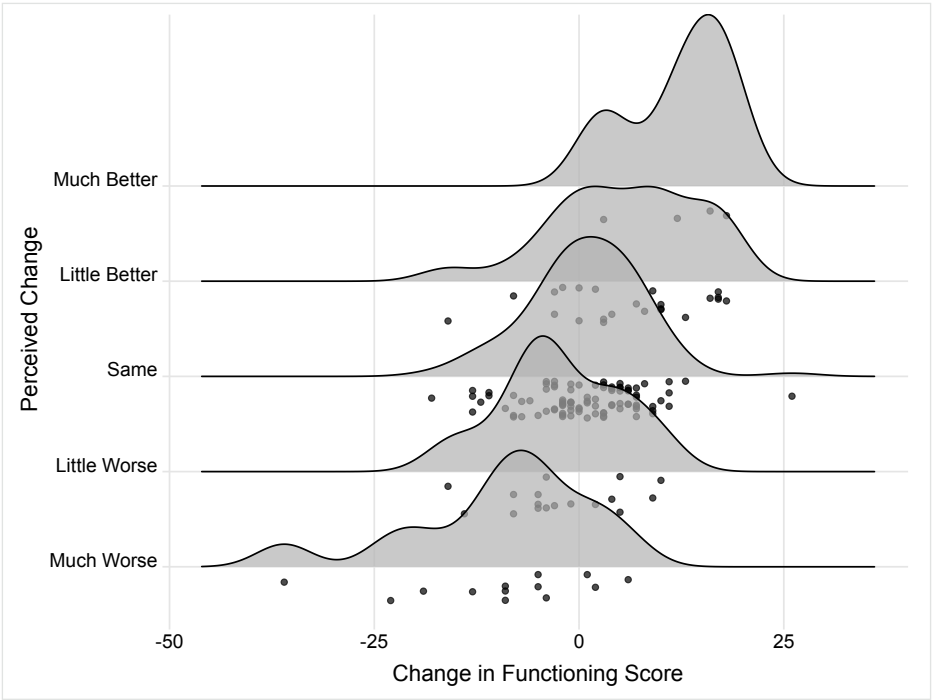

FIGURE 1. Patients rated the amount of change they perceived during their treatment for migraines, ranging from feeling much worse to much better. For each of those ratings, their actual change score for functioning measurement (i.e., current score minus baseline measurement) are shown on the x-axis. Patients who felt much worse had decreasing scores compared to patients who felt much better.

Conclusion

Researchers typically have to interpret effect sizes and other test statistics when conducting interventions or experimental studies. However, it is often unclear what a particular effect size means and whether the size of the effect is large enough to be meaningful. In this chapter, we presented the SESOI as a way to determine and interpret whether an effect is meaningful. We provided a practical primer through three intuitive ways to determine an SESOI: person as effect size, cost-benefit analyses, and the minimal important difference. These metrics do not require much statistical training to be understood: the number of people, businesses, or units that do X, the amount of value produced per unit of investment, or the smallest difference noticed by participants are as simple to calculate as they are intuitive. Moreover, specifying SESOIs reduces resource waste, which is particularly important in global development contexts where researchers work with limited resources. A SESOI should be determined a-priori in the design of a study to ensure maximum rigor and

transparency. We hope this chapter aids the development and planning of rigorous research in applied settings and facilitates the use of SESOIs in the broader scientific community.

References

Anvari, F., Kievit, R., Lakens, D., Pennington, C.R., Przybylski, A.K., Tiokhin, L., Wiernik, B. M., & Orben, A. 2023. Not all effects are indispensable: psychological science requires verifiable lines of reasoning for whether an effect matters. *Perspectives on Psychological Science* 18(2), 503-7. https://doi.org/10.1177/17456916221091565

Anvari, F. & Lakens, D. 2021. Using anchor-based methods to determine the smallest effect size of interest, *Journal of Experimental Social Psychology* 96, Article 104159, https://doi.org/10.1016/j.jesp.2021.104159.

Buchanan, E.M., Manley, H.R., Sly, J. S., Cady, R.J., Wikowsky, A., & Cunningham, A.L. 2023. Development of the functional assessment of migraine scale using a patient guided approach. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation* 32(3), 867-79. https://doi.org/10.1007/s11136-022-03279-9

Buchanan, E.M., Rodgers, L., Wikowsky, A., Sly, J., Manley, H., & Cady, R. (in preparation). The assessment of the FAMS scale and determination of cut off scores, https://osf.io/k7w46/

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Dopp, A. R., Coen, A.S., Smith, A.B., Reno, J., Bernstein, D.H., Kerns, S.E., & Altschul, D. 2018. Economic impact of the statewide implementation of an evidence-based treatment: multisystemic therapy in New Mexico. *Behavior Therapy* 49(4), 551-66. https://doi.org/10.1016/j.beth.2017.12.003

Funder, D.C., & Ozer, D.J. 2019. Evaluating effect size in psychological research: sense and nonsense. *Advances in Methods and Practices in Psychological Science* 2(2), 156-68. https://doi.org/10.1177/2515245919847202

Grice, J.W., Medellin, E., Jones, I., Horvath, S., McDaniel, H., O`Lansen, C., & Baker, M. 2020. Persons as effect sizes. *Advances in Methods and Practices in Psychological Science* 3(4), 443-55. https://doi.org/10.1177/2515245920922982

Lakens, D. 2017. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science* 8(4), 355-62. https://doi.org/10.1177/1948550617697177

Lakens, D. 2021. The practical alternative to the p value is the correctly used p value. *Perspectives on Psychological Science* 16(3), 639-48. https://doi.org/10.1177/1745691620958012

Lakens, D., Scheel, A.M., & Isager, P.M. 2018. Equivalence testing for psychological research: a tutorial. *Advances in Methods and Practices in Psychological Science* 1(2), 259-69. https://doi.org/10.1177/2515245918770963

Mao, F., Sun, Y., Wang, J., Huang, Y., Lu, Y., & Cao, F. 2021. Sensitivity to change and minimal clinically important difference of Edinburgh postnatal depression scale. *Asian Journal of Psychiatry* 66, Article 102873. https://doi.org/10.1016/j.ajp.2021.102873

Mesquida, C., Murphy, J., Lakens, D., & Warne, J. 2023. Publication bias, statistical power and reporting practices in the Journal of Sports Sciences: potential barriers to replicability.

Journal of Sports Sciences 41(16), 1507-17. https://doi.org/10.1080/02640414.2023.2269357

Riegg Cellini, S., & Kee, E.J. 2015. Cost-effectiveness and cost-benefit analysis. In K.E. Newcomer, H.P. Hatry, & J.S. Wholey (eds), *Handbook of practical program evaluation*. Jossey-Bass, pp. 636-72.

Sayette, M.A., Goodwin, M. E., Creswell, K.G., Esmacher, H.J., & Dimoff, J.D. 2022. A person-centered analysis of craving in smoking-cue-exposure research. *Clinical Psychological Science* 10(3), 570-83. https://doi.org/10.1177/21677026211032646

Smith, T.R., Sly, J.S., Martin, V.T., Kandel, E., & Cady, R.J. 2019. *Meaningful relief measurement: analysis of responses to a survey of patient priorities*. Presentation for the American Headache Society 61st Annual Scientific Meeting in Philadelphia, PA. https://doi.org/10.1111/head.13549

Walter, K. 2022. What is migraine? *JAMA* 327(1), 93. https://doi.org/10.1001/jama.2021.21857

Utility of meta-research for Global South policymaking: a reflection on education research

Anushka Ghosh (Busara)

Global development has increasingly shifted towards evidence-driven policymaking. This practice has been hailed and lauded as an effective way to tackle poverty and solve some of the most pressing issues governments face today, including education policy. Education is perceived as a public good that should be freely available to all citizens of a nation. However, despite best efforts from education providers, it remains a contentious policy issue. According to UNICEF's 'Learning Crisis Campaign' (see https://www.unicef.org/learning-crisis), only a third of 10-year-olds globally are estimated to read and understand a simple written story. Education policymakers are then confronted with the question of how to improve the quality of education. Increasingly, major players in education policy, whether governments or education committees, call for 'rigorous scientific evaluation methods,' particularly randomized controlled trials (RCT), to establish a credible link between an intervention and its desired outcomes. This article explores the indicators that policymakers and researchers in education care about when replicating and integrating research from the Global North in the context of the Global South. I also investigate how the available research influences and empowers decision-makers in the Global South to alleviate social problems. I conclude with a personal reflection about an existing experimental study in Kenya on the contextualization of research from the Global North to the Global South and reflect on this from the lens of the utility of meta-research, or research on research, and how it has been useful in helping us think deeply about our methodology and approach.

The contextualization black box

If education were indeed a public good, should governments focus on the supply or the demand of education? Boosting demand for education implies that governments find ways to incentivize parents to prioritize learning and focus on outcomes such as higher school enrollment and lower dropout rates. However, to improve the quality of education, governments also need to consider improving teacher education, investing in school infrastructure, and constantly improving assessments and curricula, among other things. However, since most nations in the Global South are constantly constrained for resources, what should they trade-off? Over a decade ago, Nobel prize laureates Abhijeet Banerjee and Esther Duflo asked this very question in an incredibly popular piece of research, their

book *Poor Economics* (2011). In this book, the authors review a dataset comprising eighteen countries, summarizing years of their work and that of others in experimental research to offer empirical evidence on what works. Their work advocates for caring more about the context in which a program is situated and viewing supply – and demand-side intervention tools as complementary. Duflo and Bannerjee posit that parents rarely have sufficient information on the real returns of an investment in education; they do not know how to respond to external interventions meant to stimulate demand for educational goods (Castillo & Wagner, 2014). They advocate for practices that combine supply – and demand-side interventions to accommodate local belief systems and behaviors. To provide an example, they review research on the extremely well-known Mexican educational scheme *Progressa* (later renamed *Oportunidades*), which offers conditional cash transfers (CCT) to households conditional on children's school attendance aimed at increasing school enrollment. An adaptation of the *Progressa* model in Malawi, which was studied through a government-funded RCT, found positive effects of the CCT program on increasing school enrolment and reducing the rate of dropouts. However, meta-analysis conducted on sixteen conditional cash transfer programs across Latin America found limited effects on long-term educational outcomes (Lomelí, 2008).

Bannerjee and Duflo review a similar example from an RCT in Kenya where twenty-five out of a hundred government schools were randomly selected to receive text books, an input that is often regarded as a positive indicator for educational achievement in low-income schools (Castillo & Wagner, 2014). However, the evaluation showed that only schools that performed well at the outset benefited from the textbooks (Glewwe et al., 2007). Further inquiry revealed that the textbooks were in English, which, despite being the official language of instruction at schools, was the third language of most Kenyan children. This indicates that context further constrains the generalizability of even the most rigorous evaluations. These contradictory findings highlight Duflo and Banerjee's theory that context matters. For a Global South policymaker or researcher, how would context be defined? The next section delves into evidence on the specific attributes that policymakers consider while consuming existing research on education.

How policymakers consume education research

To better understand how policymakers decide on what evidence to use to drive decision-making, it is helpful to take a closer look at Nozomi Nakajima's (2021) working paper shedding light on the factors that influence the preferences of education policymakers towards the consumption and use of education research. The study found that despite extensive information on the attributes of various types of research, such as their ability to establish external validity and draw causal inferences, the policymakers showed no preference for experimental studies. Nakajima can draw this conclusion from a discrete choice experiment where the policymakers are provided with ample information on the benefits of experimental studies. Maybe because of such research's high focus on technical jargon and less on applicability, policymakers showed no specific preference for such

studies in their decision-making. However, the policymakers preferred bigger studies with more sites and larger sample sizes. According to Nakajima, this is consistent with the methodological view of generalizing from 'broad to narrow' (Shadish et al., 2002), a concept which assumes that the estimates from multiple sites can be used to predict the impact on another site that is not included in the scope of the original research. The study also found that policymakers prefer studies conducted in settings with poverty rates, urbanity, and racial composition comparable to their own area of jurisdiction.

When asked about what information sources would be most useful to policymakers, 59.7% of the sample chose researcher forecasts, thus demonstrating that policymakers do have considerable interest in what researchers have to say (Nakajima, 2021). Nakajima's study also investigates the usefulness of information signals in this decision-making process, particularly if policymakers are susceptible to informational nudges on the research provided, and found that policymakers with higher scientific reasoning skills are more likely to rank researcher forecasts on top. Policymakers who are less confident in their own predictions on the effects of a particular education policy also tend to rank researcher forecasts above other sources of information. The next section draws a parallel between policymakers and researchers in the Global South, by drawing on Nakajima's study, and explores how susceptible researchers are to what kind of informational nudges. It also explores the attributes of research from the Global North that influenced the design and contextualization of an existing experimental study conducted in Kenya.

> **Global South researchers are not essentially too different from policymakers. Our pursuit of knowledge is largely influenced by the inequality that pervades our communities and is driven by our need to alleviate the conditions of the oppressed.**

From the Global North to the Global South: learnings on contextualization

To aid in reflecting on the previous section, as a Global South researcher, I present anecdotal evidence of our experience designing and running a large-scale field RCT in Kenya. This study explored the effectiveness of edutainment in last-mile distribution on literacy, social-emotional learning, and gender attitudes among children between six and nine. To help us design a large-scale study that requires strong external validity, we first tried to find as much experimental evidence as possible on the same topic. Interestingly, we could not find much experimental evidence that addresses this research question, much less in the context in which we planned to conduct the study. Our most reliable source of evidence was a meta-analysis conducted by Mares and Pan on the effectiveness of the Sesame Street television show, which was the closest in design to our intervention (Mares & Pan, 2013). As Global South researchers, what did we care about while adapting or contextualizing the existing research? Our first preference would have been congruence in the method of inquiry which was experimental research. Still, due to the lack of such, we went with the next best thing, a meta-analysis of quasi-experimental research on a similar intervention with comparable

outcomes of interest. We were constrained by the specific purpose of our study, which was to establish causality and have strong internal and external validity. Our intervention was also unique in that it was a freely available educational resource, the consumption of which could not be restricted only to a treatment group, and we needed strong guidance on how to approach the design of the study that ensured a pure treatment and control and, hence, our ability to draw a causal inference. The meta-analysis conducted by Mares and Pan, along with Busara's experience conducting previous experiments on informational nudges, guided our decision to use an 'encouragement scheme,' where the children in the treatment group were encouraged to watch the edutainment show, while those in the control group received no information about the program. Just like the findings in the Nakajima study (2021), we also cared deeply about respondent attributes such as race, urbanity, and socio-economic backgrounds and relied on similar examples from Mares and Pan's meta-study (2013) to aid our design. Lastly, regarding informational signals, the research team's affinity for statistical integrity and higher scientific reasoning skills made us rely on this particular meta-analysis for many key decisions we made. We used it to understand how the study is powered to find an impact, how to construct our cognitive and non-cognitive outcome measures and our treatment arms.

We await results from the study, not just on the impact of the intervention, but also on the efficacy of our research design in uncovering the evidence we set out to discover about how children learn. Throughout the research process, our team learned a lot about the realities of communities in Kenya, despite best efforts to 'adapt' existing research. We have learned that census data on power outages are grossly understated, that schools classified as being in urban areas are in truth lacking basic infrastructural amenities such as road accessibility, that most households do not have access to a television, despite census data on asset ownership, and that we really underestimated the learning crisis faced by children in Kenya. The hope is that one day we have enough evidence that is owned and produced by Global South researchers so that these basic contextual considerations are situated in our knowledge of our realities and that we ourselves do not become so far removed from them.

Who are we, really?

This personal reflection leads to a question of great importance, especially when we think deeply about what research does to people. What is the role of meta-research, especially at a time of greater inclination towards evidence-driven policymaking, in ensuring that the knowledge-generating scales are even? How can meta-research be leveraged as a tool of social justice and representation, especially in how we produce and own knowledge? Global South researchers are not essentially too different from policymakers. Our pursuit of knowledge is largely influenced by the inequality that pervades our communities and is driven by our need to alleviate the conditions of the oppressed. As we engage in this pursuit and strive to create better programs and policies, we find ourselves constantly in dissonance with our cultural identities and what we intuitively know to be true and the need to 'catch up' with western ways of thinking, needing to understand everything all at once and to be

better contextualized. There is a constant struggle not to recreate the wheel but at the same time not to be entrenched in colonial ways of thinking and knowledge generation. Paulo Freire aptly captured it in his Pedagogy of the Oppressed (2017): action is human when not detached from reflection. This is embodied in meta-research. Meta-research inspires us to think about how we think; it is the metacognition of knowledge generation as if the world were one interconnected web of thought or a common stream of consciousness. Meta-research creates accountability and gives us the tools to reflect on how we make meaning of the world. This is of paramount importance in education policy and research. As we strive to create better educational systems, push our boundaries of what educational achievement could look like, and create systems in which policy-makers increasingly rely on the evidence to make decisions, meta-research creates a unique opportunity for Global South researchers. It allows us to delve into why we choose to ask certain questions and how we choose to answer them. Hopefully, meta-research creates a new world where all of us learn in better ways.

References

Castillo, N.M., & Wagner, D.A. 2014. Gold standard? The use of randomized controlled trials for international educational policy. *Comparative Education Review* 58(1), 166-73. https://doi.org/10.1086/674168

Duflo, E., & Banerjee, A.V. 2011. *Poor economics: a radical rethinking of the way to fight global poverty*. PublicAffairs.

Freire, P. 2017. *Pedagogy of the oppressed*. Penguin Classics.

Glewwe, P., Kremer, M., & Moulin, S. 2007. *Many children left behind? Textbooks and test scores in Kenya*. National Bureau of Economic Research, Working paper 13300. http://www.nber.org/papers/w13300

Lomelí, E.V. 2008. Conditional cash transfers as social policy in Latin America: an assessment of their contributions and limitations. *Annual Review of Sociology* 34(1), 475-99. https://doi.org/10.1146/annurev.soc.34.040507.134537

Mares, M.-L., & Pan, Z. 2013. Effects of Sesame Street: a meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology* 34(3), 140-51. https://doi.org/10.1016/j.appdev.2013.01.001

Nakajima, N. 2021. *Evidence-based decisions and education policymakers*. Working paper. https://nozominakajima.github.io/files/nakajima_policymaker.pdf

Shadish, W.R., Cook, T.D., & Campbell, D.T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

PART THREE

How to improve how we conduct research

The democratizing effects of doubting

Adam Moe Fejerskov (Danish Institute for International Studies)

Doubt has become a deeply contentious issue in scientific milieus and society writ large. The relationship between critique and scientific knowledge is precarious in these supposed 'post-truth' times, where doubt feeds conspiracies and delegitimizes science, thereby catalyzing antidemocratic forces. Yet the complete absence of doubt does no good either, potentially nourishing an understanding of science as authoritarian, exclusionist, and undemocratic. Despite attempts to build skepticism around these and many other forms of scientific consensus, climate changes are obviously caused by humans, smoking does cause cancer, and childhood vaccines are indeed safe. But that does not mean that we, as researchers, cannot improve how we actively embrace and engage forms of doubt.

Everyone who has engaged in research, no matter how intensively, knows that the practice of science is one of assemblage, of taking things apart, putting them together, making choices, leaving concerns out, creating boundaries, and making judgments which, essentially, means that no absolute separation exists between values and epistemic conditions of science. Who we are, what we come out of, or how we are trained influences how we do science. If acknowledged and used well, this is not a weakness of science but rather an asset. And it is important because science influences policy and politics. Neither closed nor real-world laboratories have any insularity to them, and the consequences of experiments conducted in them continuously seep out into society.

When I argue here that embracing doubt does not devalue research findings – quite the opposite, it increases trustworthiness – I am inspired by American (think John Dewey, see Dewey, 1958 and 2006) and French pragmatists (think Laurent Thevenot or Luc Boltanski, see Boltanski & Thevenot, 2006), who historically saw and see doubt as a critical catalyst for changing established beliefs. For Dewey, doubt was a core tenet of reflective thinking. It, however, should not just be seen as a starting point for science but as continuously essential to the practice of research. Only by embracing doubt can we continue to question how, where, for what, or whom we conduct our research, essentially making doubt a democratic practice.

Present experimental regimes and the insularity of science

Many of the experimental scientific practices prevalent today and discussed in my book *The*

Global Lab (2022) do not embrace the values of relativism or pluralism when it comes to questions about science and evidence, thereby neglecting the very important matter of who and what can and should inform policy and thus set out future pathways for society[1]. For the hardcore proponents of experiments, especially randomized controlled trials, there is no outlook on science as a process of cumulative understanding. Either you generate your knowledge from RCTs, or your results have no legitimacy or basis.

Experimental regimes as we know them from, for example, development economics have proved immensely skillful in converting scientific and technical authority into political and moral authority. They should rightfully be lauded for the mainstream attention they have brought to issues that far too easily fall outside the realm of interest for most Westerners, Western institutions, and media that do not feel a strong inclination to report on education, water supply, or health in the Global South. Still, the methodological monopolist and disconcerting conception that exactly their – and only their – ways of doing and thinking science fall outside the realm of morals and politics makes them prone to criticism. The ideals of an expert-advised democracy are laudable. Still, the problem remains that what constitutes 'expertise' is too often interpreted by the experts themselves, just as who counts as an expert to begin with.

Present experimental practices mostly conceive of experimental subjects as potentially polluting research findings or as a nuisance whose doubts and skepticism must quickly be overcome by ethical clearance. Experimental scientists may take their trials into the social settings of the subjects. Still, here, the latter are passive observers of the star economist and his instruments, at worst, unaware of being subjects of experimentation altogether. Dyadic, reciprocal relations are substituted by one-way interactions or detached observations, and the technical proficiency of experimenters and the deep complexity of their econometric endeavors is often sufficiently abstract enough to decouple subject and experimenter by way of scientific and economic privilege; even if respondents also have their own ways of resisting or subverting directions. When some economists have difficulty in seeing through their own methodologies and scientific practices to acknowledge degrees of unequal power or social relations at play, they display a worrying and, at the same time, ignorant privilege. To present experimentation as clean, clinical, objective, and apolitical is to neglect or suppress the unequal power relations at play. This creates an opacity that serves as an excluding mechanism, defining who can have a say in experiments and strengthening a particular type of scientific authority.

Current experimental regimes often tend towards monopolist imperatives of scientific and policy truth that imply little pluralism and show almost no interest in a diversity of knowledges – it is 'our way or the highway.' This methodological totalization means there

1 For critical discussions of RCTs, see also Deaton & Cartwright (2018); De Souza Leao & Eyal (2019); Woolcock (2009); Ravallion (2009); and Donovan (2018).

are few dialogical efforts and limited room to make use of and fruitfully employ the existing multitude of methodologies. In contrast to the assumption underlining this methodological totalization, I would argue that evidence always possesses multiple meanings with no one right or pre-given way to obtain or interpret knowledge, providing much-needed room for the humility of scientific doubt, recognizing the responsibility of society to doubt, question, and not to be sure. Constructed or perceived absolute truths are inherently ethical-political because of their structuring potential. If a particular finding is upheld to a near-axiomatic level, it can determine interventions and policies far beyond the setting in which it was 'uncovered.' That is the reason why doubt can form a good counterweight by recognizing potential contingencies and uncertainties and the vulnerabilities arising from them, both along the way in the scientific production of outputs and results and the post-trial translation of statistical effects into policy recommendations. Yet, like the algorithm's computational reality of outputs almost always having to be a numeric probability, most experimenters are not satisfied with results that introduce doubt. Either there is a documented effect or there is not. Either the statistical result is significant or it is insignificant. We could say then that a core purpose of some forms of experimentation is to fundamentally eradicate doubt, condensing a multiplicity of options to a single output that may shape policy and the lives of many.

> **To present experimentation as clean, clinical, objective, and apolitical is to neglect or suppress the unequal power relations at play. This creates an opacity that serves as an excluding mechanism, defining who can have a say in experiments and strengthening a particular type of scientific authority.**

The importance of doubt and values in research

To most experimental regimes, doubt is even quite dangerous because it is seen to open a gate to the much-feared notion of subjective bias. Many experimenters assume that they can construct walls of separation between science and the *real* world, inherently detaching research from morals, values, or politics, thereby secularizing the experiment. This assumption arises from the belief that we can mobilize neutrality, objectivity, and impartiality to simply lay bare the facts of the natural world, ready for anyone who wishes to follow truths in policymaking. This production of truth is deeply connected to the value-free ideal that sees normative value commitments as a source of bias polluting pure research. The most uncompromising views on this matter maintain that no values are allowed to shape research, with more moderate views arguing that we can distinguish between epistemic and non-epistemic values. That is, values *internal* to the scientific community and related to scientific reasoning, such as the scope of research, are accepted while values seen as *external* to science, such as politics, moral, or social issues, are not. Yet, such a demarcation between epistemic and non-epistemic values is unconvincing. Although they appear 'neutral', epistemic concerns of accuracy, consistency, scope, or simplicity all relate to political and

social values on questions such as rights, redistribution, or inequality. Experimenters may argue that the results they produce and present are simply objective numbers measured in transparent ways that can, in theory, be repeated by others elsewhere. However, values are inscribed in methodological choices long before any documentation of 'neutral' results – and the values that shape what questions we ask in science, how and where we wish to pursue do not simply vanish later in the research processes.

Moral, political, social, cultural, and other assumptions are deeply embedded in the sciences in the first place, even when researchers do not pay explicit attention to them. Methodological choices and consequences – fundamentally any choice about what to research, how to do it, and why – may appear to be based on a pure interpretation of science, but they are always constituted by values. As such, there can be no real distinction between normativity and factuality, just as there is no 'free' research arena detached from socio-political worlds. All science is consequential, if in different ways and to different degrees. We have to challenge the value-free ideal for its presumptions that science can be practiced outside of a conception of society and its social, political, cultural, or economic forms. Whether directly or in derivative form, the practice of science always has implications for someone somewhere and carries with it conceptions about how society should be organized and ordered.

Experiments are not autonomous forms of theory testing but assemblages of cultural patterns, scientific practices, political economies, epistemologies, and many more coming together to create collections of ideas and practices that shape the production of truth, with no separation of the results of experimentation from the making or production of experiments. So much has been clear since the first researchers ventured into laboratories and observed the very mundane practice of experimental science, eventually confronting myths of any nature-given coherence of this. The problem with much of the experimental science we see practiced today is that it does not seem to inspire emancipation as much as to uphold unequal relations of power and resources. Not least because many scientific disciplines suffer from a conscious or implicit silencing of knowledge production at the 'margins.' Some forms of knowledge fall outside the imaginary walls of separation erected by certain authorities in these fields, whether individuals or institutions. We need to work towards a more participatory science and a practice of experimentation that recognizes its inherent inequality of engagement and works productively from and with that inequality in mind. Such a form of participatory science leaves room for substantive public contestation and, from a co-production perspective, sees doubt, and thus reflexivity, as something relational and distributed, embedded in and making use of the hybridity of expert and subject identities.

Instead of focusing on verification and singular truths, attention would be better paid to variance and how differences in results can lead us to more inclusive discussions of ideal outcomes. Rather than increasingly adopting precautionary principles or merely heightening ethical awareness while maintaining similar methodologies and practices,

we need to radically reconfigure the methodological regime based on values of doubt and democratization. Whereas most experimental regimes today do what they can to hermetically close experiments from outside influence, performing a monologue more than a conversation, we need to open up scientific processes. Instead of closing off scientific practice, we are better left with viewing experimentation as a political practice deeply rooted in democracy – a cooperative inquiry binding science and politics together. Following these thoughts would be to argue for much stronger civic epistemologies where science is increasingly socially embedded and practiced as a hybrid interaction between scientific and social actors.

References

Boltanski, L., & Thevenot, L. 2006. *On justification: economies of worth*. Princeton University Press.

De Souza Leao, L., & Eyal, G. 2019. The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society* 48, 383-418. https://doi.org/10.1007/s11186-019-09352-6

Deaton, A. & Cartwright, N. 2018. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine* 210 (August), 2-21. https://doi.org/10.1016/j.socscimed.2017.12.005

Dewey, J. 1958. *Experience and nature*. Courier Corporation.

Dewey, J. 2006. *Essays in experimental logic*. Southern Illinois University Press.

Donovan, K.P. 2018. The rise of the randomistas: on the experimental turn in international aid. *Economy and Society* 47(1), 27-58. https://doi.org/10.1080/03085147.2018.1432153

Fejerskov, A. 2022. *The global lab: inequality, technology, and the new experimental movement*. Oxford University Press.

Ravallion, M. 2009. Should the randomistas rule? *The Economists' Voice* 6(2). https://doi.org/10.2202/1553-3832.1368

Woolcock, M. 2009. Toward a plurality of methods in project evaluation: a contextualized approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness* 1(1), 1-14. https://doi.org/10.1080/19439340902727719

‘Here, my degree does not matter; you are the teacher’: ethnography, citizen ethnography, and research on global development

Ben Eyre (University of East Anglia), Sharon Acio Enon (Citizen Ethnography Collective), Dorah Adoch (Citizen Ethnography Collective), Vicky Alum (Citizen Ethnography Collective), Joel Ekaun Hannington (Citizen Ethnography Collective), Ann Gumkit Parlaker (Citizen Ethnography Collective), Ben Jones (University of East Anglia), Jimmy Ezra Okello (Citizen Ethnography Collective), Robert Oluka (Citizen Ethnography Collective) and James Opolo (Citizen Ethnography Collective)

Jimmy Ezra Okello was afraid that people who normally saw him smartly dressed and commuting to an office would look at him doing welding work. However, to carry out ethnographic research, he had to rid himself of these fears and ask Julius if he could learn to weld in his workshop. The older man told him, ‘I only stopped in senior four [secondary level education]; you are a bachelor’s degree holder. I cannot teach you!’ Okello had to convince Julius by telling him: ‘Here, my degree does not matter; you are the teacher.’ Julius laughed at this but began to show the younger man the different machines he used. Soon Okello was helping to hold door frames in position and lugging heavy machines around the workshop floor as he undertook the skilled yet strenuous work. By participating imperfectly in Julius’s work and observing over several months at the workshop, Okello came to learn about ethnography by doing it.

Okello is one of the authors of this paper. He did not learn about ethnography within an anthropology or sociology department. He was not assigned a participant observation exercise within a placement from a university course. Along with seven of the other nine authors of this chapter, Okello learned ethnography within a project about education in Uganda led by another co-author (Ben Jones) that created a space for young people to shape and conduct research rather than just being depicted in it. The project trained people with relevant lived experiences and no prior knowledge of these methods to conduct ethnographic research about ‘being educated’ in two areas of Uganda. In early 2023, they were joined by Ben Eyre, who was researching how ethnographers could better collaborate with non-anthropologists and non-academics. Now, the team has begun collaborating with academics in other disciplines, practitioners and policymakers who are increasingly interested in ethnography. This chapter introduces our group’s reflections on the value that ethnography can bring to

research about research on global development. We briefly discuss ethnographic methods in general before introducing the idea of 'citizen ethnography' as one way of describing ethnographic research that is not done (purely) by outsiders but by (mixed) teams of people with lived experiences of research topics at hand, and some prior familiarity with the research locations. We then give examples of why such approaches and methods can contribute to the study of research on global development. Although ethnographic research has its own epistemological premises, we hope to speak to those unfamiliar with ethnography and begin an interdisciplinary discussion.

Ethnography

Within contemporary development research, ethnography is often mentioned but seldom discussed in detail. Frequently, it is used loosely to connote qualitative methods in general. This lack of clarity is challenging for those who want to understand its strengths and weaknesses. Among anthropologists, ethnography has become synonymous with participant observation – where the researcher becomes a participant in the context in which she is working – conducted over the long term (Miller 2017). It suggests approaches to data gathering and a form of writing based on them that privilege insider perspectives and lived experience. As such, ethnography tends not to start with, or stick with, a hypothesis but to iteratively adapt to what matters to the people one is studying and what they do and say. What they think matters, not just what the researcher thinks is important. For this reason, ethnographers often explicitly reflect on their own limitations, and ethnographic approaches benefit from critical self-analysis and reflection. This is informed by twin concerns with context and perspective that affect all research processes but are sometimes hidden in the presentation of the results (as discussed below). These concerns are particularly important when research relations are characterized by economic inequality, language challenges, and misunderstanding between those doing research and those having it done to them.

Ethnographers aim to address such issues by paying attention to the perspectives of the people being researched. But this is easier said than done. Trust is a major issue, and it requires time and attention to know how people feel about the research, its context, the researcher, and themselves. Ann Gumkit Parlaker remembers her early attempt at data collection in Lira City: people could not express themselves because certain questions were personal and hard to answer when a neighbor might be listening and when the answers might spark gossip or feuds. Gumkit Parlaker noticed that women, in particular, were shy about talking in public. To overcome this, ethnographers build relationships over the long term. This often involves not pushing people to answer quickly. James Opolo remembers a young woman (about 25 years old) with whom he had a deep discussion at first meeting: as she told him a story about dropping out of school, she began to weep! Opolo was concerned for her, rather than simply about what she could tell him, so he started to talk about people who did well in life despite dropping out of school, and they agreed to continue her story another time. Later, they were able to have a deeper conversation because she knew Opolo was not just there to collect data but to listen to her.

Participant observation is a foundation of ethnography because it enables and builds these relationships. As such, our team became apprentice welders and hairdressers. We attended week-in-and-week-out village savings and loan associations (VSLAs) and worshiped with different church denominations over months. We did not do so with clipboard (or tablet) in hand, but dancing, singing, praying, or investing our hard-earned money in savings groups. The approach here is holistic immersion in people's lives rather than a weekly appointment. Hours spent participating and observing allow for thorough discussions built on mutual familiarity and for comparison and questions based on observations. Okello emphasizes the need for copious notes. They often run to several thousand words each day. These cover the incidental as well as the obviously relevant. They can be recorded during observations or soon after. Detailed notes slowly build up a comprehensive picture and enable analysis later through cross-referencing key themes. Participant observation might not involve many direct questions on the research topic other than those that arise organically out of activities and shared interests.

> **The self-reflexive critique of perspective and context that ethnography calls for makes it particularly useful for researching research. Ethnographers can step back and think about their own and other research practices.**

Citizen ethnography

Our collaboration began with Jones's study of education in Uganda. He sought to build on ethnography's participatory and dialogic strengths by including more closely people with lived experiences of 'being educated' who struggle to find formal employment. He began by recruiting a team of nine young people with the help of a local NGO and in partnership with Laury Ocen of Lira University. Jones and Ocen provided them with training in ethnography, and they received online training in ethical research practices from Macquarie University. This group, most of whom are co-authors of this chapter, already lived in relevant field sites and had inside experience of the research questions; they used this to actively shape the direction of the research (Aguti et al. 2024; Jones & Amongin 2023). This collaboration has grown to include Eyre, who was working with academics in other disciplines, such as behavioral economics and psychology, and with professionals beyond the academy who (might) value ethnographic insights (Eyre 2022). Collectively, we have begun constructive dialogues with practitioners and policymakers interested in ethnography to better listen to target beneficiaries, customers, or users. To summarize the strengths our diverse team affords, we have begun to refer to 'citizen ethnography.'

Citizen ethnography is not a perfect term. It risks recalling some problematic false dichotomies of 'native anthropology': foreign/home, insider/outsider, observer/ observed (Narayan 1993). Nonetheless, we (particularly those who would describe themselves as 'citizen ethnographers') find it useful to convey our distinctive contributions within a

diverse team (see also Enria 2022). We are encouraged by wide awareness and acceptance of 'citizen science' in which people are not defined by their background but speak from the specificities of their position to make a personal contribution to scientific knowledge, one that thereby incorporates more diverse perspectives. The 'citizenship' that we refer to (although it might suggest essentializing the experience of being subject to a sovereign nation-state) should metaphorically convey belonging to and ownership of research. In this way, we suggest that citizen ethnography expands on the broader ethnographic impulse to learn with and from people.

Some of the benefits of citizen ethnography are obvious. Researchers with the same language skills as their interlocutors conduct more nuanced and relaxed discussions. Those who share formative life experiences build rapport more quickly. They know how challenges they ask others to discuss can affect life, even if their circumstances are different. This is particularly important given the racialized context of a lot of ethnographic and development research. This is exacerbated by a challenge that academic researchers in many parts of East Africa and beyond often face: being compared with NGOs, which are the source of considerable resources. Some of these challenges quickly became obvious to our team. Gumkit Parlaker recalls a meeting with Jones and a group of young people. The group thought that they were from an NGO. Seeing a *muzungu* (a white person) made things worse. Everyone wanted the attention of the white man. It meant money. Dorah Adoch also came across the challenge of money. During her initial meeting with a new group of interlocutors, people wanted direct and material benefits for their time. They expected scholarships or some other aid, judging by how one mother kept trying to persuade her daughter, who was the only recent graduate in the area, to attend to avoid missing out.

Although they cannot circumvent these challenges, citizen ethnographers can negotiate them through grounded insights that other ethnographers lack. Rather than approaching them through lofty analytical or theoretical categories, they have a practiced way of being attentive during the research itself, as well as being able to reflect on intersecting challenges afterward during analysis. Opolo experienced a challenge when he realized that engaging with female interlocutors in certain places was difficult for him. He overcame this by quickly including his female colleagues in the team during such discussions. There were limits to the conversations he could have, but they were possible to negotiate as a team.

Citizen ethnography is not simply a technique to access 'community voice' or enable 'local participation.' The worthy impulse of including diverse voices has not always been pursued rigorously by asking difficult questions about who represents whom, self-reflexively critiquing research practices, and thinking about the contexts in which research is conducted. This means 'communities' often assent to statements they disagree with during a participatory exercise that obfuscates politics, inequality, and disagreement. In fact, knowledge of a place and people often challenges the idea that there is a homogeneous 'community' or a bound 'locality' to be represented. Adoch notes that her 'ethnographic journey' consistently defied

expectations about what she knew about potential interlocutors and what she would learn from them. Sharon Acio Enon expected little from Betty, who was a hairdresser. But by spending time with her, Acio Enon realized that Betty was busy networking and employing people. She was educated but in a different way. All our team noted that when we assumed who would give good data, we were often wrong. Nor are citizen ethnographers a shortcut to addressing difficult issues. Gumkit Parlaker had expected certain things from her inquiries about land conflict. She expected women to tell her that they were buying their own land. However, the problem was that people did not want to discuss the issue. Only by entering people's lives, never really asking the question, but learning holistically about many aspects of negotiating life as a young woman, were they able to come to the question much later. Land conflict came back, but only after Gumkit Parlaker had developed trusting relationships.

Researching research on global development

The self-reflexive critique of perspective and context that ethnography calls for makes it particularly useful for researching research. Ethnographers can step back and think about their own and other research practices. Several members of our team have conducted surveys and participated in focus groups. Developing expertise in ethnography has helped us to reflect on these experiences. This has been useful as we think more proactively about the strengths and weaknesses of different research methods.

Focus groups and key informant interviews are common qualitative research methods in interdisciplinary research. They also seem to share traits with ethnography because they do not necessarily follow an explicit script. But Gumkit Parlaker remembers her first experience of conducting focus groups quite differently. She requested the help of the Local Council Chair (LC1) to recruit young people aged 18-35. To her dismay, it was clear that the people that he brought together were mainly his relatives, and many were above this age. This created a difficult dynamic where young people were reluctant to contradict or even speak in front of their elders. Although Gumkit Parlaker tried to create a relaxed mood and forge good relations with everyone in the group, she struggled. The atmosphere was not friendly, and it seemed that the younger participants in particular felt they were being interrogated. Joel Ekaun Hannington experienced similar challenges in a focus group that brought together young people with different levels of education alongside leaders from the area. Those who felt 'uneducated' and unimportant were ignored by the others and did not participate fully in discussions. Although the group seemed to come to a consensus, thanks to the expertise of the research coordinator who organized the session, it was clear to Hannington that it did not account for all perspectives or life experiences.

Hannington identified other challenges when conducting survey work. Some respondents hoped to get assistance from visiting researchers if they gave particular answers to questions about their property. From what he had already observed, he thought it was likely that many gave answers that were not true. At other times, young people sometimes hid the fact that they had children. It was not always obvious why people shied away from telling the full

truth about their lives. Perhaps because surveys were often perceived as transactional. Opolo observed that both the enumerator, who conducted surveys with a mobile tablet in hand, and the survey participants sometimes described feeling like a machine themselves: an extension of the device. The room for deviation from the script was minimal, leaving people little room to add their own perspectives and insights. Even the enumerators could not amend the survey if they realized that a different question might help to better understand the situation.

Our team also experienced additional challenges. Hannington noted that when printing materials, researchers often did not receive enough money for everything. When conducting surveys and questionnaires, respondents sometimes did not even understand the question, and so guessed or sought the guidance of a neighbor or the enumerator. When using tablets or mobile phones, gaps in internet connectivity could also cause problems. Researchers would have to note responses and fill them in later. Such issues are exacerbated by the time pressures on enumerators to complete surveys extremely quickly to meet their target and find more work. The 'worst performers' are routinely let go, while 'good performers' can find opportunities for advancement through more responsibility. The contrast between doing well and poorly is based on speed, volume of work, and collection of complete and expected responses. These demands heavily incentivize enumerators to 'cook the data' (Biruk 2018) to ensure it says the right thing. This is a significant issue for those reliant on survey data, but there is no easy fix; least of all through the same methods. The inequalities encapsulated in research processes, combined with a perception that they are extractive and the idea that they can be gamed, make triangulation difficult. By building trusting relationships ensuring that any insight shared in confidence will never be traced back to them, ethnographers are well placed to explore these challenges with interlocutors who otherwise struggle to discuss them for fear of punishment or betraying others. This is even more the case for citizen ethnographers who have worked in such roles themselves but now have a different position from which their analysis is aided by reflection and by empathy for those who perpetuate the problem because of the requirements of their role. Robert Oluka suggests that he felt empowered as a citizen ethnographer because he empowered his interlocutors to talk and to be heard in ways that enabled them to shape the analysis.

If citizen ethnography helps us get to a more nuanced understanding of how things work, is this enough? In their study of village savings groups, Jones and Amongin (2023) found something surprising. Members of the group – mostly women – were saving for someone else. These members came along every Friday with someone else's money. They were, it turned out, giving up their labor for richer members of their community within the sort of institution that is often held up as an example of community empowerment. The practice of sitting for someone else is not the sort of insight that comes from an interview or a focus group discussion where people often follow an unofficial but widely understood script. While members had a sort of 'open secret' about this practice, it was not something to reveal to outsiders. But it is important to follow savings groups in their actual practice, as they

really are. By exploring what happens when poorer members sit for richer members, Jones and Amongin comprehensively analyzed why savings groups work and who they might be working for. These grounded insights have brought their research to the attention of economists working on major humanitarian interventions in Western Uganda. Amongin is now bringing her critical skills and insights to a development project along with the co-authors of this chapter. This capacity for rigorous, iterative hypothesizing can make citizen ethnographers valuable members of interdisciplinary teams.

Conclusion: (citizen) ethnography as a contribution to meta-research

Ethnography creates space for reflection on research practices. As Okello and the co-authors have found, this can feel uncomfortable at different times throughout the process. While methodological introspection is important in ethnographic research, it is (generally) less common in other research methods. For this reason, ethnographic reflection on surveys and focus groups of the type we have very briefly outlined above can seem critical and even antagonistic. That has not been our aim. Instead, we feel that citizen ethnography can complement other methods used in meta-research. This might involve triangulation in a narrow sense, stress testing, double-checking, and catching inconsistencies. But more expansively, it offers a means to discover more about what the people carrying out and being researched think of the methods used, their own role in them, and what they miss in producing knowledge about their lives. These insights about why things can go wrong might be the basis of new ideas about how to do research better, which can then be tested through various methods beyond ethnography.

Although anthropologists sometimes caricature quantitative researchers, particularly economists, as uncritical and unreflexive, we have noted growing interest from academics working in other disciplines and applied development researchers in understanding the contexts in which research about global development happens and more fully engaging with the different perspectives of people who have a stake in it. This creates an opportunity for transdisciplinary collaboration. What we have described as citizen ethnography centers on those who have experienced research from multiple positions (as people being researchers or mobilized, as enumerators or focus group participants and facilitators, as well as ethnographers). This does not neatly map on to 'better development research.' However, exploring research relationships in all their complexity enables diverse aims spanning applied goals of better data and more equitable intervention to theoretical innovation in how we think about development research. Ethnographers might consider what kind of (global) citizens they wish to be.

References

Aguti, S., Amongin, S., Jones, B., Okello, J.E., & Oluka, R. 2024. Becoming 'business class': educated youth and Pentecostal change in eastern Uganda. *Compare: A Journal of Comparative and International Education*. https://doi.org/10.1080/03057925.2023.2292518

Biruk, C. 2018. *Cooking data: culture and politics in an African research world*. New York: Duke University Press.

Enria, L. 2022. Citizen ethnography in outbreak response: guidance for establishing networks of researchers. *Social Science in Humanitarian Action (SSHAP)*. https://doi.org/10.19088/SSHAP.2022.001

Eyre, B. 2022. Action as method: ethnographic approaches for impact. *Anthropology Today*, 38, 27-8. https://doi.org/10.1111/1467-8322.12733

Jones, B., & Amongin, S. 2023. 'Money looks for money': managing financialization in eastern Uganda. *Africa* 93(4), 455-75. https://doi.org/10.1017/S000197202300061X

Miller, D., 2017. Anthropology is the discipline but the goal is ethnography. *Hau: Journal of Ethnographic Theory* 7(1), 27-31. https://doi.org/10.14318/hau7.1.006

Narayan, K. 1993. How native is a 'native' anthropologist? *American Anthropologist* 95(3), 671-86. https://doi.org/10.1525/aa.1993.95.3.02a00070

Do the randomized know that they have been randomized? A critique of the turn towards randomization in high-stakes development cooperation initiatives

Mario Schmidt (Busara & Max Planck Institute for Social Anthropology)

Development cooperation has been attacked from different angles and disciplines for various reasons for decades. Apart from critiques highlighting the sector's origin in colonial conditions that, in their postcolonial forms, still influence how development 'is done' (cf. for example, Hoffmann 2020, Ndlovu-Gatsheni 2020, see also Tilley 2011), the accusation that a lot of interventions fail to produce the effects they claim to be able to produce has worried many actors in the sector. As the efficacy of aid is particularly important for donors and philanthropists who prefer interventions that yield sustainable, innovative, impactful, and scalable results that justify the expenditure of their capital, a group of scientists and aid workers introduced methods that produce quantifiable evidence of causal effects. This group, spearheaded by the winners of the Nobel Memorial Prize in Economic Sciences 2019, Esther Duflo, Michael Kremer, and Abhijit Banerjee, particularly pushes for the use of so-called randomized controlled trials (RCTs), a method with a long history in the medical sciences. During randomized controlled trials, a targeted group is divided at random into a treatment group that benefits from the intervention and a control group that does not. Based upon the assumption that such experiments produce evidence of causal relations, they have become the 'gold standard' of development cooperation (Webber & Prouse, 2018, see also Bédécarrats et al., 2019). During my fieldwork on the experimentalization of aid interventions in Kenya, I was surprised that even very small NGOs often felt the need to conduct randomized controlled trials to attract funding.

This contribution looks specifically at randomized controlled trials conducted in the context of universal basic income and unconditional cash transfer interventions, during which participants are usually aware of whether they have been placed in a treatment or a control group: they simply receive or do not receive money on their mobile phone wallets (see also Schmidt, 2022; Lassak & Schmidt, 2024; Fotta & Schmidt, 2022). As randomized controlled trials assume that the intervention itself is the only statistically relevant difference between the control and treatment group, the ability to draw causal inferences depends on

the assumption that members of the control group are either unaware of the experiment happening or that they completely understand why they were placed in the control group, namely because a stochastic method assigned them there randomly, and, in addition to that, that they do not have a moral, political or ethical problem with their condition assignment. While the first option is, due to the ethical imperative to gain informed consent, ethically impermissible, I came across many control group members during my fieldwork who were unaware of why they were placed in the control group.

Out of the different reasons offered by my interlocutors when I asked them about why they did not receive money, I want to highlight two that risk distorting the experiment's scientific results as they change the participants' expectations and thus influence their economic decision-making, thereby potentially becoming confounding variables influencing the measured outcomes: the accusation of corruption targeting village elders or politicians, and the conviction that one has answered survey questions incorrectly and that the placement in the control group is, therefore, the respondent's own fault, resulting in self-blame and feelings of intellectual inadequacy. After describing these two potentially confounding variables in more detail, I suggest that NGOs conducting randomized controlled trials of universal basic income or unconditional cash transfer interventions should (1) actively accept their political role and aim to better understand the historical context and the role development cooperation has played in the region, (2) avoid randomizing at levels that make little sense locally, such as the household level, and (3) place more emphasis on making respondents aware of the principle and exact processes of randomization to avoid harm and to control the effect of participants' expectations on the measured outcomes.

Two reasons why respondents believe they ended up in control groups

Rumors about political corruption

While international NGOs that depend on the financial contributions of donors and philanthropists try their best to avoid being seen as corrupt, heralding the virtue of transparency instead (for a critique, see Strathern, 2000), they often appear to assume that the best way not to be seen as corrupt is to be non-political. In the case of the universal basic income and unconditional cash transfer projects that I ethnographically studied through participant observation and qualitative interviews with recipients, control group members, former field officers, local politicians, and church leaders, as well as those who rejected being incorporated into the programs, for instance, the NGO tried to focus on working together with community leaders such as chiefs and village elders who often – but far from always – have an overview of their respective home area.

These chiefs and village elders are deeply integrated into kinship relations as well as local networks of entrustment (Shipton, 2007), and thereby automatically into political alliances, which influences how they are seen by respondents who have failed to understand the complex idea of randomization. It is unsurprising that, during my interviews, several

control group members blamed village elders or politicians for interfering with the process by intentionally placing them in the control group based on probably decades-long feuds between the control group member and the village elder or between their respective families. Working together with chiefs or village elders is thus not an apolitical choice but deeply political and potentially influences how people perceive the distribution between treatment and control groups, namely not as result of randomization but as a direct effect of corruption (cf. Schmidt, 2014).

> **If research participants in randomized controlled trials studying high-stakes interventions, such as universal basic income and unconditional cash transfer programs, are unaware of the randomization mechanics and try to explain being in a control group by referring to other reasons, such as those mentioned above, the experiments risk producing questionable scientific results and causing unintended harm to the trial participants (see also Barrett & Carter, 2010).**

Self-blame and bad performances

The ethically more problematic examples I encountered were those of control group members who believed they were placed in the control group because of their performance during surveys. When we – a former field officer of a large international organization offering research services who also helped me as a research partner, Adrian Wilson, a PhD candidate from the University of Berkeley, and myself – interviewed Paul Akeyo in his unfinished hut located in a village not far from the lake-shore town Kamito, for instance, his hopes quickly turned into despair. Initially thinking that we were not merely interested in knowing more about his experiences with the NGO but actually employees of the NGO, he had hoped that we had returned to pay him the cash transfer he had been waiting to receive for several years by then. Asking him why he thinks that he never received the cash transfer, Akeyo, however, did not explain the randomization process but confided to us that he thought that he 'maybe made a mistake that I did not know of' (Dholuo, *samoro matimo mistake moro mok nang'iyo*) and that he 'was just crying, why me? Why not me?' He, therefore, felt that the intervention had been a bad thing because 'to be honest, I see that my fellow there received, that one there, but we were together, now that is why I thought maybe I answered some questions in a bad way or that I said some bad thing' (Kiswahili, *kusema ukweli sababu naona kama mwezangu pale alipata, yule pale natulikuwa pamoja, sasa hapo ndio pengine nilikuwa nafikiria pengine ili answer some questions mbaya ama nilisema kitu ingine mbaya*).

In other words, Paul had lived with self-blame and regret because he did not properly understand the methodology of randomized controlled trials, observing his friends and relatives using the cash transfer to improve their lives. At the same time, he remained behind, plagued by self-doubts for over a decade since the unconditional cash transfer program had started in 2013. It is worth highlighting his emphasis that 'we were together', whereby he

implicitly criticizes that randomized controlled trials create artificial boundaries between people who would otherwise view themselves as 'just the same'. What is a methodological benefit of randomized controlled trials, namely the comparability of the treatment and control group, is, therefore, also an existentially experienced and culturally appreciated value of equality for the recipients, which is at risk of destruction by the implementation of such experiments. The counterfactual worlds (cf. Gedeon Achi, 2020) brought forth by high-stake interventions studied by randomized controlled trials are thus worlds in which some people have been left behind intentionally, looking around themselves and seeing signs of progress they feel everyone deserves to possess.

Conclusion: avoid ethical harm and increase epistemological validity

If research participants in randomized controlled trials studying high-stakes interventions, such as universal basic income and unconditional cash transfer programs, are unaware of the randomization mechanics and try to explain being in a control group by referring to other reasons, such as those mentioned above, the experiments risk producing questionable scientific results and causing unintended harm to the trial participants (see also Barrett & Carter, 2010). While the potential of harm should be obvious from the examples of unfounded accusations of corruption and Paul's feelings of self-blame and self-doubt, the risks to scientific results might be less obvious. Yet, considering the ample evidence for the relationship between self-doubt and poor economic development (Wuepper & Lybbert, 2017), economic differences between control and treatment groups in universal basic income and unconditional cash transfer programs, which are rightfully seen as having the potential to change lives to the better (cf. Ferguson, 2015), might not only be caused by the intervention itself but also by the uncertainty as to why one has been placed in the control group.

One potential solution to this problem could be to simply avoid it by randomizing using clusters (cf. Puffer et al., 2005) that do justice to local conceptualizations of socio-spatial units. While Paul had problems understanding why his neighbor received the money and he did not, it might have been easier for him to understand and accept that no household in his village received a cash transfer when all households in a neighboring village did. This could, for instance, be understood as being caused by financial constraints on the side of the NGO. An ethically even less problematic randomization arrangement is a comparison across still wider geographical distances, such as, for instance, comparing treatment villages in one county with control villages in another county. Considering that the assumption of randomized controlled trials, however, remains that the control and the treatment group should be as similar as possible, this solution might not be accepted by scientists as methodologically robust.

Another option could be to ensure that research participants understand the randomization process better. Yet, the literature on how respondents understand the randomization process remains scarce (see Ouma, 2020, cf. Kerr et al., 2004, and Featherstone & Donovan,

2002 for a discussion of participants' understanding of randomization in medical trials), and there are not many experiments testing which types of explanation work best in the lab or the field. Considering that randomization is a concept not easily translatable into other languages and that there are a variety of different ways of explaining it – ranging from technical description, metaphorical explanations, visual demonstrations, and even enactments of the randomization process itself (for example, by flipping coins and so on) – improving respondents' understanding of randomization procedures should, alongside a more honest acceptance of NGOs' political power (cf. Das, 2020), be considered essential for creating a 'better how' in the field of development cooperation and interventions. I thus believe that it is an urgent task for scholars in the field of developmental meta-research to experimentally test the efficacy of different ways of explaining randomization, not only to assure the highest possible data quality but also to be in a better position to do the right thing – that is to adhere to the high ethical standards that come with experimenting with the lives, and futures, of socially, politically and economically already marginalized people such as Paul whose life had been disturbed by an experiment he did not even understand.

References

Barett, C., & Carter, M.R. 2010. The power and pitfalls of experiments in development economics: some non-random reflections. *Applied Economic Perspectives and Policy* 32(4), 515-48. https://doi.org/10.1093/aepp/ppq023

Bédécarrats, F., Guérin, I., & Roubaud, F. 2019. All that glitters is not gold: the political economy of randomized evaluations in development. *Development and Change* 50(3), 735-62. https://doi.org/10.1111/dech.12378

Das, S. 2020. (Don't) leave politics out of it: reflections on public policies, experiments, and interventions. *World Development* 127, 104792. https://doi.org/10.1016/j.worlddev.2019.104792

Featherstone, K., & Donovan, J.L. 2002. 'Why don't they just tell me straight, why allocate it?' The struggle to make sense of participating in a randomised controlled trial. *Social Science & Medicine* 55, 709-19. https://doi.org/10.1016/S0277-9536(01)00197-6

Ferguson, J. 2015. *Give a man a fish: reflections on the new politics of distribution*. Duke University Press.

Fotta, M., & Schmidt, M. 2022. Cash transfers. *The Open Encyclopedia of Anthropology*. http://doi.org/10.29164/22cashtransfer

Gedeon Achi, F. 2020. The embodied counterfactual. *World Development* 127, 104799. https://doi.org/10.1016/j.worlddev.2019.104799

Hoffmann, N. 2020. Involuntary experiments in former colonies: the case for a moratorium. *World Development* 127, 104805. https://doi.org/10.1016/j.worlddev.2019.104805

Kerr, C., Robinson, E., Stevens, A., Braunholtz, D., Edwards, S., & Lilford, R. 2004. Randomisation in trials: do potential trial participants understand it and find it acceptable? *Journal of Medical Ethics* 30(1), 80-4. https://doi.org/10.1136/jme.2002.001123

Lassak, M., & Schmidt, M. 2024. Free money's ideological nature: a comparative analysis of unconditional cash transfers in Eastern Africa. *Economic Anthropology* 11(1), 27-37. https://doi.org/10.1002/sea2.12293

Ndlovu-Gatsheni, S. 2020. *Decolonization, development and knowledge in Africa. Turning over a new leaf*. Routledge.

Ouma, M. 2020. Trust, legitimacy and community perceptions on randomisation of cash transfers. *CODESRIA Bulletin* 1, 25-9.

Puffer, S., Torgerson, D.J., & Watson, J. 2005. Cluster randomized controlled trial. *Journal of Evaluation in Clinical Practice* 11(5), 479-83. https://doi.org/10.1111/j.1365-2753.2005.00568.x

Schmidt, M. 2014. *'It will always be with us' – Corruption as an ontological fact among Kenyan Luo* (Global Cooperation Research Papers 7). Duisburg: Käte Hamburger Kolleg.

Schmidt, M. 2022. The gift of free money: on the indeterminacy of unconditional cash transfers in western Kenya. *Journal of the Royal Anthropological Institute* 28(1), 114-29. https://doi.org/10.1111/1467-9655.13655

Shipton, P. 2007. *The nature of entrustment. Intimacy, exchange, and the sacred in Africa*. Yale University Press.

Strathern, M. 2000. The tyranny of transparency. *British Educational Research Journal* 26(3), 309-21. https://doi.org/10.1080/713651562

Tilley, H. 2011. *Africa as a living laboratory: empire, development, and the problem of scientific knowledge, 1870-1950*. University of Chicago Press.

Webber, S., & Prouse, C. 2018. The new gold standard: the rise of randomized control trials and experimental development. *Economic Geography* 94(2), 166-87. https://doi.org/10.1080/00130095.2017.1392235

Wuepper, D., & Lybbert, T.J. 2017. Perceived self-efficacy, poverty, and economic development. *Annual Review of Resource Economics* 9, 383-404. https://doi.org/10.1146/annurev-resource-100516-053709

Manage relationships when starting and ending research with human participants

Joel Wambua (Busara), Anisha Singh (London School of Economics and Political Science), Kelvin Kihindas (Common Goal Research Center), Irene Gachungi (DIME, The World Bank) and Patrick S. Forscher (Busara)

Establishing and nurturing relationships with communities is critical for the success of development research involving human participants. By investing resources and effort into fostering these relationships, researchers can cultivate mutual trust, resulting in a shared understanding with participants and effectively mitigating potential issues such as resistance and unrealistic expectations within the community. Although relationship management is especially prominent at community entry, as this is the point of first contact when researchers and participants set expectations that will guide the research, relationships must also be managed to the study's end. At the end of the study, the researcher may have accomplished their main goal, namely to collect data, but the community may not. Many communities participate in research studies because community members believe the research will generate findings that benefit them or hope to benefit from the research materially.

Unfortunately, managing relationships in research is often overlooked, resulting in serious unintended social consequences that can create social harm. The risks of social harm are especially high when researchers and participants come from very different contexts, as these differences in context can inspire distrust and misunderstandings. Relatedly, suppose Global North researchers simply get knowledge in exchange for participant labor and incentives. In that case, fundamental power imbalances are created, potentially perpetuating colonial relationships and dynamics. From our experience, this may cause participants to feel disrespected, leading to a sense that researchers are merely using them to get funds to enrich themselves. This perception often causes participants to believe that the research will not contribute to improving their welfare, leading to careless responses during data collection and, in some cases, outright refusal to participate.

As employees of the Busara, we have more than our fair share of experience with the unforeseen risks that can arise when researchers do not manage relationships with participants effectively. As a nonprofit headquartered in Nairobi and dedicated to advancing

and applying behavioral science in pursuit of poverty alleviation, much of our research bridges the worlds of the Global North – where funders and our research partners often live – and the Global South, where much of our research takes place. We often serve as a broker between these worlds. This means we have also had a front-row seat to observe how harm can manifest when the chasms of understanding that separate these worlds are not effectively bridged.

Consider an example from our early days. We were conducting a study investigating whether transferring a large sum of money to lower-income Kenyans ($500 each) would improve people's economic and psychological well-being (Haushofer et al., 2020). The transfer of money was unconditional, which meant there were no strings attached – from our perspective and the perspective of our Global North research partners, the villagers could do whatever they wished with it. We implemented the study across several villages in Nakuru county, Kenya's third most populous county that cuts across the Great Rift Valley. Although Nakuru is home to growth and opportunity, it is also home to great poverty: in the villages where the experiment occurred, most villagers earned less than USD 200 per month.

Our study was a randomized experiment with a pure control group, meaning half the villagers did not receive the cash transfer. The problem was these control villagers did not know why. Thus, when the funds were administered after randomization, here is what the villagers outside the experimental group perceived: many of their friends, neighbors, rivals, and relatives suddenly received a windfall worth perhaps two and a half times their monthly earnings for reasons they did not understand. In contrast, they did not receive anything.

Some villagers were happy with this outcome because they could celebrate the good fortune of their friends and neighbors. However, others were distinctly unhappy – even envious. The result was utter chaos. Villagers who did not receive a payment, both in our control group and outside the experiment, started spreading rumors that the experiment had nefarious aims. For example, contrary to our intent that villagers perceive the money as unconditional, some claimed that the money was 'blood money' (cf Schmidt 2022, see also Schmidt, this volume), which compelled the recipient to perform a blood sacrifice. Others claimed that the experiment was a plot concocted by satanic devil worshippers locally known as the Illuminati and that the money was an inducement to recruit villagers into their ranks. In some cases, the intense feelings of anger, envy, fear, and jealousy even escalated to the point of domestic violence, as some men discovered that their wives took part in the experiment – and therefore received more than two months' earnings – without their knowledge. Tensions across the villages threatened to escalate into inter-household conflict. We therefore halted our activities during the study endline and spent some time trying to understand what had gone wrong and what we could have done to prevent it.

What went wrong in Nakuru County

We halted our study for a total of three months. During that time, we performed an extensive debrief. We interviewed villagers, engaged community leaders, and held informal community meetings, or barazas as they are known in Kenya, where we explained the study's objectives, introduced the research team, and allowed the villagers to air their concerns. Crucial to these proceedings was the fact that the people who we hired to help carry out the project, our enumerators, were from Nakuru county and, we assume, therefore, carried a degree of trust in the eyes of our participants – although many enumerators were not from the villages included in our study.

We learned from this reflection that we had incorrectly managed our relationship with our participants. Instead of holding our barazas and engaging the community before we started our study, allowing us to establish the baseline level of mutual knowledge and trust necessary to disburse such large sums of money to a random subset of villagers, we held them afterward. This allowed rumors to fill the gap in understanding and context between the source of the money and the recipients. Instead of engaging the community after the funds were disbursed to remind them that they would not receive any money, we left them to wonder when and if their cash would be received. Instead of treating our potential participants – both those who made it into our final sample and those who did not – as partners who could help us identify issues in the research process and help co-create solutions, we treated them as objects from whom insights are extracted. In short, we had failed our responsibility of bridging the gaps in culture, levels of income and wealth, job status, and other sources of context that exist across the research world, which the Global North heavily influences, and the world of our participants.

Lessons in relationship management

Our story of the cash transfer study in Nakuru County has a positive ending. After we learned about the relationship management mistakes that we had made and implemented corrective actions, we completed the remainder of the study without incident. We conducted community exit barazas to notify the community about the conclusion of the study and to address their expectations regarding cash transfers proactively. As a result, we have successfully carried out other studies in the region without the same resistance we faced during the endline phase of this project. This suggests that our barazas successfully managed our relationships with these communities by showing them we respect them enough to return the knowledge we gained from the study to them.

Yet, our story also illustrates the harm and challenges that can occur when researchers do not prioritize developing a foundation of trust with their participants. Ineffective relationship management increases the risk of social harm, and research that does harm is unethical. Researchers, therefore, have to plan their studies to minimize the risk of this harm. If that means that studies become more expensive, that is a cost that researchers and funders have an ethical duty to shoulder.

Better relationship management strategies

Cultivating a productive relationship with community members is a resource-intensive endeavor as it requires gathering insights from and communicating with diverse individuals. Researchers often engage a specific subset of individuals, typically community leaders, to represent the broader community. While this approach may initially seem efficient in terms of time and resource management, it often falls short in providing researchers with a comprehensive understanding of potential challenges that may arise when conducting research within a specific community. Furthermore, these leaders may not consistently relay information about the study to the rest of the community.

For instance, in our study in Nakuru, we maintained regular contact with community leaders throughout the study. However, this communication did not effectively translate into a broader understanding of our research activities within the community. When we completed the baseline and intervention phases of the study, we signaled the conclusion to the community leaders. Still, we failed to adequately inform the study participants, particularly those who received no cash. This oversight resulted in significant hostility and disengagement when we returned to collect end-line data, even sometimes putting our enumerators at risk. As mentioned earlier, we were forced to halt the study and organize debrief sessions with the participants and community leaders. In these sessions, we assessed whether the expectations set at the beginning of the study were met and to clarify why not everyone received financial support. This experience underscores the importance of managing relationships effectively, from community entry and maintenance to exit.

Finally, we emphasize that the 'communities' that play host to research are not homogeneous but are rife with all the complex political and social tensions that one might expect from moderate to large groups of humans. This point underscores the importance of not treating communities as homogeneous wholes and thereby unintentionally perpetuating colonialist stereotypes. Instead, you should take the time to understand the political and social dynamics at play in your workplace.

Below, we highlight the steps we consider crucial for managing community relationships.

Step 1: Hire Locally. Make sure your enumerators are locals. They know the context better than you and will help you flag potential harm in advance.

Step 2: Identify focal points. To gain support and positive relationships from the local community, it is important to involve the respected figures of the community such as local administration, village heads, community group leaders, and religious leaders. As they have expertise in the local culture, traditions, and context, they could assist in resolving any issues that may arise during your research activities.

Step 3: Involve the wider community. Hold town hall meetings (in Kenya, barazas) with the community members. These meetings are important because they inform villagers about study objectives and discuss expectations, culture, and so on with the goal of a mutual understanding of the study.

Step 4: Keep an open line of communication. This nurtures the relationship. Researchers should regularly engage with community members, updating them on study events like incentive distribution and interventions to prevent misconceptions and reduce tensions. Obtain consent at each study stage (baseline, intervention, end-line) to avoid misunderstandings about procedures.

Step 5: Reflect and update. Conduct end-of-study debriefs with participants and community leaders to evaluate if initial expectations were met. Document and address any issues for reference in future research-community relationships. For instance, in our case study, debriefs strengthened our relationship with the communities, enabling Busara to conduct successful follow-up studies.

Step 6: Share study results. This is not only ethical but important to the community. In recent qualitative research, participants expressed their desire to receive study results because when they agree to participate in research studies, they hope to learn something they can share with the rest of the community. Second, If researchers from the Global North simply get knowledge in exchange for participant labor and incentives, fundamental power imbalances are created, perpetuating colonial relationships and dynamics. We do not want that. The least we can do is ensure any knowledge gained from research goes back to participants in a manner that is comprehensible and useful to them.

The last steps, sharing research results and debriefing the community members, often prove to be particularly difficult to achieve. Before we conclude, we therefore want to highlight some guidelines for conducting debriefs and disseminating research findings effectively:

Plan in advance. Plan for community debriefs and sharing results during the design phase of your study. Include these activities in your protocol to hold yourself accountable. Debriefs and results sharing are resource-intensive exercises. building them into your design allows you to plan them adequately, and ensure you set aside enough resources to follow through.

Contextualize your approach. There is no one-size-fits-all way of conducting debriefs and sharing results. Embrace dialogue when planning. Engage the community members to identify context-relevant, effective, and meaningful debriefing methods. For example, in some communities, we have an official meeting day where members meet to discuss issues affecting their communities. By engaging community members during your planning phase, you get to know such information, and you can leverage these meetings to reduce the costs of debriefing and maximize impact.

Co-design solutions. People are experts in their own lives. They know what works and what does not work in a given context. If issues arise during the debrief sessions, ensure a clear way forward before you continue with other research activities. Be honest with what is achievable. Do not make promises you cannot keep to achieve high recruitment rates. In the long run, deceiving participants may build distrust, leading to high attrition in future study phases or resistance during future engagements.

Learn and update. Sharing results is not only ethical but also allows you to check your interpretation of the data you collect. Use this opportunity to improve the accuracy of your findings, which will improve their reliability and credibility.
Embrace transparency. Include an ethics appendix in your final publication that discusses any special ethical issues or considerations that arose over the course of your research. This enhances the replicability of your study findings by providing other researchers with better insights into the participants' interactions with your study. Discussing ethical issues explicitly in paper appendices can also help clarify and improve ethical norms, as Asiedu, Karlan, and Udry (2021) suggested.

Conclusion

Through our experience and those of other organizations (cf MacPhail et al., 2013; Iguna et al., 2022; Wein et al., 2023), we have come to understand research with humans as an ongoing relationship between the researcher and the participants. When the researcher and researched share context and expectations, this relationship is easy to manage. However, the greater the chasm in context, the more actively that relationship must be managed, lest the research brings about unintended social harms due to misunderstandings resulting from the context gap and expectations. Sometimes, third parties that have a foot on both sides of the context chasm are necessary to eliminate these understandings. This is a role that Busara often plays – but only because we have learned through hard experience the harms that result when the chasms that separate different worlds are not effectively bridged.

References

Asiedu, E., Karlan, D., Lambon-Quayefio, M., & Udry, C. 2021. A call for structured ethics appendices in social science papers. *Proceedings of the National Academy of Sciences* 118(29), e2024570118. https://doi.org/10.1073/pnas.2024570118

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, Working Paper 28106. https://doi.org/10.3386/w28106.

Iguna, S., Getahun, M., Lewis-Kulzer, J., Odhiambo, G., Adhiambo, F., Montoya, L., ... & Camlin C.S. 2022. Attitudes towards and experiences with economic incentives for engagement in HIV care and treatment: qualitative insights from a randomized trial in Kenya. *PLOS Global Public Health* 2(2), e0000204. https://doi.org/10.1371/journal.pgph.0000204

MacPhail, C., Adato, M., Kahn, K., Selin, A., Twine, R., Khoza, S., ... & Pettifor, A. 2013. Acceptability and feasibility of cash transfers for HIV prevention among adolescent South African Women. *AIDS and Behavior* 17, 2301-12. https://doi.org/10.1007/s10461-013-0433-0

Schmidt, M. 2022. The gift of free money: on the indeterminacy of unconditional cash transfers in western Kenya. *Journal of the Royal Anthropological Institute* 28(1), 114-29. https://doi.org/10.1111/1467-9655.13655

Wein, T., Lanthorn, H., & Fischer, T. 2023. First steps toward building respectful development: three experiments on dignity in aid in Kenya and the United States. *World Development Perspectives* 29, 100485. https://doi.org/10.1016/j.wdp.2023.100485

Why applied psychologists should consider their work's value-laden context

Jason M. Chin (Australian National University)

This essay argues that applied psychologists should openly consider the broader context of their work and, in particular, the consequences of being mistaken. More specifically, they should transparently consider the consequences of erroneously declaring that their work supports or rejects a hypothesis. That is, they should openly debate their work's 'inductive risk'. This involves thinking beyond psychological conventions (for example, alpha = .05), to the risks specific to their work's broader context. After a brief review of inductive risk, I apply this background to a case study of a large government-funded legal psychology study in which external inductive risks were particularly salient and tractable. Yet, the gravity and facileness of internal psychological customs were apparently too strong. The researchers failed to think about the broader consequences of their research. This does not bode well for research with less obvious societal implications. In the conclusion, I zoom out to consider other ways in which researchers might need to consider the broader context of their work.

Non-epistemic values, inductive risk, and transparency

Scientists, especially applied scientists, play a special role in society. They are seen to possess specialized knowledge, which grants their claims extra authority. This authority leads to the expectation that their statements will 'have an impact, to be believed (by at least some) and (potentially) to serve as a basis for action' (Douglas, 2017, p. 87). As a result, they should take care when conducting and reporting their research and think about the effects of their work. Beyond these role responsibilities per se, scientists should consider the costs of being mistaken. Because not all scientific claims carry the same consequences, applying a uniform standard of evidence may not be appropriate merely because that is what is conventionally done in a field (Douglas, 2017, p. 89). These considerations have led some philosophers of science to say that researchers have a moral obligation to consider the inductive risk involved in their work (see Elliot, 2017).

If scientists knowingly incorporate non-epistemic values into their work (as they should), they should also be clear about this when they report their work. This promotes public accountability (Elliot, 2017; Douglas, 2017). However, some researchers may deny that non-epistemic factors inform their work. These researchers should still be transparent about

how their research is conducted, detailing their processes and assumptions transparently, which would allow 'democratic decision makers to make a justified decision given the best and most comprehensive scientific understanding of the system at hand' (Betz, 2017, p. 101). In other words, the research users can use this transparency to determine if the research is safe for their purposes.

A publicly-funded study to determine the effect of joining child sexual assault trials

To better understand these principles, I will apply them to a consequential law and psychology study (see reviews and critiques of this study in Hunter & Kemp, 2017; Robinson, 2017; Chin et al., 2022; Chin, 2023). It occurred in the context of a broader inquiry conducted by the Australian Government in response to widespread concern about the prevalence of child sexual abuse. This inquiry, the Royal Commission into Institutional Responses to Child Sexual Abuse (RCIRCSA), lasted approximately five years. It ended in 2017 (https://www.childabuseroyalcommission.gov.au/) and cost approximately half a billion Australian dollars (Wright et al., 2017). Law reform based on its findings is ongoing, with Australian jurisdictions continuing to consider and implement its recommendations (Hamer, 2021). One of the RCIRCSA's questions was how to prosecute child sexual abuse more effectively. These proceedings often run into difficulties because the assaults occur in private, and the prosecution occurs decades after the alleged acts. Lengthy individual trials can cause delays (RCIRCSA, 2017, p. 634). As a result, the RCIRCSA was interested in exploring rules that would enable more joint trials, allowing the court to hear multiple allegations at once, thus giving the judge or jury additional evidence. Joint trials could also be more efficient. On the other hand, joint trials may present prejudice to the accused, making the case against them seem stronger than it was and causing the factfinder to conflate various strands of evidence. These questions inspired our case study, a large mock jury study examining, among other things, the effect of joining trials versus hearing the same allegations in separate trials (Goodman-Delahunty et al., 2016). The study found that juries in the joint trial condition were convicted somewhat more than in the single trial condition (about 15% depending on the scenario, see ibid., p. 194). However, there was no statistically significant difference. This contributed to the authors' conclusion that 'no joinder effect was found' (ibid., p. 194). Accordingly, they determined that 'there is negligible risk to the defendant of a conviction based on reasoning logically unrelated to the evidence' (ibid., p. 265).

Non-epistemic values, inductive risk, and transparency in our case study

I will now evaluate this study based on the principle that researchers ought to consider the inductive risk related to their work openly. As I have noted, one way of managing inductive risk is to set standards of evidence sufficiency, such as the alpha level (that is, the probability of observing a positive result when the null hypothesis is true) and the beta level (that is, the probability of observing a negative result when the null is false). Psychologists tend to set their alphas at .05 and betas at .20. This is based on internal psychological customs rather than broader considerations about the meaning of the study in the wider world. Those who developed the statistical frameworks that psychologists now use did not intend this:

> These two sources of error can rarely be eliminated completely; in some cases, it will be more important to avoid the first and, in others, the second. We are reminded of the old problem considered by Laplace, which is the number of votes in a court of judges that should be needed to convict a prisoner. Is it more serious to convict an innocent man or to acquit a guilty? That will depend upon the consequences of the error, whether the punishment is death or a fine, the danger to the community of released criminals, and the current ethical views on punishment. From the point of view of mathematical theory, all we can do is show how the risk of the errors may be controlled and minimized. The use of these statistical tools in any given case, in determining just how the balance should be struck, must be left to the investigator. (Neyman & Pearson, 1933, p. 296)

Psychology's convention to set beta at .20 likely comes from Jacob Cohen, who suggested it as a default on the premise that Type I errors are generally four times more serious than Type II errors (i.e., .20 / .05 =4). He added that his suggestion should be ignored whenever there was a reason to set the error based on the research context (Cohen, 1988). Philosopher of science Heather Douglas echoes this reasoning: 'Having a flexible standard (that one must have some good evidence, but exactly how strong varies) allows scientists to consider the context in which they are doing their work and its potential implications concerning both false positive and false negatives' (2017, p. 89).

When research, psychological or otherwise, is divorced from obvious real-world contexts, it may be difficult to think about the consequences of erroneously claiming a hypothesis was supported or rejected. However, as Neyman and Pearson indicate in the quote above, sometimes non-epistemic justifications will be more tractable and, consequently, articulable. Our case study, for instance, was commissioned to inform a change in the law. As a result, many external considerations at play are not based on psychological conventions. Some consequences of error, for instance, are laid out in Table 1. That table shows that by mistakenly claiming that the data does not support a joinder effect, society would lose the chance to obtain more correct convictions and have more efficient trials. It also shows that more wrongful convictions may result by mistakenly claiming there was no joinder effect.

ERROR	CONCLUSION	GROUND TRUTH	RECOMMENDATION TO LAWMAKERS	ERROR CONSEQUENCE
False positive (Type I)	Joinder effect exists	No joinder effect	Retain current rules preventing joint trials	Lost chance of more correct convictions and efficient trials
False negative (Type II)	No joinder effect	Joinder effect exists	Change law to facilitate joint trials	Increased wrongful convictions

TABLE 1.

The reasoning in Table 1 is not present in our case study. In fact, the authors were not very clear in their justifications for many of their statistical choices. However, they did explain why they chose their alpha level: 'To assess the effects of trial type on jury decisions in this study, a formal power analysis determined the sample size; the effect size was determined based on the magnitude of effects observed in past studies; and the significance level, as is customary in social scientific practice, was set at 95%' (Goodman-Delahunty et al., 2016, p. 20). In other words, they relied on customs rather than considering any external reasons for their study's error tolerances.

The authors' unwillingness to consider anything beyond custom inspires a potentially useful thought question: What would attentiveness to non-epistemic matters, specifically the cost of error, look like if the authors had turned their minds to it? To begin with, the study's planners might have recognized, as Neyman and Pearson did in the block quote above, that an inductive error could lead to the convictions of innocent people. Specifically, failing to find evidence against the null when the null should be accepted leads to the recommendation of changing the laws related to joinder (a Type II error). The authors might then consider what error rate is acceptable here. If they mistakenly accept the null, they will recommend a law that can produce wrongful convictions. So, how sure do they want to be? There are many justifiable approaches to this question, such as looking at public opinion surveys about how seriously the public takes wrongful convictions or asking the Royal Commission to hold a roundtable with stakeholders to discuss this.

Then, the authors should reason that this question of error requires grounding in the magnitude of the error. In other words, if no statistically significant effect is found, we would like to know both the Type II error rate and the effect size that we can now rule out with that error rate. Fortunately, routine statistical reasoning girds this thought process: 'When designing an experiment, one goal might be to choose a sample size that provides a desired Type II error rate for an effect size of interest. This can be achieved by performing an a-priori power analysis' (Lakens, 2022). In other words, the authors will want to consider what effect size the Type II error rate should be associated with. For example, they may set their beta at .05 and effect size to equal 5% more convictions due to joinder. This allows the inference that if the joinder effect is 5% or greater, the maximum Type II error rate is 5%. The effect size I have described here is the 'smallest effect size of interest' or the smallest effect that is practically or theoretically worth finding (Anvari & Lakens, 2021).

Determining the smallest effect size of interest can be very difficult, especially when research is divorced from practical implications (Riesthuis et al., 2021). But, in the context of our case study, it is both a tractable question and one that the researchers had an obligation to undertake due to the authority and potential political and legal implications of their research (Douglas, 2017). Once again, the authors could have turned to the legal stakeholders, asking them what differences in joint versus individual trials (i.e., wrongful convictions), all else equal, are socially justifiable in light of the advantages of joint trials. Is it 10% more

convictions, 5%, or something else? Perhaps the stakeholders would return that there is no difference that is acceptable. In that case, the case study should be aborted because running a study powered to find such a minute difference would never be practical. In any case, these are moral and democratic questions that require contemplation.

This was not the path the case study authors followed. As noted, they relied on psychological customs to determine their error rates. To determine the effect size associated with that error rate, they chose the 'magnitude of effects observed in past studies' (Goodman-Delahunty et al., 2016, p. 20). This is both a custom within psychology but also a justification inapposite to the goal of designing a study that can inform the law. In other words, the authors may be able to articulate a Type II error rate associated with the effect sizes found in prior studies. But that error rate has nothing to do with the safety or unsafety of joining trials. It only speaks to the prior literature.

Finally, Douglas and others say that to be fully accountable, researchers should make clear to users the non-epistemic values they use in their research or, at least, fully detail their analysis. In our case study, transparency could be advanced by more thorough reporting in several ways. For instance, the authors could conduct the above power analysis and report the results and analysis code used to produce it. They could also engage with stakeholders and transparently report the elicitation processes they used, such as the IDEA protocol (Hemming et al., 2017). Alternatively, they may have to report that they did not conduct the study with a meaningful error rate in mind, so the null result is uninterpretable. All of these approaches would advance transparency.

As may be clear, the authors were not transparent in their reasoning. Rather, they performed a 'formal power analysis' (Goodman-Delahunty et al., 2016, p. 20) based on an effect size from previous studies. However, they did not provide that analysis in their report or make their analytic code or data openly available. They also did not say what the effect size they used for that calculation was. Without this information, users of the research cannot assess it to determine the risks involved in applying their decision to change the laws in their jurisdictions.

Beyond inductive risk

The case study above demonstrates the gravity of custom and how it can draw attention away from critical thinking about a study's context. Specifically, the authors of the RCIRCSA study followed entrenched psychological thinking about error rates when designing their study. They did this even though the error consequences in their study were much more salient than in standard applied psychological work. Ironically, one potential consequence, namely wrongful convictions, in that study is the same one Neyman and Pearson used to illustrate their unease with a uniform standard.

But, while inductive risk is one area in which applied researchers seem unwilling to broaden

their thinking, it is far from the only one. This essay's analysis could easily be applied to several other studies in which researchers robotically followed the epistemic customs of their field (see, for example, Chelwa & Muller, 2019 who describe randomized controlled trials that, among other things, do not manage the risks associated with the research or consider whether any knowledge gained will generalize such that the benefits of the research outweigh the costs). Researchers – applied ones especially – should be more willing to engage with the broader literature surrounding the values inherent in the research and reporting process. They should embrace the 'social and ethical priorities' inherent in their work (Elliot, 2017, p. 1). And they should do so openly and transparently.

References

Anvari, F., & Lakens, D. 2021. Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology* 96, Article 104159. https://doi.org/10.1016/j.jesp.2021.104159

Betz, G. 2017. Why the argument from inductive risk doesn't justify incorporating non-epistemic values in scientific reasoning. In K. C. Elliot & D. Steel (eds), *Current controversies in values and science*. Taylor & Francis, pp. 94-110. https://doi.org/10.4324/9781315639420

Chelwa, G., & Muller, S. 2019. *The poverty of poor economics*. Africa is a Country. https://africasacountry.com/2019/10/the-poverty-of-poor-economics

Chin, J.M. 2023. Law and psychology must think critically about effect sizes. *Discover Psychology* 3(3). https://doi.org/10.1007/s44202-022-00062-2

Chin J.M., Nakagawa S., & Lagisz M. 2022. Where is the evidence in evidence-based law reform? *University New South Wales Law Journal* 45(3),1124-53. https://doi.org/10.53637/FMHK9486

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Douglas, H. 2017. Why inductive risk requires values in science. In K. C. Elliot & D. Steel (eds), *Current controversies in values and science*. Taylor & Francis, pp. 81-93. https://doi.org/10.4324/9781315639420

Elliot, K.C. 2017. *A tapestry of values*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190260804.001.0001

Goodman-Delahunty, J., Cossins, A., & Martschuk, N. 2016. *Jury reasoning in joint and separate trials of institutional child sexual abuse: an empirical study*. https://www.childabuseroyalcommission.gov.au/sites/default/files/file-list/Research%20Report%20-%20Jury%20reasoning%20in%20joint%20and%20separate%20trials%20of%20institutional%20child%20sexual%20abuse%20An%20empirical%20study%20-%20Government%20responses.pdf

Hamer, D. 2021. Myths, misconceptions and mixed messages: an early look at the new tendency and coincidence evidence provisions. *Criminal Law Journal* 45, 232-52.

Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., & Wintle, B.C. 2017. A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution* 9(1), 169-82. https://doi.org/10.1111/2041-210X.12857

Hunter, J., & Kemp, R.I. 2017. Proposed changes to the tendency rule: a note of caution. *Criminal Law Journal* 41, 253-60.

Lakens, D. 2022. *Improving your statistical inferences*. https://doi.org/10.5281/zenodo.6409077

Neyman, J., & Pearson, E.S. 1933. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 231(694-706), 289-337. https://doi.org/10.1098/rsta.1933.0009

Riesthuis, P., Mangiulli, I., Broers, N., & Otgaar, H. 2022. Expert opinions on the smallest effect size of interest in false memory research. *Applied Cognitive Psychology* 36(1), 203-15. https://doi.org/10.1002/acp.3911

Robinson, P.M. 2017. Joint trials and prejudice: a review and critique of the report to the Royal Commission into Institutional Child Sex Abuse. *Monash University Law Review*, 43 (3), 723-59.

Royal Commission into Institutional Responses to Child Sexual Abuse. 2017. *Criminal justice report*, parts III-VI. https://www.royalcommission.gov.au/system/files/2021-01/criminal_justice_report_-_parts_iii_to_vi.pdf

Wright, K., Swain, S., & McPhillips, K. 2017. The Australian Royal Commission into Institutional Responses to Child Sexual Abuse. *Child Abuse & Neglect* 74, 1-9. https://doi.org/10.1016/j.chiabu.2017.09.031

Cultural context and ecological validity in global development research

Symen Auke Brouwers (Independent Researcher, Potchefstroom, South Africa) and Floriza Freire Gennari (George Washington University)

Aid effectiveness and sustainability in global development have been debated extensively in recent decades. Negative perspectives have ranged from critiques about the ineffectiveness of foreign aid for global development (Williamson, 2008) to observations about its potential harms, such as the deterioration of social cohesion or indigenous cultural identities (MacLachlan et al., 2010). Others posit that, while valid, global development programs could benefit from long-term sustainability considerations, greater participation of the communities receiving the interventions, and overall, closer assessment and accountability (Bhutta & Aleem, 2013; Biesma et al., 2009). At the same time, the economists Abhijit Banerjee and Esther Duflo received a Nobel Prize in Economic Sciences for their contributions to poverty alleviation research, signaling the value society places on the specification of aid effectiveness, preferably quantified (IPA, 2019). This chapter evokes the principle of contextualization and the related need for ecological validity to advance research that carefully assesses the nature and extent of community development challenges and the impact of interventions targeting those challenges. By following the notion of contextualization, global development programs may be more applicable, inclusive, and ultimately more sustainable and effective.

Contextualization is a principle from the fields of cultural and cross-cultural psychology that posits a close connection between, on the one hand, behavior and mental or psychological processes (for example, emotion, cognition, decision-making, knowledge, beliefs, attitudes, norms, and values), and, on the other hand, the socio-cultural context in which they develop (Brouwers & Van de Vijver, 2015). In terms of its reach, contextualization is placed between the theoretical perspectives of relativism (assuming that processes and mechanisms are always unique to specific cultures, giving rise to its entire incarnation) and universalism (assuming that processes and mechanisms are always shared across cultures with context only affecting its shape or strength). For instance, some researchers subscribing to universalism posit that while children may receive more stimulation in certain cultures than others, they will likely all reach the same cognitive milestones. Contextualization is particularly important when one considers research that attempts to extract 'decontextualized' processes and mechanisms from the human behavioral repertoire, such as intelligence quotient

(IQ). Similar to child development, proponents of IQ will say that psychological processes develop in the same abstract way regardless of socio-cultural context and preferred ways and strategies of processing information.

Ecological validity, a concept closely related to contextualization, refers to the extent to which project designers' understanding of problems and solutions match reality within the receiving context (Bronfenbrenner, 1977). In other words, it assesses the extent to which a stimulus validly reflects people's feelings and skills in that context. In cross-cultural research, the focus has mostly been on the inappropriate use of stimuli within measurement tools and the assessment protocols. A study by Jean Lave (1997 [1980]) in Liberia illustrates the potential danger of inappropriate stimuli usage. She compared the maximum cognitive performance of unschooled tailors with adults who had received formal schooling. The findings showed that tailors performed better on tasks that included tailoring stimuli, while the schooled adults performed better on tasks that included formal schooling stimuli. Failing to adapt the measurement tools or protocols could have led to the erroneous assumption that tailors without formal schooling had less cognitive capacity than adults who had received formal schooling. Greenfield (1997) uses the phrase 'you can't take it with you' ('it' meaning ability assessments) to describe this fundamental assessment problem across cultures; social and cultural experiences shape the meaning people create.

> **Rather than decontextualizing and isolating a problem or barrier, preparation for an intended program should start with a critical, systemic analysis of the problem and its associated elements.**

Entrenchment in context

The foundation for the idea of contextualization is clearly found in the ecological systems theory of Bronfenbrenner (1977; also see Berry et al., 2002), also known as the socio-ecological model. The model espouses a central position of the human individual with their intrapersonal mental states, knowledge, and beliefs, surrounded by ever-increasing circles that capture widening ecological systems: firstly, a microsystem with peers, family, community, secondly a mesosystem for comparing their own peers, families, church, and school with others, thirdly an exosystem with broader social institutions such as politics, neighborhoods, and media, and finally a macrosystem with values, attitudes, and 'culture.' The model's simplicity is very illustrative of the position of individual human beings in larger systems, their sources of influence, and the many elements that might be relevant to their behaviors (Brouwers, 2018). It is also noteworthy that the many levels of influences shaping the behaviors of an individual will differ widely by setting and cultural context, thus requiring the 'contextualization' of measures and tools for understanding them.

The theory of generative entrenchment posits that elements of behavior become robust when used to construct other behaviors (Wimsatt & Griesemer, 2007). An example of generative

entrenchment is the construction of railroads in the early 1900s in Ghana during the British colonial past and its impact on economic activity (Jedwab & Moradi, 2016). In the early 1900s, Ghana, like the rest of Sub-Saharan Africa, did not have any modern transportation technology. Using economic and geographic data at eight-time points from 1891 to 2000, Jedwab and Moradi show that the building of railroads by the British colonizers did not just reorganize existing economic activity but created new ways of economic trade, followed by new agricultural activities. Railroads led to the location of new cities. When the railroads and road networks fell into disarray after independence from the British in 1960, the effects of railroad construction on economic activity persisted. In economic theory, this is called path dependence; that is, economic investment in a particular area leads to a particular development trajectory. Using Wimsatt's terminology, we can say railroads are scaffolds that, by generating new adaptations, become entrenched in the Ghanaian economy.

Rather than a person being passive and simply receiving culture, they are conceived as fully in control, selecting objects they need from their sociocultural and physical environment while ignoring and rejecting others. According to Wimsatt, more commonly used elements will become differentially more entrenched than parts that are not. Wimsatt and Griesemer (2007) call this process scaffolding, employing historical and social institutions and ideational entities as building blocks for new behavior (Wimsatt, 2013). Scaffolds create trajectories of development through nearby points of stability by (a) lowering fitness barriers to existing performance or achievements, (b) making accessible new competencies that become the self-scaffolding of later development, and (c) the creation of dependencies through the increased use of these competencies (Bickhard, 1992). Renewal, maintenance, and resistance against change are inherent concerns in this process: scaffolds provide space for new elements to be held, but at the same time, they may resist something when it threatens other competencies resting on that scaffold.

Using Wimsatt's theory of generative entrenchment, we can say that a global development intervention, behavior, or activity newly introduced into a community will be successful when it can be attached to an existing scaffold – and later become a scaffold itself. The step from leaning on a scaffold to turning into a scaffold carries significant implication: a newly introduced behavior requires generalization of this one behavior to all behaviors that lean on it, thus implying transmission from one individual to another as others recognize the new behavior as successful and come to use it as a scaffold for themselves, as well as the societal mobility of the individual as more and more of his compatriots share the same successful behaviors as him or her. This requires attention to a constellation of cultural objects, not only the behavior of interest, but anything it rests on and anything that rests on it. For instance, changing the dietary habits of infants and children – even if only concerning a specific group – requires an understanding and appreciation of the entire sociocultural and economic ecosystem related to growing, acquiring, preparing, and consuming food, including the different individuals, interpersonal and community roles involved, the contexts in which different foods are offered, why they are offered, and by whom.

Creating ecological validity

A careful examination and analysis of the principles of contextualization and ecological validity, as they are meant to be fully understood, lead us to claim that a good match between global development programs and existing cultural objects and practices is often missing. Programs are frequently developed far from the community in question, using broad standards and objectives, non-contextualized measures of success, and, for the most part, target specific, individual-level behaviors only. The initial positive effects of a program not rooted in contextualization and ecological validity might fade after the implementing agency leaves. Further, it is theoretically unclear how self-organization among people in low-resource communities is affected by outside funding (Gugerty & Kremer, 2008). Indeed, externally imposed programming may threaten accepted social community participation and inclusivity (Ager et al., 2010), including weakening self-organized criticality that typically exists in societies (Brunk, 2002). However, as argued here, biases in the underlying research and the measures used may have simply resulted in a misunderstanding about the target behavior and the constellation, or system, within which it occurs (Norenzayan & Heine, 2005): the language used in assessment can inhibit this broader perspective. The desired behavioral response may be entirely unavailable in the community where the study is set because the proposed behavior is alien to that cultural context, and desired behaviors function differently.

> **A valuable lesson that can be drawn from cross-cultural psychology about the research, design, and implementation of development programs is that these are rarely universal and that placing the individuals we aim to serve at the center of interventions before, during, and after implementation is essential for success**

There are clear, practical applications for global development research and programs once we move beyond the tendency to isolate differences in our research and start attending to context: the presence of variations in many cultural objects points to 'local symbiosis' in which most of the social, psychological, and physical elements work together with individuals to create a traversable habitat with infrastructure and socially accepted ways of living together. Existing, observable behavior is never maladaptive, which we might assume when we examine it with our own cultural lens of what is good and bad, what is acceptable, but rather always an adaptation in a given context.

If the generative entrenchment theory teaches us anything at all, it is that people resist change. A particular constellation of various cultural objects in a local habitat guides the absorption and retention of new, alien cultural objects that might follow from global development aid and, at the same time, provides protection from threats that might be perceived as arising from alien cultural objects. For example, in wanting to better the inclusion of children with disabilities in schools, strategists are faced with some very

different cultural objects and mechanisms that, despite their differences, work together to keep children with disabilities out of school, such as beliefs that the origin of disability lies in misbehavior of the mother or the family of the mother, exclusion of fathers from critical decision-making mechanisms in their community, limited competence of teachers to deal with children with disabilities, anxiety of mothers that their children with disability will be mistreated, and bullying, all conspire to resist any change in the status quo (Brouwers, 2021). For strategists to simply address one problem but not any others might not lead to the intended effect of disability inclusion. Any powerful constellation of cultural objects can dramatically impact aid effectiveness; because of elements in the culture that cause friction, aid activities might be distorted or lose their intended meaning and thus miss their desired contribution. It is thus essential to develop a comprehensive, qualitative understanding of each context's constellation or system.

Rather than decontextualizing and isolating a problem or barrier, preparation for an intended program should start with a critical, systemic analysis of the problem and its associated elements. This entails the identification of scaffolds and other behaviors that depends on identifying the one behavior that outsiders might perceive as a problem. It is true that while such thoughtful formative research may require more time and resources than available in emergency situations and in responses to certain threats, such as during the COVID-19 pandemic, a more culturally-sensitive approach may be more cost-effective long-term, resulting in research tools and interventions with a longer shelf-life than is typically seen. Further, once good contextualization is created, researchers can apply the measurement tools and interventions with many more people in the community, rather than only with those meeting specific criteria. In effect, over time, high-quality material will be available for building and re-tooling interventions, like a doctor's bag full of timely and useful supplies. The overarching approach for arriving here should be people-centered and based on empathy and cultural humility. The aim should be to understand the lived experiences of the individuals whose behaviors we seek to change, and thus the rationale behind the behavioral scaffolds. Conducting thoughtful, formative research alongside local partners and stakeholders can provide valuable insights, supplemented with knowledge about psychological and behavioral mechanisms in academic literature. At this stage, the analysis should be participatory, observational, and where traditional qualitative methods are used, favor non-leading and open-ended questions. For example, in a recent attempt by UNICEF Eastern & Southern Africa to better understand the different social norms related to the possibility of children with disabilities going to school, parents, teachers, school directors and social workers were asked questions along the lines of: 'What do people say are the disadvantages of keeping children with a disability out of school?', and 'Who is important to consider when thinking or talking about keeping children with disabilities out of school? Why so?' (Brouwers, 2021). Such types of questions seek a connection with the targeted problem (inclusivity in school) with larger mechanisms that are at play in the community (beliefs, embarrassment, decision authority, fear of harm). Once behaviors of interest have been mapped out with their influences, behavioral scaffolds, and constellation well-

understood, a contextualized intervention ideally includes the co-development of prototypes, or interventions, that stem from the community members themselves.

Conclusion

Critical analysis of context – in terms of the ways in which the targeted behavior depends on other behaviors or the ways other behaviors depend on the targeted behavior – has the potential to provide deep insights about individual and social change. Together with the optimization of ecological validity, obtained through the well-executed fusions of qualitative and quantitative research techniques, this approach can contribute to a much more accurate assessment of change from baseline through program implementation. With research that truly reflects the mechanisms of behavior change in communities, recommendations can be made that bring us closer to the outcomes desired by both the community and program implementers.

A valuable lesson that can be drawn from cross-cultural psychology about the research, design, and implementation of development programs is that these are rarely universal and that placing the individuals we aim to serve at the center of interventions before, during, and after implementation is essential for success. The demands placed on programs and the barriers towards success can be monumentally specific to a context, such that using non-contextualized or non-ecologically valid interventions even within the same region or country can backfire. The meaningful participation of recipients is thus essential to ensure that interventions align with the existing constellation or system of behaviors. The bottom line is to never assume you can simply 'take it with you' – be it a measure or program that was successful somewhere else – and apply it where you are working now; thorough contextualization is key.

References

Ager, A., Stark, L., Akesson, B., & Boothby, N. 2010. Defining best practice in care and protection of children in crisis-affected settings: a Delphi study. *Child Development* 81(4), 1271-86. https://doi.org/10.1111/j.1467-8624.2010.01467.x

Berry, J.W., Poortinga, Y.H., Segall, M.H., & Dasen, P.R. 2002. *Cross-cultural psychology: research and applications* (2nd ed.). Cambridge University Press.

Bhutta, Z.A., & Aleem, S. 2013. *A review of external assistance and aid effectiveness for maternal and child health: challenges and opportunities*. United Nations University – World Institute for Development Economics Research, Working Paper No. 2013/099.

Bickhard, M.H. 1992. Scaffolding and self-scaffolding: central aspects of development. In L. T. Winegar & J. Valsiner (eds), *Children's development within social contexts: research and methodology*. Lawrence Erlbaum, pp. 33-52.

Biesma, R.,G., Brugha, R., Harmer, A., Walsh, A., Spicer, N., & Walt, G. 2009. The effects of global health initiatives on country health systems: a review of the evidence from HIV/AIDS control. *Health Policy Plan* 24(4), 239-52. https://doi.org/10.1093/heapol/czp025

Bronfenbrenner, U. 1977. Toward an experimental ecology of human development. *American Psychologist* 32(7), 513-31. https://psycnet.apa.org/doi/10.1037/0003-066X.32.7.513

Brouwers, S.A. 2018. The positive role of culture: what cross-cultural psychology has to offer to developmental aid effectiveness research. *Journal of Cross-Cultural Psychology* 49(4), 519-34. https://doi.org/10.1177/0022022117723530

Brouwers, S.A. 2021. *The role of social norms in decisions to provide schooling to children with disabilities*. UNICEF ESARO SBC. Downloaded from: https://www.unicef.org/esa/reports/social-norms-disability

Brouwers, S.A., & Van de Vijver, F.J.R. 2015. Contextualizing intelligence in assessment: the next step. *Human Resource Management Review* 25(1), 38-46. https://doi.org/10.1016/j.hrmr.2014.09.006

Brunk, G.G. 2002. Why do societies collapse? A theory based on self-organized criticality. *Journal of Theoretical Politics* 14(2), 195-230. https://doi.org/10.1177/0951692802014002003

Greenfield, P.M. 1997. You can't take it with you: why ability assessments don't cross cultures. *American Psychologist* 52(10), 1115-24. https://psycnet.apa.org/doi/10.1037/0003-066X.52.10.1115

Gugerty, M.K., & Kremer, M. 2008. Outside funding and the dynamics of participation in community associations. *American Journal of Political Science* 52(3), 585-602. https://doi.org/10.1111/j.1540-5907.2008.00331.x

IPA. 2019. *IPA congratulates 2019 Nobel Prize winners Abhijit Banerjee, Esther Duflo, and Michael Kremer*. https://poverty-action.org/ipa-congratulates-2019-nobel-prize-winners-abhijit-banerjee-esther-duflo-and-michael-kremer

Jedwab, R., & Moradi, A. 2016. The permanent effects of transportation revolutions in poor countries: evidence from Africa. *Review of Economics and Statistics* 98(2), 268-84. https://doi.org/10.1162/REST_a_00540

Lave, J. 1997 [1980]. What's special about experiments as contexts for thinking? In M. Cole, Y. Engeström & O.A. Vasquez (eds), *Mind, culture, and activity: seminal papers from the Laboratory of Comparative Human Cognition*. Cambridge University Press, pp. 57-69.

MacLachlan, M., Carr, S.C., & McAuliffe, E. 2010. *The aid triangle: recognizing the human dynamics of dominance, justice and identity*. Zed Books.

Norenzayan, A., & Heine, S.J. 2005. Psychological universals: what are they and how can we know? *Psychological Bulletin* 131(5), 763-84. https://psycnet.apa.org/doi/10.1037/0033-2909.131.5.763

Van de Vijver, F.J.R., & Leung, K. 1997. *Methods and data analysis for cross-cultural research*. Sage.

Williamson, C.R. 2008. Foreign aid and human development: the impact of foreign aid to the health sector. *Southern Economic Journal* 75(1), 188-207. https://doi.org/10.1002/j.2325-8012.2008.tb00898.x

Wimsatt, W.C. 2013. Entrenchment and scaffolding: an architecture for a theory of cultural change. In L. Caporael, J. Griesemer & W.C. Wimsatt (eds), *Developing scaffolding in evolution, cognition, and culture*. MIT Press, pp. 77-104.

Wimsatt, W.C., & Griesemer, J.R. 2007. Reproducing entrenchments to scaffold culture: the central role of development in cultural evolution. In R. Sansome & R. Brandon (eds), *Integrating evolution and development: from theory to practice*. MIT Press, pp. 228-323.

Consent, open-ended questions, and feedback loops: empirical insights into research ethics in the Global South

Joel Wambua (Busara)

The history of research ethics is a journey marked by both progress and setbacks, shaped by ethical dilemmas, scientific advancements, and societal changes. Various international organizations, regulatory bodies, and institutions have established guidelines, codes of conduct, and oversight mechanisms to ensure the ethical conduct of research. These include the establishment of Institutional Review Boards (IRBs), informed consent procedures, and standards for data management and privacy.

Despite these measures, there has been considerable criticism of how global development research is carried out. Prime among these criticisms are those that focus on the role of experimentation. Although experiments are no doubt extremely useful for generating knowledge (Falk & Heckman, 2009; Duflo & Banerjee, 2013), critics have raised doubts about both the efficiency and the ethics of over-relying on experiments for knowledge. In fact, the March 2020 issue of World Development (Volume 127), devoted an entire special section to these issues, including important critiques from Hoffmann (2020) and Kaplan et al. (2021). These criticisms extend beyond the methodological cases made by Deaton and Cartwright (2018) and Pritchett (2018). Though sometimes couched in ethical terms, these older criticisms mostly focused on misallocating resources and problems of causal inference. More recent criticisms have centered on the power imbalances present in much research and how these imbalances are reinforced (Amarante et al., 2021). They argue that much of the research conducted in the Global South replicates and perpetuates colonial relationships. They note that almost all this research's money, theory, and hypotheses come from the West. Arguing that the recommendations and findings generated by this work are insufficiently grounded in the realities and experiences of those in the Global South – and thus have limited validity and value to the people and places they ostensibly seek to serve – critics point out that the knowledge produced might serve Western academic career advancement more than actual development policy.

Catalyzed by these discussions, many researchers are beginning to realize the importance of upholding a higher standard of ethical practice in research. As Crow et al. (2007) put

it, a good ethical practice enhances the quality of data collected because when research is genuinely inclusive, it better represents the voices and worlds of those we seek to study (Cornwall & Jewkes, 1995). Furthermore, when research is conducted poorly, it can harm research participants and research staff in the field (Steinert et al., 2021). As Hoffmann (2020) makes clear, if a research project cannot be conducted ethically, with accountability structures to ensure it is, it should not go ahead.

Several researchers whose work has been founded on experimentation have responded to these criticisms with suggestions for more ethical conduct (Humphreys, 2020). Glennerster and Powers (2016) highlight how the US Belmont Principles of ethical research, such as evaluating the risks and benefits of a study, can be applied to the particular challenges of experimental research in the Global South. Asiedu et al. (2021) seek to ensure we pay attention to these concerns with a call for more comprehensive reporting of ethical considerations in paper appendices. Groves Williams (2016) charts the many inconsistencies in approaching this problem. Only Cronin-Furman and Lake (2018) move beyond the standard Belmont Principles to examine what research for equity would look like (specifically relating to conflict research). Still, their recommendations have wider applicability for research in the Global South. While many researchers agree on the importance of avoiding unethical research, meaningful ethical practice is hard to implement. Ethical practice easily lapses into a box-ticking exercise, partly due to the technocratic structure of the measures put into place to protect participants. Many of the ways researchers protect participants have little regard for the actual preferences of those participants and are instead governed by far-off IRBs, following principles focused on and articulated by actors in the Global North.

Moving towards empirical research ethics

Regarding the specific debates about experimental research, the criticisms and responses rarely feature the participants' voices. However, there has been a growing recognition among researchers of the significance of involving participants in ethical deliberations. (Wambua et al., 2020). As we explore an inclusive approach to raising ethical standards in research, there are key questions we need to ask ourselves. How do participants perceive the constructs and principles that underlie ethics, such as dignity and respect (Wein, 2020)? Are the options suggested by various scholars well-aligned with the interests of those they seek to protect? We conducted an empirical research study to prioritize the various ameliorative options based on the input of research participants with these questions in mind. We sought to generate a preference-driven empirical basis for our ethics-focused practices centered on those who participate in social science research. We sought to empirically test the relative effectiveness of changes to research protocols aimed at improving participants' well-being and upholding higher standards of ethical practice. To find out what problems we should focus on to make our research more ethical and respectful, we conducted a qualitative study (Nyaga et al., 2022) to understand the experiences and preferences of our research participants in low-income areas in Nairobi. In this study, we learned that participants want to be involved in research meaningfully and want researchers to share results and

improve consenting processes. Based on these findings, we sought to empirically test the relative effectiveness of revised research protocols and forms of participant engagement. Specifically, we focussed on three aspects of the research that can be modified to improve participant experience: the consenting process, the inclusion of qualitative questions in the research design, and the addition of procedures to provide participants with feedback about the outcomes of the research process.

Modifications to consent

In previously conducted qualitative research, our participants called for better consent practices (Nyaga et al., 2022). Informed consent entails that a prospective research participant receives clear, unambiguous information about the study so that a voluntary and educated decision can be made on whether to participate. Several studies have shown, however, that the documentation of informed consent, operationalized by a signature on a consent form, may not necessarily indicate informed consent.

There have been few published empirical studies of better consent. Most that do exist show null or only small effects. Knepp (2018) demonstrates that few US participants pay attention to consent forms. A series of changes had only very small effects on attention in an experiment by Ripley et al. (2018), and this holds across the literature (Crouse Quinn et al., 2012) – though longer consent forms can help (Nishimura, 2013). Unfortunately, the things shown to increase attention rely on person-to-person interaction (Cohn & Larson, 2007; Flory & Emanuel, 2004; Sudore et al., 2006), which has been made more difficult by the pandemic. Although some of these studies focus on minority populations, all focus on participants from the United States, and all use attention as their primary outcome variable. Our study aimed to determine whether a better consent design is effective in an online setting and whether it affects other outcomes beyond just attention, such as felt dignity.

The inclusion of qualitative questions in the research design

In our qualitative study, participants stated their preference for more open-ended questions. According to them, these questions are more respectful because they feel free to air their views. Strange et al. (2001) argue that including qualitative sections within experimental research helps bring them closer to participants' preferences, and helps us 'listen' better. However, we are unaware of any work measuring the effect of open-ended questions on felt dignity, hence our decision to test this empirically. We are also aware that the inclusion of open questions involves some trade-offs. Reja et al. (2003) suggest that open-ended online survey questions can increase dropout rates. Smyth et al. (2009) find that open-ended questions are somewhat sensitive to design features and can be hard to respond to if the questions are not well-framed.

Feedback about the outcomes of the research process

Feedback to participants is frequently suggested as a way of making research practice more ethical (Asiedu et al., 2021). It is a central feature of many qualitative and ethnographic

approaches, which engages in member checking to build relationships and validate findings (Birt et al., 2016), and a common practice in medical research (Bredenoord & van Delden, 2012). In our study, we focussed on the content of feedback and the effect on felt ethical dignity. We built on previous qualitative work where Busara's participants requested more feedback after studies.

Study overview

To test the impacts of different modifications to the structure of the research process on participant experience, we conducted an online study that bundled together three separate experimental modifications: a consent modification, a qualitative questions modification, and a feedback modification. We assessed the impact of all experimental modifications on a variety of measures of participant experience, including their feelings, their perceptions of how they were treated, and their perceptions of our organization. The study also included a rank-order preferences survey where we asked participants to prioritize ideas to improve ethical practice.

The consent modification, qualitative questions modification, and preferences surveys were tested in a separate study phase from the feedback modification. Implementing the feedback modification required assessing the results of the other parts of the study so they could be synthesized into feedback for the participants. The feedback phase, therefore, occurred about three months after the initial phase of the study.

Participants

We recruited 2079[1] participants from the existing Busara participant database in low-income areas of Nairobi. The age of the participants ranged between 18 and 67 years. The majority of the participants were generally well-educated, with 60% having at least a college education. The participants were more likely to be men with 65% of the sample being male. The feedback phase of the study had a sample of 740 participants. The high dropout rate for this phase was likely due to the three-month gap between this phase and the other parts of the study.

Design and procedure

The participants were required to complete an online survey developed using Qualtrics. The consent and qualitative experiments were combined and presented to participants individually. Firstly, the participants were randomly assigned to one of the consent experimental groups – the control group, in-depth video consent, and extended consent with an FAQ. They were randomly assigned again to determine whether they would receive the open-ended question.

1 The initial sample size was larger due to the link circulating online. Some of the participants we had recruited shared the link online and with their friends resulting in the need to gain consent from the people we had not recruited. Due to this, the final sample was 2079 participants

Three months after this initial phase, we contacted the participants to share the study results for the feedback experiment. Participants were randomly assigned to one of four experimental groups to determine whether and how they received this feedback (control group, no results shared; results only; results + recommendations for action; and results + Busara's commitment to action).

We assessed the impact of all experimental modifications using an array of measures of participant experience. These measures are detailed in Table 1.

CONSTRUCT	SCALE	ITEMS	SAMPLE	MINIMUM	MAXIMUM
Respectfulness	Respect	1-5	Busara has treated me with dignity	1 - Strongly Disagree	1 - Strongly Agree
Fairness	Fairness	1-5	I feel that Busara will keep its promises.	1 - Strongly Disagree	1 - Strongly Agree
Negative Affect	PANAS	1-5	Read every word carefully and select the number that indicates how you feel at this time - Afraid	1 - Not at all	5 - Extremely
Positive Affect	PANAS	1-5	Read every word carefully and select the number that indicates how you feel at this time - Active	1 - Not at all	5 - Extremely
Willingness to recommend Busara	NPS	1-5	How likely are you to recommend your friend to participate in research with Busara	1 - Completely unlikely	1 - Completely likely

TABLE 1.

Experimental modification 1: Consent

Research question

By implementing this experimental modification, we sought to investigate what consent design has the most positive effect on participant experiences. Specifically, we wanted to determine the effect of improved consent on felt ethical practice. We improved how we presented our consent forms, varying the mode of presentation and simplicity of the materials included in the consent form to improve felt ethical practice.

Experimental groups

The participants were randomly assigned to one of three experimental groups. Participants

in each of the three experimental groups were presented with a different consent design as follows:

- Control group: received a standard (but plain language) consent screen.
- FAQ: received a standard consent + option to view an FAQ with more information.
- In-depth video consent: received a video of standard consent + a commitment to reading and responding to participants' feedback from a lab officer.

Analysis

We used planned contrasts to estimate the effects of the improved consent design on participant experiences. To test the impact of different types of consent, we used planned contrasts to compare our two improved consent forms to a standard consent, then compared the two improved consent forms to each other.

Results

The two improved consent forms did not produce a detectable impact on perceived respectfulness, $b = -0.012$, $t(2073) = -0.33$, $p = 0.73$, perceived fairness, $b = 0.0009$, $t(2075) = 0.03$, $p = 0.976$, negative feelings, $b = 0.02$, $t(2075) = 0.48$, $p = 0.62$, positive feelings, $b = -0.01$, $t(2075) = -0.37$, $p = 0.73$, or willingness to recommend Busara, $b = -0.004$, $t(2075) = -0.16$, $p = 0.87$; see Figure 1. As shown in Table 2, the mean values for all variables were nearly identical across the consent form versions. Due to the large sample size achieved in this experiment, we can be fairly confident this result is not a false positive.

Experimental modification 2: Open-ended question

Research question

By implementing this experimental modification, we sought to investigate whether including a qualitative response option positively affects participant experiences.

Experimental groups

The participants were randomly assigned to one of two experimental groups. The treatment group saw an open-ended question during the survey, during which they were invited to provide feedback or share opinions on how we can improve research processes, while the control group did not receive this open-ended option. Participants then filled out a respectfulness and fairness scale to assess their felt ethical practice.

Analysis

We used a contrast to estimate the effects of including a qualitative response option of felt ethical practice. We compare the inclusion of the qualitative option versus not including a qualitative option.

Due to the potential compounding effects of the bundled treatments, we also tested for interactions between treatments.

Results

The inclusion of an open-ended question did not produce a detectable impact on perceived respectfulness, $b = -0.04$, $t(2075) = -0.13$, $p = 0.9$, perceived fairness, $b = 0.03$, $t(2075) = 0.97$, $p = 0.17$, negative feelings, $b = -0.05$, $t(2075) = -1.52$, $p = 0.13$, positive feelings, $b = -0.04$, $t(2075) = -1.37$, $p = 0.17$, or willingness to recommend Busara, $b = -0.009$, $t(2075) = -0.39$, $p = 0.69$; see Figure 1.

Potential compounding effect

We considered the possibility that the impact of including qualitative questions in the study design compounds with the impact of improving the consent procedure, resulting in an improved participant experience only when these two study enhancements are included together. To assess this possibility, we tested the interaction between the primary contrast that represented the comparison between the two improved consent forms and the primary consent form and the contrast representing the inclusion of a qualitative question. We found no significant interactions on any variable (respectfulness, $b = 0.03$, $t(2073) = -0.44$, $p = 0.66$; perceived fairness, $b = -0.01$, $t(2073) = -0.19$, $p = 0.85$; negative feelings, $b = -0.10$, $t(2073) = -1.25$, $p = 0.21$; positive feelings, $b = 0.12$, $t(2073) = -1.99$, $p = 0.047$; or willingness to recommend Busara, $b = -0.005$, $t(2073) = -0.11$, $p = 0.91$; note that though the p-value for positive feelings is below .05, this result is almost certainly a false positive as the pattern of means is not interpretable).

		NO QUALITATIVE			COMPLETED QUALITATIVE		
Variable	**Consent condition**	N	MEAN	SD	N	MEAN	SD
Respect	Standard	325	4.61	0.78	372	4.58	0.73
	More Info	331	4.58	0.73	343	4.63	0.68
	Video Consent	336	4.58	0.82	372	4.55	0.80
Fairness	Standard	325	4.68	0.71	372	4.66	0.65
	More Info	331	4.68	0.64	343	4.67	0.62
	Video Consent	336	4.69	0.70	372	4.63	0.75
Negative feelings	Standard	325	1.58	0.91	372	1.59	0.87
	More Info	331	1.63	0.92	343	1.56	0.85
	Video Consent	336	1.68	0.97	372	1.57	0.81
Positive feelings	Standard	325	4.60	0.56	372	4.48	0.65
	More Info	331	4.54	0.61	343	4.52	0.66
	Video Consent	336	4.51	0.65	372	4.53	0.60
Recommend Busara	Standard	325	4.76	0.53	372	4.75	0.49
	More Info	331	4.77	0.47	343	4.75	0.46
	Video Consent	336	4.74	0.57	372	4.74	0.56

TABLE 2. Means and standard deviations for all variables across each experimental condition

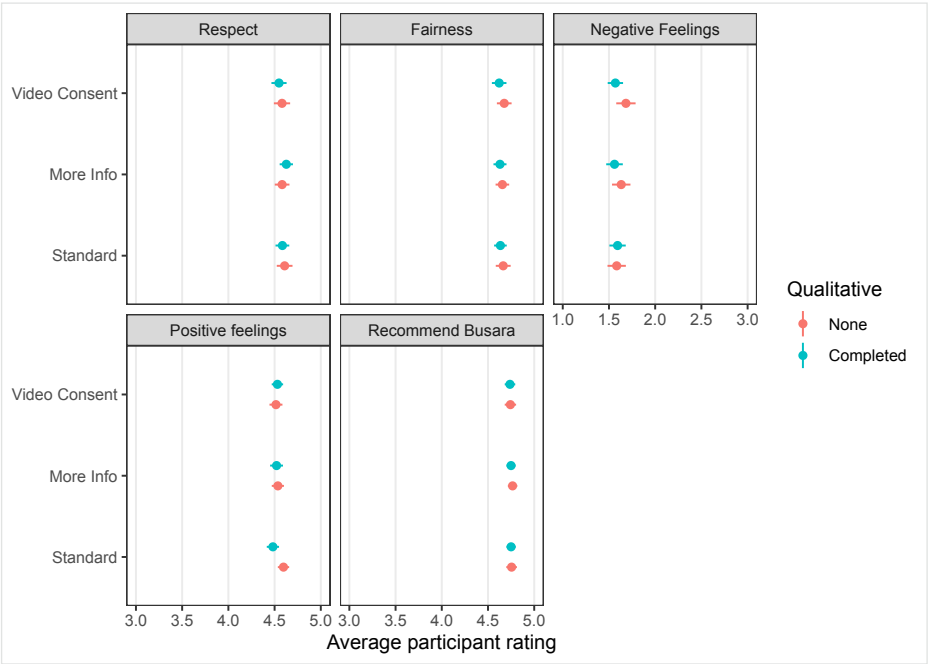

FIGURE 1. Means and 95% confidence intervals for the different consent and qualitative question conditions for all major study variables.

Experimental modification 3: Feedback

Research question

By implementing this experimental modification, we sought to investigate what feedback content positively impacts perceived ethical practice and respectfulness.

Experimental groups

The participants were randomly assigned to one of four experimental groups. Participants in each of the four experimental groups were given feedback as follows:

Results only: Participants receive a follow-up notification describing the results of the study after the analysis of results is completed.

- Results + recommendations: Participants receive a follow-up notification describing the results of the study and our recommendations from this study, including a recommendation to make no changes to our research protocols if that is our decision, but with no commitments to make those changes (in contrast to T3), after the analysis of results is completed.
- Results + Busara action: Same as results + recommendations, but participants additionally receive an assessment from us on what research process Busara is likely to change or why they will not change if that is our conclusion after the analysis of results is completed.
- Control: No feedback.

After receiving the results of the study, participants then filled out the different scales that comprised our study variables. Even though the control group did not receive any feedback, they also filled out the different scales.

Results

We found that receiving feedback about the result of the study did not produce a detectable impact on perceived respectfulness, $b = -0.02$, $t(736) = 0.35$, $p = 0.72$, perceived fairness, $b = -0.01$, $t(736) = -0.22$, $p = 0.82$, negative feelings, $b = -0.08$, $t(736) = -1.68$, $p = 0.09$, positive feelings, $b = 0.05$, $t(736) = 1.29$, $p = 0.2$, or willingness to recommend Busara, $b = -0.08$, $t(736) = -0.15$, $p = 0.88$ (see Table 3).

	CONTROL			RESULTS			RESULTS + RECOMMENDATION			RESULTS + BUSARA ACTION		
Variable	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD
Respect	190	4.61	0.89	178	4.66	0.82	189	4.64	0.81	183	4.60	0.91
Fairness	190	4.67	0.79	178	4.69	0.73	189	4.60	0.87	183	4.66	0.80
Negative feelings	190	1.69	0.87	178	1.57	0.77	189	1.52	0.86	183	1.62	0.81
Positive feelings	190	4.50	0.58	178	4.62	0.49	189	4.56	0.56	183	4.51	0.65
Recommend Busara	190	4.76	0.51	178	4.77	0.44	189	4.73	0.58	183	4.77	0.50

TABLE 3. Mean values for all variables across each experimental condition.

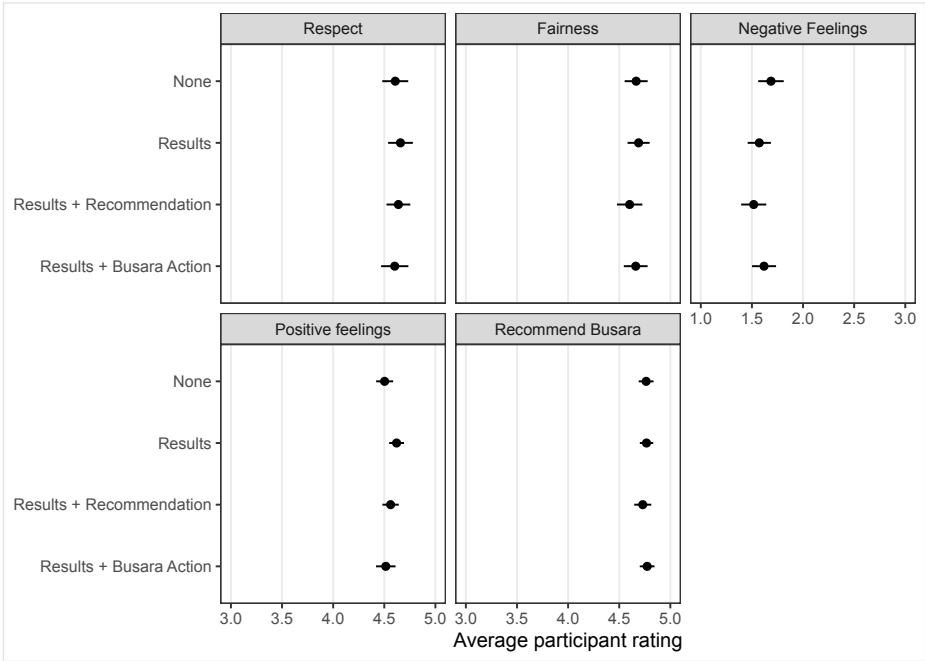

FIGURE 2. Means and 95% confidence intervals for the different feedback conditions for all major study variables.

Preferences survey

In addition to the experiments, we asked participants to prioritize suggested ideas to improve ethical research practice through a rank-order survey. We wanted to know which ideas participants prefer that researchers prioritize. According to the results, closing the feedback loops is one of the most important things participants want researchers to do more about. Participants also ranked clear consent procedures, including qualitative questions, anonymity, shorter surveys, codesigned studies, and more African PIs as important issues

Priority	Count	Proportion of respondents with this answer	Proportion of respondents with this answer
Share results	1569	75.47%	10.01%
Clear consent	1395	67.10%	8.90%
Qualitative questions	1374	66.09%	8.77%
Anonymity	1103	53.05%	7.04%
Shorter surveys	1064	51.18%	6.79%
Codesign studies	1059	50.94%	6.76%
More African PIs	1052	50.60%	6.71%

TABLE 4.

to improve ethical practice.

Discussion

The findings in this study show conclusive null results across all our outcomes for all interventions. All the interventions we tested had no detectable impact on participant experience, including their felt emotions, perceived ethical practice, and perceptions of our organization. These results stand in stark contrast to the results of the preferences survey, where results sharing, clear consenting procedures, and the inclusion of qualitative questions emerged as priority items for improving ethical practice in research among our participants.

We have some hypotheses as to why this pattern may have emerged. As discussed earlier, this study was built on a qualitative study where we investigated what participants care about regarding ethical practice. The null results could simply mean a potential misalignment in how we interpreted what participants meant in this qualitative study. For example, we interpreted mentions of the importance of open-ended questions as meaning that it was important to balance quantitative measures with qualitative ones. However, the respondents could have wanted studies that are purely qualitative, or they may have wanted to answer questions via a dialogue with a person rather than through an online textbox. Regarding the consenting experimental modification, the null results could mean that consenting should be a continuous process, not a single event in time. In our study, the consent was a one-off exercise, which is one of the study's weaknesses.

The null results in the feedback experimental modification could mean that alternative methods of sharing results are needed beyond what we tested in this study. In our case, we were limited to sharing our results via SMS, which may not have been an especially impactful way of receiving results. The null result could also mean the duration between the study and when the results are available to participants matters. In this study, the waiting period was about three months, which may have further diminished the impact of results-sharing.

It could be that no modification was going to make much of a difference because the experimental setting was low-stakes, especially compared to the settings that have sparked debates about ethics in development research. In cash transfer RCTs, for example, participants in the experimental group stand to gain as much as two months of income. No such stakes were present in our experiment. Overall, on the basis of this experiment, we believe that small adjustments to study procedures are unlikely to make much of a difference, and that more radical changes are needed to make a measurable impact on participant experience.

Though our experimental modifications did not produce an impact on participant experience, this does not mean that improvements to ethical practice are not achievable.

Nor does it mean that these improvements could not be achieved through other ways of providing feedback, including open-ended ways for research participants to express their opinions, and improving consenting procedures. The preferences survey we conducted suggests that our participants care about improvements in all of these aspects of research. We therefore flag two areas for future research that could uncover better ways of improving ethical research practice:

1. Engaging more directly with participants to better understand their preferences for ethical practice.
2. Testing larger, more substantial changes in research protocols, perhaps especially in research settings with high stakes for participants.

Upholding a higher standard of ethical practice in research matters. There is a need to build empirical evidence on participants' preferences regarding ethical practice and how to measure and monitor improved ethical practice. We hope this will bring about a more just research ecosystem, better evidence, and wiser policies grounded in this evidence – for everyone's benefit.

References

Amarante, V., Burger, R., Chelwa, G., Cockburn, J., Kassouf, A., McKay, A., & Zurbrigg, J. 2021. Underrepresentation of developing country researchers in development research. *Applied Economics Letters* 29(17), 1659-64. https://doi.org/10.1080/13504851.2021.1965528

Asiedu, E., Karlan, D., Lambon-Quayefio, M.P., & Udry, C.R. 2021. *A call for structured ethics appendices in social science papers*. National Bureau of Economic Research, Working paper 28393. https://doi.org/10.3386/w28393

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. 2016. Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research* 26(13), 1802-11. https://doi.org/10.1177/1049732316654870

Bredenoord, A. L., & van Delden, J.J. 2012. Research ethics in genomics research: feedback of individual genetic data to research participants. In J. Schildmann, V. Sandow, O. Rauprich, & J. Vollmann (eds), *Human medical research: ethical, legal and socio-cultural aspects*. Springer, pp. 127-36.

Cohn, E., & Larson, E. 2007. Improving participant comprehension in the informed consent process. *Journal of Nursing Scholarship* 39(3), 273-80. https://doi.org/10.1111/j.1547-5069.2007.00180.x

Cornwall, A., & Jewkes, R. 1995. What is participatory research? *Social Science & Medicine* 41(12), 1667-76. https://doi.org/10.1016/0277-9536(95)00127-s

Cronin-Furman, K., & Lake, M. 2018. Ethics abroad: fieldwork in fragile and violent contexts. *PS: Political Science & Politics* 51(3), 607-14. https://doi.org/10.1017/S1049096518000379

Crow, G., Wiles, R., Heath, S., & Charles, V. 2007. Research ethics and data quality: the implications of informed consent. *International Journal of Social Research Methodology* 9(2), 83-95. https://doi.org/10.1080/13645570600595231

Deaton, A., & Cartwright, N. 2018. Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine* 210, 2-21. https://doi.org/10.1016/j.socscimed.2017.12.005

Duflo, E., & Banerjee, A. 2013. *Poor economics: rethinking poverty and the ways to end it*. Random House.

Falk, A., & Heckman, J.J. 2009. Lab experiments are a major source of knowledge in the social sciences. *Science* 326(5952), 535-8. https://doi.org/10.1126/science.1168244

Flory, J., & Emanuel, E. 2004. Interventions to improve research participants' understanding in informed consent for research: a systematic review. *JAMA* 292(13), 1593-1601. https://doi.org/10.1001/jama.292.13.1593

Glennerster, R., & Powers, S. 2016. Balancing risk and benefit: ethical tradeoffs in running randomized evaluations. In G.F. DeMartino & D.N. McCloskey (eds), *The Oxford Handbook of Professional Economic Ethics*. Oxford University Press, pp. 367-401.

Groves Williams, L. 2016. Ethics in international development evaluation and research: what is the problem, why does it matter and what can we do about it? *Journal of Development Effectiveness* 8(4), 535-52. https://doi.org/10.1080/19439342.2016.1244700

Hoffmann, N. 2020. Involuntary experiments in former colonies: the case for a moratorium. *World Development* 127, 104805. https://doi.org/10.1016/j.worlddev.2019.104805

Humphreys, M. 2020. How to make field experiments more ethical. *Washington Post*, December 30. https://www.washingtonpost.com/news/monkey-cage/wp/2014/11/02/how-to-make-field-experiments-more-ethical/

Kaplan, L., Kuhnt, J., & Steinert, J.I. 2020. Do no harm? Field research in the Global South: ethical challenges faced by research staff. *World Development* 127, 104810. https://doi.org/10.1016/j.worlddev.2019.104810

Knepp, M.M. 2018. Using questions to improve informed consent form reading behavior in students. *Ethics & Behavior* 28(7), 560-77. https://doi.org/10.1080/10508422.2017.1320665

Nishimura, A., Carey, J., Erwin, P.J., Tilburt, J.C., Murad, M.H., & McCormick, J.B. 2013. Improving understanding in the research informed consent process: a systematic review of 54 interventions tested in randomized control trials. *BMC Medical Ethics* 14, 28. https://doi.org/10.1186/1472-6939-14-28

Nyaga, R., Wanjiku, A., Wambua, R., Juma, C., Wambua, J., & Wein, T., January, 2022. Prioritizing dignity in practice. *The Busara Blog*. Retrieved from https://medium.com/busara-center-blog/prioritizing-dignity-in-practice-74d86b65c867

Quinn, S.C., Garza, M. A., Butler, J., Fryer, C.S., Casper, E.T., Thomas, S.B., ... & Kim, K.H. 2012. Improving informed consent with minority participants: results from researcher and community surveys. *Journal of Empirical Research on Human Research Ethics* 7(5), 44-55. https://doi.org/10.1525/jer.2012.7.5.44

Pritchett, L. 2018. *The debate about RCTs in development is over: we won. They lost.* Presentation at the Development Research Institute, New York University. https://lantpritchett.org/wp-content/uploads/2019/01/The-Debate-about-RCTs-in-Development-is-Over_cambridge.pdf

Reja, U., Manfreda, K.L., Hlebec, V., & Vehovar, V. 2003. Open-ended vs. close-ended questions in web questionnaires. *Developments in Applied Statistics* 19(1), 159-77.

Ripley, K.R., Hance, M.A., Kerr, S.A., Brewer, L.E., & Conlon, K.E. 2018. Uninformed

consent? The effect of participant characteristics and delivery format on informed consent. *Ethics & Behavior* 28(7), 517-43. https://doi.org/10.1080/10508422.2018.1456926

Smyth, J.D., Dillman, D.A., Christian, L.M., & McBride, M. 2009. Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? *Public Opinion Quarterly* 73(2), 325-37. https://doi.org/10.1093/poq/nfp029

Steinert, J.I., Nyarige, D.A., Jacobi, M., Kuhnt, J., & Kaplan, L. 2021. A systematic review on ethical challenges of 'field' research in low-income and middle-income countries: respect, justice and beneficence for research staff? *BMJ Global Health* 6(7), e005380. https://doi.org/10.1136/bmjgh-2021-005380

Strange, V., Forrest, S., & Oakley, A. 2001. A listening trial: 'qualitative' methods within experimental research. In S. Oliver, & G. Peersman (eds.), *Using research for effective health promotion*. Open University Press, pp. 138-53.

Sudore, R.L., Landefeld, C.S., Williams, B.A., Barnes, D.E., Lindquist, K., & Schillinger, D. 2006. Use of a modified informed consent process among vulnerable patients: a descriptive study. *Journal of General Internal Medicine* 21(8), 867-73. https://doi.org/10.1111/j.1525-1497.2006.00535.x

Wambua, J.M., Owsley, N.C., & Wein, T. 2023. *Participant voice first: the Busara research agenda on ethical research in the Global South*. Busara. https://www.busara.global/our-works/participant-voice-first-the-busara-research-agenda-on-ethical-research-in-the-global-south-2/

Wein, T. 2020. *Dignity and development: a review of the literature on its application, definition and measurement*. Working Paper. https://dignityproject.net/wp-content/uploads/2020/11/Wein-Dignity-Literature-Review-291120.pdf

List of contributors

Advantages and challenges of Global North/Global South research collaborations an emphasis on sub-Saharan African research infrastructure

Dana M. Basnight-Brown is a research scientist and associate professor at the United States International University – Africa (USIU-Africa), whose primary research focuses on the cognitive processes surrounding human memory, emotion, and language, particularly within the domain of multilingualism. She has a strong interest in team science initiatives and in international research development.

'Should you study abroad?' The mechanisms and utility of educational emigration from Southeast Asia

Amirun Haqqim bin Eldeen Husaini (Haqqim) is an interdisciplinary researcher pursuing his PhD in Analytics at the University of Nottingham, Malaysia, focusing on social media algorithms and polarization. He also works at COMTESSA, a research group attached to the University of the Bundeswehr, Munich, focused on analytics, computer science and mathematics. His research interests primarily include interdisciplinary research concerning technology and human flourishing, although he also has work and interests in the environmental sciences.

Miguel Silan (Université Lumière Lyon 2) is the co-founder and chief behavioral strategist of the Annecy Behavioral Science Lab, a research start-up that does multi-country research on loneliness, social connection and human flourishing especially for vulnerable populations.

Building local networks for open science: a case study of the Chinese Open Science Network (COSN)

Zhiqi Xu is a Development Studies researcher at the International Institute of Social Studies (Erasmus University Rotterdam) who explores policy-making and local engagement in rural community-driven development (CDD) in the Global South, with a focus on rural China. Her interdisciplinary work weaves culture and psychology into the study of these processes.

Yue Wang is an emerging developmental psychology researcher at the Department of Educational Psychology at The Chinese University of Hong Kong dedicated to open science practices, and committed to lifelong learning.

Liangyou Zhang is a PhD candidate at Utrecht University's Helmholtz Institute. His interest lies in investigating the interaction between attention and numerosity using fMRI and the Population Receptive Field model.

Wenqianglong Li is a molecular genetics researcher at the Nuffield Department of Clinical Neuroscience (University of Oxford) with a background in developmental neuroscience,

genetics, and psychiatry.

Chenghao Zhou is currently a master's student and will continue into the PhD program at the Perceptual and Mnemonic Decision-making Lab at the Department of Psychology (New York University), focusing on decision-making, change of mind, and confidence.

Xi Chen is a human factor engineer in OPPO. His research interests include psychometrics accuracy, statistical methods decision, semantic measurement, ML/DL application in psychology and open/meta science.

Hu Chuan-Peng is a professor at the School of Psychology, Nanjing Normal University. His research interests include meta-science, Bayesian hierarchical modeling, and its applications in self cognition and mental health. He values equality, diversity, and inclusivity and actively promotes open scholarship both in the Chinese-speaking community and the international community.

Bridging the majority and minority worlds: liminal researchers as catalysts for inclusive open and big team science

Alma Jeftić is a postdoc at the University of Copenhagen and a Research Fellow at the Peace Research Institute, International Christian University in Tokyo. She is a social psychologist and a co-founder of ABRIR, a consortium of psychological researchers from the Global South.

Flavio Azevedo is a Brazilian assistant professor of Interdisciplinary Social Sciences at Utrecht University. His research primarily focuses on the political psychology of ideological attitudes and their psychological underpinnings and integrating open science into higher education. He was a research associate at the University of Cambridge Social Decision-Making Lab and a Fulbright fellow at New York University, and obtained his PhD in Political Science at the Center for Comparative Politics of the University of Cologne

Nadia Saraí Corral-Frías is an associate professor at the University of Sonora in Mexico, Assistant Director in the Psychological Science Accelerator (PSA), founding member of ABRIR, and co-chair of the MDD workgroup of the Latin American Genomics Consortium (LAGC). She obtained her Ph.D. in Neuroscience at the University of Arizona. She is a fervent proponent of increasing diversity in psychological, genetics and neuroscience research and prides herself in her outreach activities to increase knowledge about the mind as well as the brain and its function.

Marc Yancy Lucas is a lecturer in psychology, anthropology, and social sciences at the University of Sonora. He holds a doctorate in psychology (with a focus on environmental psychology) from the University of Sonora and specializes in applied research in the area of environmental psychology and urban anthropology. He has extensive experience in urban planning and the reciprocal effects of environment and human behavior and affect. His teaching and research emphasize the importance of justice in physical and social environments.

Contextually grounded research in postgraduate research training in Africa: why and how

Aurelia Muthoni Munene is a social policy researcher, a gender equality and social inclusion specialist, and a sexual reproductive health professional. She is the founder of Eider Africa and the executive director of the Roots and Wings Research and Development Organization. She is deeply committed to nurturing a generation of competent, critical, reflexive, collaborative and ethical African researchers and co-creating research ecosystems in the continent that are responsible and meaningful.

How inclusive and equitable is research in clinical psychology that focuses on the Global South?

Helen Niemeyer is postdoctoral research fellow in clinical psychology at the Freie Universität Berlin and interested in sociology, meta-research and open science. She completed her doctorate on the topic of publication bias, collaborated in interdisciplinary projects and was visiting researcher in sociological departments (Cultural Sociology and Center for the Study of Rationality, Hebrew University Jerusalem, Israel; Sociology of Science, University of Chicago, USA), and is active in different open science initiatives.

Louis Schiekiera is a trained psychologist and sociologist at the Division of Clinical Psychological Intervention (Freie Universität Berlin). He has a diverse range of research interests that span machine learning, psychotherapy, meta-research and historiography.

Collaborating to support a more inclusive evidence ecosystem

Maya Ranganath is the associate director of Global Networks and Inclusion at the Center for Effective Global Action.

Behavioral public policy for global challenges

Sanchayan Banerjee is assistant professor of Behavioural and Environmental Economics at Vrije Universiteit Amsterdam and a visiting fellow of the London School of Economics. His research focuses on developing and testing citizen-oriented participatory behavioral public policies, notably nudge+.

Matteo M. Galizzi is associate professor of Behavioural Science in the Department of Psychological and Behavioural Science at the London School of Economics and Political Science (LSE). He is director of the LSE Executive MSc in Behavioural Science and co-director of the LSE Behavioural Lab.

Building research capacity in Africa via big team science: challenges and lessons learned from the ManyLabs Africa initiative

Adeyemi Adetula (Alex Ekwueme Federal University Ndufu Alike) is a Nigerian experimental,

forensic, and social psychologist interested in social behaviors, big team science replication, and meta-research. He is a PhD candidate at the Université Grenoble Alpes working on the generalizability of psychology from an African perspective to improve Africans' participation in global research.

The gender data gap in development policy research

Tommie Yeo Thompson is a quantitative researcher at the Charter Cities Institute, an organization working to improve urbanization in sub-Saharan Africa. Before joining CCI, he implemented large-scale impact evaluations of USAID and MCC projects in Africa, Latin America, and Southeast Asia. Tommie also worked as a Research Specialist at Busara, where he led behavioral experiments to address poverty in Kenya and Nigeria.

Winnie Mughogho is pursuing a PhD in Economics at the Queen Mary University of London. Her research interests are in development and labor economics. She is particularly keen on using rigorous research methodologies to generate evidence that informs policy. She previously worked as a Research Specialist at Busara in Kenya, where she led the Methods research agenda.

Anisha Singh is currently a PhD student at the London School of Economics and Political Science (LSE) applying behavioral insights to managerial decision-making within workplaces. Prior to joining the PhD program, she was a Vice President at Busara in Kenya where she led a portfolio of academic collaborations and experiments that focused on creating new knowledge around behavioral mechanisms and interventions in the Global South.

Assessing deliberative polling methodology for decolonizing and indigenizing research

Dennis Chirawurah is a senior practice-based community development scholar, actively engaged in teaching and research at the University for Development Studies in Tamale, Ghana, for over twenty years. He has worked in diverse capacities within international and national research networks and is teaching both undergraduate and graduate courses in the fields of development, health systems, environment and resilience.

Niagia F. Santuah (Millar Institute for Transdisciplinary and Development Studies/Millar Open University, Bolgatanga, Ghana) is a multi-skilled development communication scientist who leads learning and experimentation to broaden the evidence base for action-oriented decision making. With over 25 years of experience in research into indigenous knowledge and local food systems across West Africa, he has expertise in design, monitoring, evaluation and learning (DME&L).

Evaluating interventions: a practical primer for specifying the smallest effect size of interest

Hannah K. Peetz is a PhD student at the Behavioural Science Institute at the Radboud University. Her current research focuses on attitudes of children towards their classmates

but she is also very interested in open science practices, meta-science and methodology.

Maximilian A. Primbs is a PhD student in social psychology at the Behavioural Science Institute, Radboud University. He investigates how our social and cultural environment shapes our implicit biases. He is also the Assistant Director for Translation and Cultural Diversity at the Psychological Science Accelerator.

Leonie A. Dudda is a PhD candidate in meta science with a focus on open science, reproducibility, and methodology at the Department of Psychology, Education, and Child Studies, Erasmus University Rotterdam.

Pia K. Andresen is a quantitative methodologist currently pursuing a PhD at Utrecht University in the Netherlands. Her expertise lies in time series modeling, the measurement of behavioral dynamics, and statistical consultancy for social scientists.

Charlotte R. Pennington is an experimental social psychologist interested in how people's complex social environments impact upon health and wellbeing. She is also an expert in open science. She is Aston's Local Network Lead for the UK Reproducibility Network (UKRN), Action Editor for Registered Reports at Peer Community In, and author of A Student's Guide to Open Science: How the Replication Crisis can Reform Psychology, published by Open University Press.

Samuel J. Westwood (King's College London) is a cognitive neuroscientist investigating non-pharmacological treatment for neurodevelopment conditions as well as researching the quality of research in psychiatry.

Erin M. Buchanan teaches a variety of statistics and language courses at the Harrisburg University of Science and Technology. Her research involves computational linguistics, natural language processing, statistical thinking, and best practices in scientific research.

Utility of meta-research for Global South policymaking: a reflection on education research

Anushka Ghosh is an educator and researcher with a keen interest in evidence-driven, impactful programs in early childhood learning for communities on the margin. At Busara, she uses her knowledge of pedagogy and applied cognitive neuroscience in learning and teaching to build meaningful research-practice partnerships where programs and policies are evidence-driven, and social scientists are guided in their pursuit of knowledge through contextual expertise from the field. She is passionate about building safe and joyful learning spaces for children.

The democratizing effects of doubting

Adam Moe Fejerskov is a senior researcher at the Danish Institute for International Studies, where he works on the terrains and ramifications of contemporary global inequalities. His most recent book is *The Global Lab* (Oxford UP).

‘Here, my degree does not matter; you are the teacher’: ethnography, citizen ethnography, and researching research on global development

Ben Eyre is an ethnographer focused on global development, climate change, and sustainable business. His research spans from the boardrooms of investors in North America and Europe to the fields of smallholder farmers in East Africa. He is a Leverhulme Early Career Fellow based at the School of Global Development at the University of East Anglia.

Sharon Acio Enon is a member of the Citizen Ethnography Collective. She has a bachelor’s degree in Arts with Education from Kampala International University and comes from Kole District in Northern Uganda. She has taught in several schools and worked as a researcher on topics including education, humanitarianism, and youth.

Dorah Adoch is a member of the Citizen Ethnography Collective. She holds a diploma in law from the Law Development Centre in Kampala, Uganda, and is currently pursuing a bachelor’s degree in Law at Gulu University. She has worked as a primary school teacher, research assistant, and legal clerk.

Vicky Alum is a member of the Citizen Ethnography Collective. She is a 34-year-old single mother of an 11-year-old child. She comes from Amolatar District in Northern Uganda. She is the first born of nine children and holds a degree in Development Studies from Kampala University. She is currently pursuing a post graduate diploma in public administration and management science from Lira University.

Joel Ekaun Hannington is a member of the Citizen Ethnography Collective. He has a bachelor’s degree with honors in Public Administration from Gulu University. Since 2018 he has worked as a research assistant on a number of projects about education, local politics, humanitarianism, and entrepreneurship in Eastern and Western Uganda.

Ann Gumkit Parlaker is a member of the Citizen Ethnography Collective. She is from Kitgum District in Northern Uganda and holds a bachelor’s degree in public administration and management from Lira University. Since 2019 she has worked as a researcher on different projects for organizations including The University of East Anglia, Lira University, and the Field Lab Uganda. She also runs a small tailoring business.

Ben Jones is associate professor in the School of Global Development at the University of East Anglia. He is currently researching the experiences of educated ‘unemployed’ youth in eastern Uganda. His recent publications include ‘Education as identity: the scaffolding of “being educated” in eastern Uganda’ (published in *American Ethnologist*).

Jimmy Ezra Okello is a member of the Citizen Ethnography Collective. He has a bachelor’s degree in Management Science from Kyambogo University. He has participated in research projects in Kyegegwa and Ngora districts, and is an entrepreneur who has built his own shoe business.

Robert Oluka is a member of the Citizen Ethnography Collective and comes from a humble family background and left school after his first year of secondary education to support his parents through selling fish. After several years, he returned to education and has been involved in social research since 2019. He is studying for a Masters in Global Development at the Open University.

James Opolo is a member of the Citizen Ethnography Collective. James is originally from Lira, northern Uganda, and earned his undergraduate degree in Information and Communication Technology from Gulu University in Uganda. He is now pursuing an MSc course in Global Development at The Open University. Over the past seven years, he has been actively involved in research for various NGOs and universities including the University of East Anglia in the UK and Lira University in Uganda.

Do the randomized know that they have been randomized? A critique of the turn towards randomization in high-stakes development cooperation initiatives

Mario Schmidt is a senior research specialist at Busara and an associate researcher at the MPI for Social Anthropology in Halle, Germany. He is interested in the effects of evidence-based development aid interventions across East Africa and the epistemological and ethical foundations of the behavioral sciences with a special focus on cash transfers, masculinity, and developmental meta-research.

Manage relationships when starting and ending research with human participants

Joel Wambua is a research specialist at Busara. His work focuses on creating and implementing ethical guidelines that promote transparency, accountability, and fairness in research practices. Joel's dedication to ethical research is driven by his commitment to advancing knowledge that can improve the lives of individuals in developing regions while maintaining rigorous ethical standards.

Anisha Singh is currently a PhD student at the London School of Economics and Political Science (LSE) applying behavioral insights to managerial decision-making within workplaces. Prior to joining the PhD program, she was a Vice President at Busara in Kenya where she led a portfolio of academic collaborations and experiments that focused on creating new knowledge around behavioral mechanisms and interventions in the Global South.

Kelvin Kihindas is a researcher at the Common Goal Research Center and specializes in experimental design projects focusing on the biases and heuristics that affect actions and decision-making aimed at advancing behavioral science in the Global South.

Irene Ngina Gachungi (DIME, World Bank) is a seasoned researcher specializing in development research, with extensive experience conducting impact evaluations and lab experiments from inception to completion.

Patrick S. Forscher is the Director of the Culture, Research Ethics, and MEthods (CREME) developmental meta-research team at Busara, where he and his team are working to make behavioral science in development more robust, useful, and fair. Before this, he was the funding lead for the Psychological Science Accelerator, a non-profit that does large, multi-national, 'big team science' psychology projects.

Why applied psychologists should consider their work's value-laden context

Jason M Chin studies and teaches evidence law at the College of Law, Australian National University.

Cultural context and ecological validity in global development research

Sijmen A. Brouwers is an independent researcher from Potchefstroom, South Africa. He is a Cross-Cultural Psychologist with a PhD from Tilburg University. Based in South Africa since 2014, he is dedicated to researching and teaching culture and behavior in the developing world. His work includes research for UNICEF on disability inclusion and vaccine demand, as well as publications on contextualization in Human Resource Management Review and the Journal of Cross-Cultural Psychology. Brouwers also offers classes, workshops, coaching, and training on intercultural communication and culture care.

Floriza Gennari is a Brazilian behavioral scientist and gender specialist based in Geneva, Switzerland. Her expertise is in the nexus between public health, behavioral science, and the prevention and response to violence against women and children. Floriza has worked for over 15 years with organizations such as the World Bank, PAHO/WHO, UNICEF, NGOs, and was previously based in Mozambique, Eswatini, Botswana, Thailand, and Laos. Floriza holds a DrPH in Health Behavior from The George Washington University, and a Master's Degree in Public Health in Developing Countries from the London School of Hygiene and Tropical Medicine.

Consent, open-ended questions, and feedback loops: empirical insights into research ethics in the Global South

Joel Wambua is a research specialist at Busara. His work focuses on creating and implementing ethical guidelines that promote transparency, accountability, and fairness in research practices. Joel's dedication to ethical research is driven by his commitment to advancing knowledge that can improve the lives of individuals in developing regions while maintaining rigorous ethical standards.

This volume on developmental meta-research contains twenty contributions that turn meta-research's behaviorally-informed critical lens toward topics that have traditionally been the focus of global development practitioners. It explores intersections between development practitioners and meta-researchers and mobilizes a new community around the question of how we can improve the ways in which research is done in development.

Busara is a research and advisory organization, working with researchers and organizations to advance and apply behavioral science in pursuit of poverty alleviation. Busara pursues a future where global human development activities respond to people's lived experience; value knowledge generated in the context it is applied; and promote culturally appropriate and inclusive practices.

To accomplish this, we practice and promote behavioral science in ways that center and value the perspectives of respondents; expand the practice of research where it is applied; and build networks, processes, and tools that increase the competence of practitioners and researchers.

Busara publications present information, analysis, reflection and key recommendations on issues relating to behavioral science and global development.

Patrick S. Forscher is the Director of the Culture, Research Ethics, and MEthods (CREME) developmental meta-research team at Busara, where he and his team are working to make behavioral science in development more robust, useful, and fair. Before this, he was Funding Lead for the Psychological Science Accelerator, a non-profit that does large, multi-national, 'big team science' psychology projects.

Mario Schmidt is a senior research specialist at Busara and an associate researcher at the MPI for Social Anthropology in Halle, Germany. He is interested in the effects of evidence-based development aid interventions across East Africa and the epistemological and ethical foundations of the behavioral sciences with a special focus on cash transfers, masculinity, and developmental meta-research.

ISBN 978-9914-766-06-6

9 789914 766066