This post is also available on the In Development website.

In 1993, I was sitting on a mud floor with a small group of women in a village in Karnataka, trying to start a conversation about their lives. I had a survey for them to answer: heavy on questions that required numerical answers about consumption and family structure — who lived in the household, how old they were, what they had eaten, what the roof was made of. This is standard stuff in development economics.

We were a few minutes in when the door slammed open. One woman's husband stormed into the room, grabbed her by the hair and dragged her out, shouting that lunch wasn't cooked and she was wasting her time with us.

I was a young economist then, with a freshly minted PhD. My questionnaire had no item for what I had just seen. We had not come to study domestic violence; we had come to collect evidence on more prosaic questions of how sociocultural and economic systems shape marriage markets and living standards.

This is a vague enough remit that, in principle, almost anything should have qualified as relevant; it should have been easy to retool and collect more data that accurately reflected women's lives. But the survey we had brought with us — with all its inherent assumptions — was calibrated to register what could be measured cleanly and not much else. Our disciplinary training was not fit for purpose.

It took us a week of staying in that village, drinking buttermilk and coffee, sitting through many silences, before one of the women finally opened up and said: "You have become our friends and we can't lie to you anymore. We feel like we are in jail. Our husbands beat us all the time. They spend the family money on alcohol. No one helps."

That was new information. We rewrote our questionnaire on the spot, added items on wife-beating, which resulted in a mixed-methods analysis of domestic violence and one of the first economics papers on the subject in a developing country.

Vijayendra Rao (in blue shirt) with team in rural Karnataka, 1993. © Vijayendra Rao

I have thought about that week for more than 30 years. Not because the story is unusual — anyone who has done serious fieldwork has a version of it — but because of what it taught me about my own discipline. I am an economist, and economics is well known for its rigor, its emphasis on quantitative methods — and its distance from its subjects. Economics tends to focus on the measurable, which can exclude what is important; if it cannot be included in a survey module, it is not worth studying. But life is about more than survey modules, and the wall between economist and subject is not a methodological convenience. It is the central problem with the discipline.

I have spent my career as a kind of spy — an economist smuggling anthropology's methods across the wall. I have written the careful academic version of this argument. What follows is the version I would talk about over a drink.

How We Got Here

Things did not start this way. In the late 19th century, Charles Booth, a shipping magnate turned amateur sociologist, took it upon himself to find out how the poor of London actually lived. He spent 17 years on the job. His team gathered information on 4 million Londoners: school inspectors, factory owners, clergymen, policemen and many more. They did surveys, but they also did much more: they took notes, they did open-ended interviews, they made color-coded maps. Booth used both numbers and stories because his question — what does poverty look like in London? — demanded both. He and his team stitched all of it into 17 volumes of "Life and Labour of the People in London," showing, block by block, who was wealthy and who was destitute and who was somewhere in between. This integration of both story and data shaped economic and social policy in Britain for a generation.

Over the 20th century, that integration came apart. Economics, in its push for scientific status, narrowed itself to the analysis of quantitative data. Cultural anthropology split off into the ethnographic tradition — some of it deeply insightful, and some of it consumed by critique and navel-gazing. Sociology fragmented into many parts and a spiral of internal arguments over methods. Psychology went experimental. Political science was increasingly influenced by economics both in theory and method but retained an openness towards mixing qualitative and quantitative methods. By the 1980s, the disciplines distinguished themselves from one another mainly by what they refused to look at.

In economics, there were a few attempts to bridge this gap that did not have much of an impact. Back in 1984, Pranab Bardhan organized a seminal conference on "Conversations Between Economists and Anthropologists" on data and mixed methods. In 2002, I tried to make a case for "participatory econometrics."

Then came the credibility revolution. In an effort to make research more credible and reliable, causal inference became the lodestar of empirical economics. This meant that economists could say with confidence that A caused B, but it also meant neglecting the types of questions that could not be answered with these tools.

I do not want to be misread. The credibility revolution is a major leap forward for the social sciences. We can now answer questions that were genuinely beyond us 30 years ago. But when the only acceptable tool is a hammer, one tends to look for questions that could use a nail. Economists tend to prefer questions that can be answered cleanly and leave aside those that cannot. In development, the consequences have been particularly stark. We can now estimate to three decimal places how a cash transfer affects a child's height. We have almost nothing to say about how the program worked (or did not work) and we almost never hear directly from the child and their family about what they had to say about it — in their own words.

The Distance Problem

In 2002, the sociologist Michael Burawoy drew a hard line between what he called positive science and reflexive science. Positive science — which is roughly economics — insulates itself

from its subjects. Data collection is standardized so it does not matter who is collecting the data. The researcher does not have to be on site; a survey firm can administer a survey as well as s/he can. The world is to be observed, not participated in. Reflexive science — ethnography, in its best form — embraces the opposite point of view. The interview is not separate from the intervention; it is part of it. The researcher is not separate from the research; rather, her presence is a key part of the process.

Burawoy thought these two ways of doing social science were so different that they could never be reconciled. I disagree, and I have spent much of my career trying to marry the twain that he argues could never meet. But he is right about one thing. The distance between researcher and researched is real, and in development economics it is stark.

Almost everyone I know who studies poverty has never been poor. We are often from different countries, different social classes, different castes and races, often speaking different languages, from the people we write about. None of that disqualifies you from working on poverty — it would be a strange profession that insisted only the poor could study poverty — but it should give us some humility. Most development economists do not know what the lives of poor people are actually like, day in and day out.

What the discipline rewards instead is a performance of objectivity: Researchers should not influence their subjects; they should stay at arm's length and use the same tools as everyone else. Your subjective suspicions should not contaminate your analysis. You are not a part of the subject group, and you should not be; your influence upon them would make your research less valid.

In short: the more distance we have from those we study, the more clearly we will see them. The opposite is closer to the truth. Distance does not produce clarity; it produces the kind of clarity you get from squinting through a small circle that you have cleaned inside a dirty window. It convinces you that what you can make out from where you are standing is all there is to see.

Four Things We Could Actually Do

So what would it take? I suggest four things economics (development or otherwise) could do to limit researchers' distance from their subjects.

Cognitive Empathy

The sociologist Mario Small uses the phrase "cognitive empathy" to mean the ability to understand a person's predicament as they understand it. It does not suffice to think about how you would feel if you were in their position; you must think about it as *they* understand it from theirs. It is much harder than it sounds. It requires you to take seriously the possibility that your subjects' theory of their own lives is more accurate than yours.

Some of the best development economists I know have this in spades. Jean Drèze has lived for decades in the rural India he writes about, refused funding from institutions like the World Bank that he thought would compromise his independence and successfully pushed for one of the largest rural employment guarantees in the world. Despite this, he almost never publishes qualitative analysis; his work is overwhelmingly quantitative. But every sentence is shaped by years of having actually listened. He is not alone among economists; Christopher Bliss and Nicholas Stern spent eight months in the village of Palanpur in the mid-1970s and inspired what is now a multigenerational research project.

The point is not that empathy must show up as a qualitative paragraph in the paper. Merely having done a focus group that you report in a footnote for color (and to signal your field creds) is not enough. Many economists already do this; it clearly isn't sufficient to reduce the distance between the researcher and the researched.

It is that the work has to be shaped by it. And here is the awkward truth: Almost all of contemporary development economics is not. We design experiments based on the input of other economists. We read papers, we deliberate in seminars, we come up with new models based on economic theory rather than experience. The intervention happens through an implementing partner, and the endline survey through a survey firm. Even the analysis might be handled by a research assistant (or, now, coding agent). The result is a body of work that is technically immaculate and substantively thin. To paraphrase an old chestnut: It is an expensive way to be precisely trivial rather than vaguely right.

Narratives Are Data

People do not talk in numbers. They talk in words. They tell stories, they contradict themselves, change the subject and circle back. They forget things and remember them halfway through the next topic; they get distracted and tell you about something interesting but unrelated. Survey instruments are a kind of violence against this — useful violence, often, but violence nonetheless. Instead of the messiness of human life, you get a small number of tick boxes. If the boxes are well designed, you get a lot of information. But you never get everything, and frequently, you will miss the most important parts.

The book "Portfolios of the Poor" is what every economist should be made to read on this point. The authors — a development economist (Jonathan Morduch), an anthropologist, a microfinance practitioner and a finance expert — gave up on standard surveys and instead visited 250 households in Bangladesh, South Africa and India at least twice a month for a year, building "financial diaries" out of long, open-ended conversations. They estimated that one-shot surveys were missing about half the financial activity of a poor household. *Half.* The poor were not, it turned out, simply consumption smoothing. They were trying to manage portfolios of assets under conditions of grotesque uncertainty, with very little room for error. No standard consumption survey would have shown the level of complexity in how the poor managed their households; researchers had to let households describe their finances in their own words.

I have done versions of this myself. With Paromita Sanyal, I spent 10 years analyzing transcripts of 300 village meetings in South India — what we called "oral democracy" — to understand whether poor, low-literacy, deeply unequal communities could actually deliberate in any meaningful sense. (Spoiler: They can, and the quality of that deliberation depends much more on state government policy than on the village's literacy rate.)

Ten years is a long time. It is also why so few economists do this kind of work. Tenure evaluation is often slower than that; investing in something that might not bear fruit until after tenure is a difficult choice for many early-career faculty. Worse yet, this kind of data does not always lead to the kind of publication that gets you tenure. My work was eventually published as a book, not a top-five journal article — a much lower-value thing when one is up for tenure.

And until very recently, the technology to scale up the analysis of narrative data simply did not exist. The 10 years were a function of the technology of the time, not a requirement of the method. With today's technology — recording devices, AI for transcription — collecting qualitative data is easier than ever. Embedding a few weeks of open-ended interviews in a standard RCT, reading the transcripts your survey firm's enumerators could be collecting anyway, piggybacking on existing qualitative data — these fit inside a dissertation timeline, and the tools for scaling them are getting cheaper every year.

The deeper point is the disciplinary reflex. Open-ended narrative is still routinely treated in economics as "anecdote" rather than data. Researchers who mix quantitative and qualitative methods are, to quote the political scientist Atul Kohli's wonderfully insightful joke, "stuck between a rock and a soft place." Reviewers reject them because of a perceived lack of rigor, and editors do not see the added value.

I have had this happen to me directly. The two early papers on domestic violence I mentioned earlier — one using a combination of ethnographic and econometric methods, the other building a game-theoretic model of dowry violence and testing it with survey data — were originally a single paper. When I included qualitative methods in the paper, it was rejected from several economics journals. My co-author Francis Bloch and I gave up and split it in two — one of the resulting papers was accepted into one of the most prestigious journals in economics.

This distaste for the qualitative is one of the stranger superstitions in the discipline; there is no methodological reason a transcript is less informative than a Likert scale.

Take Process Seriously

Empirical economics, especially since the credibility revolution, has become almost monomaniacally focused on outcomes. Did the intervention work? By how much? For whom? These are good questions, but they are not the only ones. It is rare for empirical economics papers to spend much time focusing on how an outcome happened. Who said what to whom to start the process of change? Who were the early adopters, and who had to be brought along later? What did people think about the intervention? An RCT can tell you if an intervention

worked; ethnography can tell you why. Mechanisms often get short shrift in empirical economics. You do your RCT, you get your result, you come up with some plausible explanations, you write up a model for how those explanations would work (if they're right). This is fine as far as it goes, but it does not go very far.

The map of mechanisms you can construct from theory alone is a small subset of the mechanisms that actually operate in the world, and reasoning from outcomes back to mechanisms is a notoriously unreliable exercise. The alternative is to actually go, talk to people and observe.

A few years ago, colleagues and I worked on a randomized trial in rural Karnataka to test whether intensive training in participatory planning would improve village governance. The intervention assigned 50 villages to the participatory training, planning and monitoring exercise and 50 to control. Instead of just relying on a baseline and endline survey (which we also did), we did the unusual thing of embedding five trained ethnographers in matched treatment-control pairs of villages for the duration of the study. They produced 240 monthly reports over four years.

Participatory planning in Raichur District, Karnataka © Vijayendra Rao

The headline result was a null. The intervention did not produce statistically significant gains over the comparison villages. In the standard economics genre, that would have been the end of the story — a disappointing null. Perhaps the paper would include a theoretical model on why participation fails, but there would be very little to learn here.

But because we had the ethnographies, we could see what actually happened: The “failure” was not a failure of the idea but of the conditions under which it was tested. The quality of implementation varied a lot: Some facilitators were excellent and some were poor. Higher officials in the government did not take the intervention seriously and frequently transferred and replaced facilitators for reasons that had nothing to do with the intervention. Persistent caste-based inequality was actively chewing through any gains the training produced. The ethnography was the core of the paper, not just the color.

Kripa Ananthpur conducting an interview about the intervention. © Vijayendra Rao

Respondents as Analysts

If our purpose as researchers is to help people become better off, then the people themselves should at minimum be told what we found. Development practitioners have been talking about

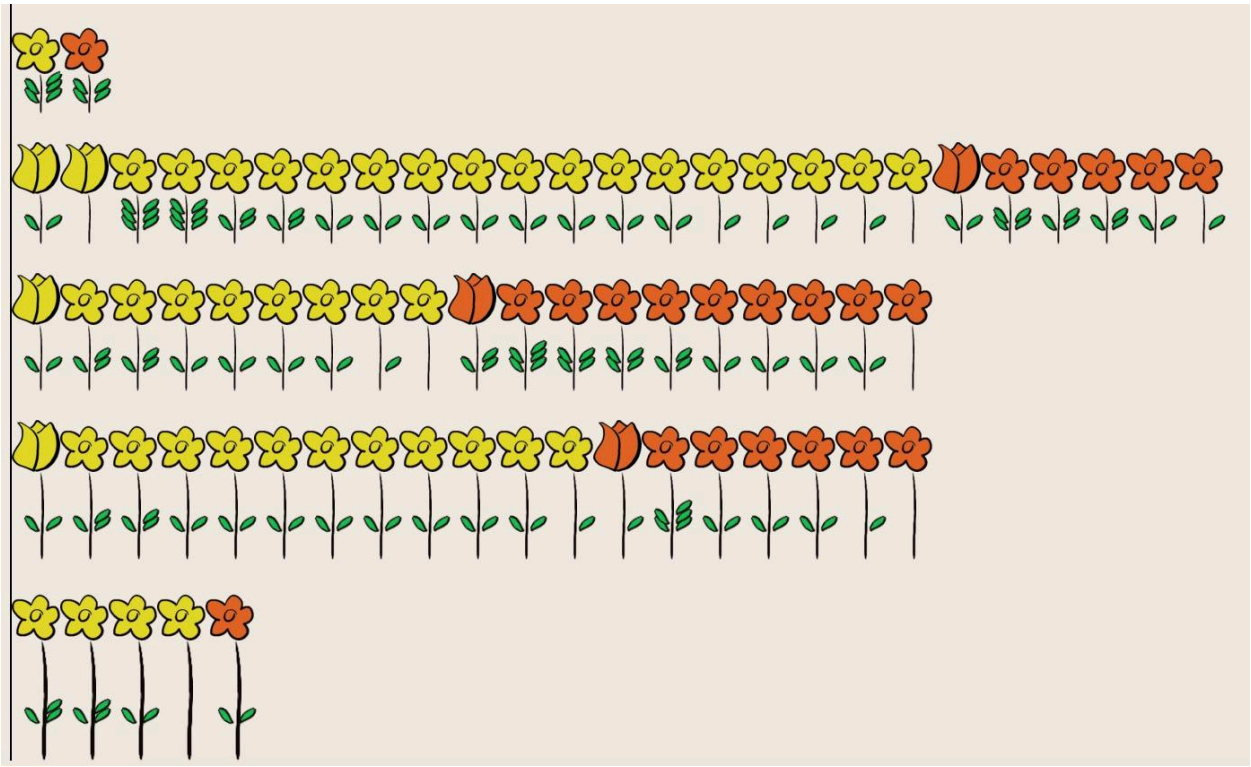
including those they study in the process for decades — and yet, this is rarely implemented. And reporting findings would only be step one. It would be better if they helped design what we ask. Best of all, they should be able to conduct research on their own lives without our mediation at all.

In 2014, with a group of colleagues at the World Bank, I tried this. We worked with representatives of more than 200 tribal villages in South India to co-produce a method we called participatory tracking. The villagers spent weeks deciding for themselves what counted as the good life, turned those ideas into survey questions, tested the questions in their own villages and then — using tablets and a system of video-based training — surveyed their own neighbors. In a single district, we conducted a census of 32,000 households in about six weeks.

Only 17% of the questions overlapped with the standard Indian National Sample Survey. The villagers were asking different things because they wanted to know different things. For instance, in order to assess whether a household was poor, they devised the following question, which proved extremely effective: “Did the last person to eat in the family ever go hungry in the past week?” To a respondent this was obviously a question directed at the mother, and mothers in food-constrained households often went hungry in order to feed the adult men and children in the family.

Literacy rates were low, so we collaborated with the villagers to produce data visualizations of the results. We iterated until people who could not read could see at a glance how their village was doing on what they cared about. The data then got used in actual village meetings. The quality of those meetings improved noticeably, because everyone was working from the same picture instead of arguing about the facts.

Instead of the usual LaTeX table, here is a visualization of marriage patterns that we co-produced.



Each woman in the village is represented as a flower. The height of the flower is the age at which the woman got married. The number of leaves is the number of children resulting from the marriage. A red flower indicated that the woman had married a blood relative, and a yellow one meant that she had not. A flower bud meant that the marriage was not consensual, and a bloomed flower was consensual. The mappings between the flower's appearance and meaning were tweaked by the women to make them more relevant to their local contexts, and therefore more intuitive.

Some villagers who saw this visualization immediately disregarded the high proportion of red flowers they saw, as marriage within families is accepted and widely practiced. However, other communities wanted to reduce the ratio of red flowers to yellow flowers by educating future generations of women about the problems related to intra-family marriage.

Women in Rural Tamil Nadu discussing the visualizations © Vijayendra Rao

Trust me, I know how this sounds. The rest of the field would file this experiment under “charming but unscalable.” That is a little bit true; a participatory, co-produced survey process is slower and more laborious than a standard questionnaire. We could not use a standard set of questions, because the standard set of questions wasn’t what people actually wanted to know. And the alternative was the status quo: where we would show up, ask people a bunch of questions about their lives — that they really didn’t feel were all that relevant to them anyway — and write a paper that no one surveyed would ever read. Given that, co-produced surveys seem worth the time.

The LLM Problem

The technology for coding open-ended interview questions has also gotten much better in the last decade. Large language models are particularly adept at going through large amounts of

unstructured text and pulling out themes; Claude's current abilities would have seemed like science fiction in 2014. Now, they can essentially replace a research assistant in coding English-language text. I have used LLMs too. Some of my recent work piggybacks on a panel of Rohingya refugees and Bangladeshi hosts where we conducted long open-ended interviews with 2,000 respondents and developed a method we call iQual. It analyzed a small subsample using interpretative sociological qualitative coding and then used machine learning to scale up the codes to the full sample. This kind of project was simply not feasible before.

But using LLMs outside of Western contexts isn't always simple. We compared our "bespoke" method to LLM-based coding and found that LLMs gave us highly biased results — possibly because they are not trained on Bengali and Rohingya dialect text. The bias of a poorly designed survey is at least legible. The bias of a frontier language model trained on the internet is not. Hopefully this is a temporary problem; with efforts in place to improve AI with under-resourced languages, LLMs will get better at this with time.

So I should be excited, and on most days I am. But I want to be careful here. It's true that there is a version of the future in which LLMs democratize narrative analysis in the same way that calculators did for arithmetic. They are faster and cheaper and they can vastly expand who can do qualitative analysis. Unfortunately, however, I fear it is more likely that they will be used to widen the distance between researcher and respondent. If a researcher never spends time with the people they study and simply reads the LLM's summary, there is no cognitive empathy. The LLM is the only one listening — not the researcher. My rule of thumb is that LLMs cannot substitute for a human. They can extend your reach — help you process data more quickly — but they cannot replace you. You still need to spend time in the field; you still need to talk to people about their lives and their needs. You must read the transcripts yourself; you should know the people and the context well enough to be able to tell if the LLM's coding has gone awry. LLMs might be powerful enough to do 90% of the job — but that remaining 10% should not be automated away. Empathy is not a Claude skill.

What It Would Actually Take To Listen to Respondents

Some of this can be accomplished by individual researchers deciding to do their work differently. Quite a lot more of it requires the discipline itself to change. Right now, the professional incentives still push young scholars towards "business as usual."

Journal editors have to stop reflexively rejecting qualitative material. Graduate programs will have to expand beyond teaching just econometrics into teaching qualitative methods as well. Hiring committees will have to start to value time spent in the field. Funders have to accept that good mixed-methods work can be expensive and slow — certainly slower than running a regression from a Cambridge office. But it will also make economics a stronger discipline. Spending time with people will produce insights that can be obtained no other way.

In the short run, I am not optimistic. Disciplines are stubborn things, and economics is more stubborn than most. It is a discipline already struggling to diversify beyond a few top programs and economics seems to be particularly prone to reproducing class hierarchies. But I am less pessimistic than I might be, because it seems things are already beginning to change.

A generation of younger development economists is more comfortable doing fieldwork than mine was. Adjacent disciplines — political science, sociology, public health — have been mixing qualitative and quantitative methods for a long time without losing their disciplinary identity. Some of these disciplines even co-author with economists, adding depth to the rigor of an economics paper. And some funders are also starting, slowly, to ask harder questions about whose voices are represented in the work they pay for.

My argument is, fundamentally, not that complicated. If you study people, you should listen to them. This is doubly true if you study people who live lives that are very different than your own. It is on you, as the researcher, to try to close the distance between you and your subjects — building bridges instead of walls.

Thirty years ago, Robert Chambers asked an important question: "Whose reality counts?" It has still not been answered. Economics has preferred to duck the question, pretending that empirics could substitute for an answer. I do not think it should continue to do so. Indeed, I think this remains one of the most important open questions in development research.

The woman in Karnataka dragged out of our focus group by her hair was telling us something — that her reality was not captured by our questions. It took us a week to be able to hear it, and the only reason we eventually did was because we were still in the village a week later. There is no shortcut for this.

Numbers help. So do words. But most important is sitting in the village and listening long enough, and with enough cognitive empathy, that someone tells you the truth.

Vijayendra Rao is a development economist and social scientist who combines econometric methods with ethnography. He spent 27 years as a Lead Economist in the Development Economics Research Group at the World Bank, and is currently the Roberta Buffett Distinguished Scholar-Practitioner in Residence at Northwestern University. His interests include gender, culture, participation, deliberative democracy, political economy, and innovations in mixed methods. His recent methodological work with Julian Ashwin, Monica Biradavolu, Aditya Chhabra and others uses natural language processing to analyze open-ended interviews at scale and compares it to LLM based qualitative analysis. His latest book, Revolution by Stealth: How Women's Groups Catalyzed a Cultural Transformation in Bihar (with Shruti Majumdar and Paromita Sanyal), is forthcoming from Cambridge University Press in September 2026.

If you have comments on this article, or wish to contribute to the discussion, please email them to letters@indevelopmentmag.com. Responses will be featured in a letters section.