

# Diversity in Study Participants and a Critique of the “Representative Sample” in Human-Computer Interaction Research

Advait Sarkar

Microsoft Research, University of Cambridge, University College London  
United Kingdom  
advait@microsoft.com

**Abstract**— We reflect on our experiences in improving the diversity of participants in our research, focusing on geographic diversity and countering WEIRDness. Our reflections are grounded in four studies conducted over two years, with more than 100 total participant engagements across more than 100 hours of user studies. Our samples included participants from the UK and the USA, but also from the Republic of Ireland, the Netherlands, Kenya, Nigeria, Ghana, Armenia, Israel, and Japan. We reflect on some of the challenges we encountered and what we have learnt about the benefits of geographic diversity. Finally, we discuss the scientific ideal of “representativeness” and consider whether it is possible, or appropriate, in small sample studies ( $n \sim 20$ ). We propose that representativeness is antagonised by the ideal of diversity. We seek alternative ways of understanding and articulating the epistemic value of diversity in Human-Computer Interaction (HCI) research.

**Index Terms**—diversity, representation, end-user programming, epistemology, statistics, design research, philosophy of science

## I. INTRODUCTION

For a variety of reasons that have been much better articulated by Goel et al. [4], research in end-user programming (EUP) has a diversity problem. In EUP, most researchers are from WEIRD (Western, Educated, Industrialised, Rich, and Democratic) nations. Most research participants are from WEIRD nations. Developments in EUP follow developments in industrialised societies. As a result, for reasons that also have been better articulated elsewhere and need not be recapitulated, research suffers.

In our research group, which is based in a WEIRD nation and whose studies have hitherto been dominated by WEIRD participants, we have been attempting to improve the geographic diversity of our participants. Through collaboration with partner organisations we have recruited participants outside Europe and the United States to constitute a non-trivial proportion (though still, apart from one exception, not the majority) of our participant sample. We have conducted four such studies so far. An overview of these studies is given in Table I. Participant numbers in this table reflect only the final set of participants who successfully completed the study and whose data were included in analyses. The full number of participants we recruited for each study was greater (discussed in Section II).

In these studies, our non-WEIRD participants came from the following countries: Ghana, Kenya, Nigeria, Israel, Japan and Armenia (though the latter is considered geopolitically European). They constituted between 20.8% and 53.3% of the sample in these studies. We also attempted to improve the diversity of our WEIRD group beyond the UK and the USA, e.g., by recruiting participants from the Netherlands and the Republic of Ireland, but that is of secondary importance here.

To associate a participant with a country, we ask for their primary country of residence as part of a demographics questionnaire. We acknowledge that there are limitations to this; there are many good reasons that someone ought not to be held representative of that country for the mere fact of residing in it. And good reasons to critique geography as a measure of diversity at all. People move around, to cite one simple reason. To cite another, individuals are intersections of racial, ethnic, cultural, and linguistic identities (among others), and to associate countries with identities is to subscribe to a particular 19<sup>th</sup> century ideology of the nation. However, gathering more detailed information about ethnic and cultural identities, or immigration history, was well beyond the scope of our end-user programming research studies, and of questionable ethical propriety. We therefore use country of residence as a proxy for geographic diversity, which is itself a proxy for the many ways in which people of the world are diverse.

All the aforementioned studies followed similar methods which will be familiar to experienced practitioners of HCI research. The sessions were conducted remotely using Microsoft Teams, with participants given access to any experimental software running on the researcher’s computer via Teams screen sharing and remote control features. Study sessions typically lasted 45-60 minutes.

Data collection typically included think-alouds during the study, or guided reflections during interaction with the tool, pre-experiment demographics questionnaires, standard usability and cognitive load questionnaires, ad hoc questionnaires about feature preferences and confidence etc., and post-experiment semi-structured interviews. Questionnaires were administered using Microsoft Forms. Telemetric logs of user interactions with the tools were recorded. Audio and video of the screen share, think aloud, and interview data were recorded and transcribed.

TABLE I  
DESCRIPTION AND PARTICIPANT COMPOSITION OF FOUR STUDIES CONDUCTED OVER 2 YEARS. PARTICIPANT NUMBERS REFLECT ONLY THOSE INCLUDED IN THE FINAL ANALYSIS, NOT ALL THOSE WHO WERE RECRUITED.

| Dates conducted       | Study description and methods   | Total N | Non-WEIRD N (%) |
|-----------------------|---|---------|-----------------|
| July 2022             | A between-subjects, randomised controlled study of two variations of an experimental interface for improving end-user programming with large language models in spreadsheets [1]. | 24      | 5 (20.8%)       |
| August-September 2022 | A within-subjects, randomised controlled study of two variations of an experimental interface for improving end-user debugging of AI-generated code in spreadsheets [2].          | 24      | 5 (20.8%)       |
| June 2023             | A “participatory prompting” study eliciting opportunities and challenges for generative AI-assisted data analysis [3].  | 15      | 8 (53.3%)       |
| February-March 2024   | A between-subjects, randomised controlled study of two variations of an experimental interface for improving data analysis.   | 24      | 7 (29.2%)       |

## II. CHALLENGES AND BENEFITS

Our efforts to diversify our participant samples in our regular cadence of research, as exemplified by the four studies in Table I, have taught us several things that we outline here, presented in no particular order. Some of these reflections might seem trivial, or obvious, especially to those accustomed to studying non-WEIRD populations. On the other hand, as evidenced by the careful work of Goel et al. [4], those accustomed to studying non-WEIRD populations are conspicuously underrepresented in end-user programming research.

Moreover, in programming research, studies of such populations are typically conducted in the context of a project that seeks to serve the particular needs of a specific community (e.g., Corbett’s *Cree* [5], or Blackwell et al.’s *Visualising Bayesian Probability in the Kalahari* [6]). Efforts such as these are typically *highly* reflexive and critical practices. Starting with specific communities is an absolutely essential aspect of decolonising methodologies [7]. In a sense, in such studies, diversity is the epistemic centrepiece, the starting point and the destination.

In contrast, our studies exemplify a complementary and equally important aspect of addressing the diversity problem in EUP research: namely, incorporating the voices of a diverse range of participants *even when they are not a specific target population*. It is these types of research projects that are particularly susceptible to falling into WEIRDness traps, due to the “inconvenience” and poor institutional incentives for sample diversification. As such, we believe these reflections, while simple, are nonetheless valuable contributions to the community.

*a) Did you just assume my device?:* Prior to these studies, our research protocols made the assumption that participants would join our remote studies using laptop or desktop devices, as opposed to smartphones or tablets. This is because the small size and touch input make smartphones and tablets unsuitable for remote controlling a screen-shared prototype, because interface elements designed for a larger screen become too small, and hover affordances do not function properly with touch input. However, we never specified a device type in our instructions for participants, though we did mention in the pre-study communications that remote control screen sharing would take place.

Unlike our participants from Europe and the USA, we found that the majority of our participants from African countries joined our studies from mobile phones and tablets. This exposed our culturally biased assumption that participants would infer the need to join from a laptop or desktop device from the pre-study communications, and we quickly rectified it by stating and explaining our device type constraints in the pre-study materials.

This experience also raised deeper issues regarding the inclusivity of the entire approach of the project. Some of our participants dialled in from a smartphone not out of preference, but because it was the only computing device in their possession. Excluding participants for not owning or being able to access a laptop or desktop computer seemed to defeat the purpose of our efforts to improve diversity. We have explored at least two ways to address this problem. First, to design the study and any prototypes to work across a wider range of device types. Second, to design an alternative version of the study for participants who are using devices that are unsuitable for the study. For instance, instead of a fully interactive prototype, a click-through prototype with enlarged fonts could be used on a mobile device. Depending on the particularities of each study, such as the technologies being investigated, the software/hardware stack used to develop prototypes, etc., one option may be easier than the other.

*b) Hello? Can you hear me now?:* The studies revealed another set of assumptions we had been making about infrastructure, Internet connectivity and the electric grid. African participants often joined the call through Internet connections that were low-bandwidth, or unreliable. They did not always have access to an uninterrupted supply of electricity. Kenya was experiencing frequent blackouts throughout the studies we conducted in the summers of 2022 and 2023, for instance [8]. We attempted to compensate for poor bandwidth by turning video feeds off during the studies, though this likely came at a cost to the interactivity of the session and the depth of our subsequent qualitative analysis that is hard to characterise. We learnt to adapt our study protocols for sudden termination of the video call (e.g., due to an unplanned power outage) by saving progress as we went along, and introducing caching mechanisms into prototypes when possible.

Here again we noticed systemic incentives for erasure and exclusion. For instance, if a participant only completes part of a study and cannot be rescheduled to complete it, the “safest” option (i.e., considered the most rigorous by our disciplinary standards and thus most defensible at peer review) is to discard data from this participant and recruit a new one to take their place. Indeed, our recruitment partners, who gained a commission for each successfully completed study, repeatedly offered “replacement” participants in the stead of those who could not complete the study due to connectivity issues.

Some of our studies were heavily overprovisioned owing to the high rate of unsuccessful completion amongst non-WEIRD participants. For one study in 2022, we scheduled and conducted *seventeen* sessions with non-WEIRD participants, but data from *just five* could be used in our final analysis. This required us to revise the timelines for our studies, to take into account the high volume of studies that would not complete successfully and would need to be rescheduled.

c) *Lost in transcription*: Like many research groups, we rely on automatic transcription of study audio, followed by a manual pass to correct automatic transcription errors.

While all our participants were proficient in spoken English, the accuracy of our transcription software was much lower for English spoken with any accent other than American or British (acknowledging that there is a huge diversity of accented speech even within these two categories). As a result, the manual pass to correct transcription errors was much more time-consuming than we had initially anticipated and planned for. As with study session overprovisioning, we learnt to account for this slower manual transcription phase in the timelines for subsequent studies. This is a problem we hope will diminish with future improvements in more considerate and equitable transcription technologies.

We also encountered and had to develop solutions for issues in compensating participants fairly, since our normal approved methods of compensation (Amazon store vouchers or American Express vouchers) were not available in all countries in which our participants were resident.

d) *Diverse data, common threads*: Recruiting participants from a wider variety of countries has the obvious effect of increasing the diversity of user contexts and tasks we are exposed to over the course of our studies. For instance, in one study of data analysis tasks, we gathered examples of user tasks such as studying the relationship between education levels and job preparedness in Nigeria, analysis of workplace discrimination in Kenya, sharing software expertise across an organisation in Ghana, and performing date-time calculations in the Armenian calendar.

We encountered the limitations (albeit not unanticipated) of using geographic diversity as a proxy for diversity more broadly. For instance, studies of affluent and well-educated participants from non-WEIRD countries had much in common with the typical study of a WEIRD participant. The converse is also true. There is plenty of “first world” in the “third world”, and plenty of third world in the first.

We developed a renewed appreciation for the overused and much maligned practice of recruiting students or employees from the researchers’ institutions. While there is much to criticise about the potential homogeneity and representativeness of such samples, it is also true that students at large research institutions, and employees at large multinational corporations, can be increasingly diverse and international. Fifty years ago, a sample of 20 British University students might have been ruinously homogenous. Today the situation is rather different, and now such a sample might well contain students with multiple ethnic, cultural, racial, linguistic, etc. backgrounds.<sup>1</sup>

Recruitment diversity should not just focus on seeking *differences* (an epistemic bias that could be seen as a form of “othering”), but also look for similarities. We found considerable value in observing user problems that spanned WEIRD and non-WEIRD participants, because such observations provide a broader empirical basis for reasoning about the generality of insights that could quite plausibly have cultural specificities.

### III. DIVERSITY AND “REPRESENTATIVENESS”

A broader empirical basis for the generality of insights. Sounds good, in theory. But in practice, this feels like too great a leap to make from studies with such relatively small samples. The more we extended our efforts to recruit participants from more countries, and the more we critically reflected on what such diversity could plausibly achieve, it became apparent that we were encountering difficulties in articulating any core epistemic principle that supported seeking geographic diversity in a small sample. What can it possibly mean, from a knowledge-production perspective, to have a diverse and representative set of participants when most nations in your sample were represented by only one or two individuals?

#### A. *The paradox of diversity for representation*

Broadening the scope of HCI research beyond the concerns of WEIRD communities is an abiding endeavour of our research community, evidenced by the growing currents of feminist, queer, critical race-theoretic, and postcolonial HCI, and interest in many overlooked groups besides. In these research themes, the following idea is effectively axiomatic: such groups must not be orientalist into a homogenous “other”, that not all queer people are alike, nor women, nor people of colour. Individuals in these groups belong to an infinite fractal lattice of subgroups, with hierarchical and non-hierarchical relationships. They experience intersectionalities as they participate in these multiple groups and perform their various identities.

Yet, in many small sample study techniques which form the core of our research toolkit, such as interviews, controlled lab experiments, diary studies, etc., we strive for the scientific ideal of “representativeness” in our samples, and for consequent ideals in our data, such as “saturation”. Small sample studies are *by far* the most common type in HCI, and therefore

<sup>1</sup>Alas, this has far less to do with the progressive multiculturalism of 21<sup>st</sup> century Britain, and far more to do with the avaricious addiction of British Universities to international students’ tuition fees.

constitute the bulk of our disciplinary knowledge. The modal size of a user study sample reported in papers published in the proceedings of the ACM CHI conference is 12 [9]. The mean sample size for in-person qualitative studies is 14, and for quantitative studies 20. Following statistical logic that HCI research inherits from experimental psychology, the term “sample” implies a larger, homogeneous population, whose existence is refuted by the aforementioned research themes.

Using someone to represent a population is an exercise in reduction and abstraction. The participant is reduced to the smallest denominator of properties that they are assumed to share with the population. The individual is abstracted to a symbol or archetype. In the logic of representation, you look “through” them to a (constructed) group, not “at” them as a person.

This is a paradox. We cannot aim for representativeness while simultaneously acknowledging the complexity and irreducibility of the people we wish to study. It would seem that the ideology of diversity is antagonised by the ideology of representativeness, if not entirely incommensurate with it. From our perspective, the “representativeness” criterion may be yet another misapplication of scientism in HCI [10].

It is difficult to articulate, therefore, precisely *what* function is being performed by diversifying recruitment in a small sample. This is essential if we are to assess whether we have been successful in diversifying a sample, and if not, how to alter our recruitment strategies. If by recruiting non-WEIRD participants we are not making our sample more *representative*, what exactly are we doing? The answer cannot simply be that it is (or feels) fairer, or more ethical to do so, because while these motivations are important, they are not informative enough for us to make fine judgements about our diversification strategy. Is it fairer or more ethical to recruit a sample of participants from Ghana, Nigeria, and Kenya, or from India, Bangladesh, and Nepal? Is one sample more diverse than the other? These questions do not make any sense, particularly when asked in the context of much EUP research, which like ours, as mentioned before, is not centred on the needs of a specific target population.

We need alternative criteria for evaluating the epistemic strength of small samples, and the results of qualitative analyses of these samples. We need alternative notions of rigour that can provide a foundation for engaging with and studying diverse communities, and allow us to reason confidently about the kinds of knowledge produced through such engagements.

In the remainder of this paper, we will briefly discuss approaches that strike us as possible starting points for new epistemic foundations for studies of small, yet diverse samples. The aim is to draw attention to the need for such foundations, not to claim that we have found them.

### B. Beyond the representative sample

One approach is to reframe our objective from representation to simply presentation. Rather than aiming to represent a diversity of people, we can aim to simply present a diversity of people – focusing on the specific people that we do, in

fact, study. This shifts the emphasis from maximising the probability that people are seen, to increasing the possibility of being seen. Rather than asking of our sample how probable it is that we have obtained general knowledge, we instead ask what possibilities we have created for obtaining specific knowledges. Even a single anomalous observation can help problematise theories of human behaviour. It is through following such anomalies, for example, that it was discovered that the Müller-Lyer line length illusion, far from being a physiological universal, was in fact subject to cultural differences, differences in the built environment, and even whether the illusion was presented with the concomitant task of describing it verbally or through gestures [11, 12]. Accessibility research often motivates the extended value of designing for inclusion with reference to the “curb-cut effect” [13], where accessibility features designed for people with disabilities also benefit others: for instance, curb cuts for wheelchairs are also useful for people pushing strollers, delivery carts, or suitcases; subtitles for people with hearing impairments also benefit foreign language speakers or those watching a video on a device or in an environment where they cannot hear the audio.

This approach is not entirely satisfying. The “curb-cut effect” seems rather like the case made by proponents of trickle-down economics, but in reverse (a “trickle-up”, if you will). It has the effect of decentering the minority community and recasting their value in terms of benefits to the majority. To cite this as a benefit of diversifying participant pools is to admit that we are still interested in serving WEIRD populations and are extractively mining non-WEIRD peoples for unexpected ways of doing so. We must also be cautious of tokenism and the “one and done” fallacy (research shows for example that inclusion of exactly one single minority job applicant in candidate pools, which can satisfy diversity requirements, almost never leads to the hiring of that applicant [14]). Inclusion itself can be problematic: individuals included on the basis of diversity are often subject to associated burdens, and Ahmed characterises commitments to diversity as “non-performative”; they do not achieve what they claim [15]. Ferreira has critiqued the related idea of targeting the “local” in ICT4D research, finding that “1) communities are often essentialized in agency-depriving ways, 2) researchers claim substantial discretionary power in representing communities, and 3) participatory approaches are framed as inherently beneficial, obscuring compromises.” [16]. Inclusion is not itself sufficient, and we must be diligent in the ongoing work of attending to, analysing, and following up the nuances in the individual voices.

In *Research for the Very Particular* [17], Bertelsen et al. present another approach to the paradox of diversity and representation. It begins by altogether discarding the problematic notion of representativeness and heading instead in the opposite direction: the ultimate specificity of particular individuals. Research for the very particular aims to arrive at a deep understanding of an individual in their unique context. Unfortunately, in sidestepping the issue of representativeness, it also stops short of supplying an answer for diversity. To take



knowledges produced in such particular contexts and apply it to the design of tools meant for others might involve developing new methods for moving between different particulars, as suggested by Bertelsen et al., or finding ways to cumulate knowledge from multiple particular cases without losing the essence of their specificity.

Moreover, while we might reject the idea of generality, it cannot be escaped as a concept with political cachet. The beliefs of decision makers about the generality of knowledge matters. To effect change in the design of technology at a large scale, research needs to be persuasive to decision makers in organisations and governments. As Clancy explains, “*If people believe research done elsewhere isn’t applicable to their context, then that research is less likely to inform their decisions. That’s true even if the research actually is applicable, but people don’t believe it*” [11]. Any approach rooted only in particulars faces institutional challenges to generalisability and broader impact. While Bertelsen et al. make a compelling case for regarding generalisability not as the primary goal but as a secondary bonus outcome, this stance conflicts with the systemic incentives for producing a narrative of generalisability imposed by funding bodies, corporations, governments, and the wider academic community.

Hayes proposes *transferability* as an alternative to generalisability [18]. Transferability, as a core principle of action research, shifts the focus from generalisability in all contexts to the ability to transfer knowledge from one local context to another. By focusing on rich, contextual descriptions and local solutions, transferability allows researchers to acknowledge and embrace the complexity of diverse populations and still argue for the applicability of the knowledge produced without falling into the trap of representativeness. On the other hand, the collaborative, long-term nature of action research required for effective transferability may not be feasible for all research projects or researchers, conflicting as it does with the “*pressure on design to make its value evident on the basis of brief and shallow encounters*”, and “*the annual drumbeat of conference submission cycles*” [19].

Expanded to its greatest scope, what we are struggling with here is the postmodern challenge posed by diversity to the modernist value of representation. The accounts of queer, feminist, critical race, and postcolonial theory offer a genealogy (per Nietzsche) or deconstruction (per Derrida) for representativeness that problematises its status in research aiming to be un-WEIRDed. Unfortunately, it seems to be the nature of postmodern critique to create problems without solutions. In particular, here it does not readily offer a compelling alternative story for how we are to reason about the knowledge produced in studies of small diverse samples. What we require, then, can perhaps be described as a *metamodern synthesis* [20, 21] of the epistemology of such studies. We wish to retain the baby of high quality knowledge but discard the bathwater of exclusion and oppression. We need to seek, from a place of sincerity, a new “grand narrative” for small sample studies and a vision of progress that explains whether and how geographic diversity in small samples produces better knowledge.

## IV. CONCLUSION

We have reflected on our efforts to improve the geographic diversity of small samples of participants recruited for end-user programming research. Concretely, we conducted four studies between July 2022 and March 2024 where between 20.8% and 53.3% of our sample consisted of participants resident in non-WEIRD countries. The challenges we encountered (among others) included: needing to account for a wider range of device types used by participants, poor network connectivity and other infrastructural issues, and limitations of automatic transcription for a diverse range of accented speech. These challenges caused us to critically examine our own biases and assumptions about study participants, and prompted adaptations in our study protocols.

In considering how to evaluate the strength of a small but diverse sample it becomes evident that the scientific ideal of “representativeness” is inappropriate for, and even counter to, the ideology of diversity. The paradox of diversity for representation, in a nutshell, is this: the pursuit of a diverse sample is motivated by the principle that it is problematic to assume that the experiences of one group reflect those of another, but the pursuit of a representative sample is motivated by precisely that problematic assumption.

If not representativeness, what epistemic frame can we apply to understand the value of diversity in small samples? We have pulled at a few threads: shifting from representation to presentation, research for the very particular, and transferability rather than generalisability. In wrestling with the postmodern challenge posed by diversity to the modernist value of representation, we have clarified the need for a “metamodern synthesis” that can explain how geographic diversity contributes to better knowledge production in small sample studies.

Our journey toward making our research practices more global and equitable continues. The questions we raise have no simple answers. We believe that critical examination of the epistemic foundations of small sample studies in human-computer interaction research is essential, not merely as an intellectual exercise to resolve the paradox of diversity for representation, but ultimately to produce more nuanced and theoretically well-founded understandings of the diverse world in which we live and for which we design.

## ACKNOWLEDGMENTS

Thanks to Andy Gordon, Carina Negreanu, and Siân Lindley, for supporting our endeavours to diversify study participants. Thanks to Kagonya Awori and Mide Olatoye (under the aegis of Jacki O’Neill) for their assistance in recruiting participants resident in Africa. Thanks also to my collaborators and interns: Ian Drosos, Michael Xieyang Liu, Kasra Ferdowsi, and Xiaotong (Tone) Xu, for persevering through the vagaries of experimental recruitment pipelines.

## REFERENCES

- [1] M. X. Liu, A. Sarkar, C. Negreanu, B. Zorn, J. Williams, N. Toronto, and A. D. Gordon, ““What It Wants Me

- To Say”: Bridging the Abstraction Gap Between End-User Programmers and Code-Generating Large Language Models,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544548.3580817>
- [2] K. Ferdowsi, J. Williams, I. Drosos, A. D. Gordon, C. Negreanu, N. Polikarpova, A. Sarkar, and B. Zorn, “COLDECO: An End User Spreadsheet Inspection Tool for AI-Generated Code,” in *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 2023, pp. 82–91.
- [3] I. Drosos, A. Sarkar, X. Xu, C. Negreanu, S. Rintel, and L. Tankelevitch, ““It’s like a rubber duck that talks back”: Understanding Generative AI-Assisted Data Analysis Workflows through a Participatory Prompting Study,” in *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*, ser. CHIWORK '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: <https://doi.org/10.1145/3663384.3663389>
- [4] H. Goel, A. Kumar, and S. S. Ragavan, “End-User Programming is WEIRD: How, Why and What to Do About It,” in *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 2023, pp. 41–50.
- [5] J. M. R. Corbett, “Indigenizing computer programming for cultural maintenance,” in *Companion Proceedings of the 2nd International Conference on the Art, Science, and Engineering of Programming*, ser. Programming '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 243–244. [Online]. Available: <https://doi.org/10.1145/3191697.3213802>
- [6] A. F. Blackwell, N. J. Bidwell, H. L. Arnold, C. Nqeisji, K. Kunta, and M. M. Ujakpa, “Visualising Bayesian Probability in the Kalahari.” in *PPIG*, 2021.
- [7] L. T. Smith, *Decolonizing Methodologies: Research and Indigenous Peoples*. Zed Books, 1999.
- [8] A. L. Dahir, “Kenya suffers third major power blackout in four months,” Dec 2023. [Online]. Available: <https://www.nytimes.com/2023/12/11/world/africa/kenya-blackout.html>
- [9] K. Caine, “Local Standards for Sample Size at CHI,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 981–992. [Online]. Available: <https://doi.org/10.1145/2858036.2858498>
- [10] S. Reeves, “Human-computer interaction as science,” in *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives*, ser. CA '15. Aarhus N: Aarhus University Press, 2015, p. 73–84. [Online]. Available: <https://doi.org/10.7146/aahcc.v1i1.21296>
- [11] M. Clancy and C. Fry, “When Research Over There Isn’t Helpful Here,” *New Things Under the Sun*, nov 17 2023, <https://www.newthingsunderthesun.com/pub/0xbymz4>.
- [12] K. Cherry, “How The Müller-Lyer Illusion Works,” <https://www.verywellmind.com/how-the-muller-lyer-illusion-works-4111110#Does%20Everyone%20See%20The%20M%C3%BCller-Lyer%20Illusion?>, 2023, accessed 26-06-2024.
- [13] M. A. Rothberg, *Designing for Inclusion: Ensuring Accessibility for People with Disabilities*. Cham: Springer International Publishing, 2019, pp. 125–143. [Online]. Available: [https://doi.org/10.1007/978-3-319-96906-0\\_7](https://doi.org/10.1007/978-3-319-96906-0_7)
- [14] S. K. Johnson, D. R. Hekman, and E. T. Chan, “If There’s Only One Woman in Your Candidate Pool, There’s Statistically No Chance She’ll Be Hired — hbr.org,” <https://hbr.org/2016/04/if-theres-only-one-woman-in-your-candidate-pool-theres-statistically-no-chance-shell-be-hired>, 2016, accessed 26-06-2024.
- [15] S. Ahmed, *On Being Included*. New York, USA: Duke University Press, 2012. [Online]. Available: <https://doi.org/10.1515/9780822395324>
- [16] P. Ferreira, “Examining the “Local” in ICT4D: A Postcolonial Perspective on Participation,” in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, ser. CHI '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: <https://doi.org/10.1145/3613904.3642748>
- [17] O. W. Bertelsen, S. Bødker, E. Eriksson, E. Hoggan, and J. Vermeulen, “Beyond generalization: research for the very particular,” *Interactions*, vol. 26, no. 1, p. 34–38, dec 2018. [Online]. Available: <https://doi.org/10.1145/3289425>
- [18] G. R. Hayes, “Taking action in your research,” *Interactions*, vol. 19, no. 4, p. 50–53, jul 2012. [Online]. Available: <https://doi.org/10.1145/2212877.2212890>
- [19] A. Sarkar, “Should computers be easy to use? questioning the doctrine of simplicity in user interface design,” in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544549.3582741>
- [20] T. Vermeulen and R. van den Akker, “Notes on metamodernism,” *Journal of Aesthetics & Culture*, vol. 2, no. 1, p. 5677, 2010. [Online]. Available: <https://doi.org/10.3402/jac.v2i0.5677>
- [21] H. Freinacht, *The Listening Society: A Metamodern Guide to Politics, Book One*, ser. Metamodern Guides. Metamoderna ApS, 2017, vol. 1, Kindle Edition.