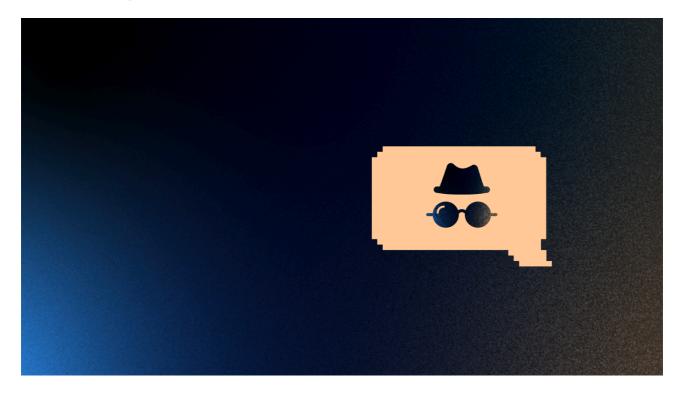# Chatbots in Disguise

How Participants Use GenAI to Hack Qualitative Research

When researchers Ranjit Singh, Livia Garofalo, Briana Vecchione, and Emnet Tafesse set out to interview people about their use of mental health chatbots, they encountered a particular kind of AI-enabled deception.

June 11, 2025

At first, everything seemed routine. Our research team at Data & Society set out to study how US-based individuals use chatbots — tools like ChatGPT, Replika, and Woebot — to manage their mental health and emotional well-being. We recruited participants and compensated them for their time. But from the very first interviews, something felt off.

Participants confidently described extensive (and expensive) use of multiple chatbot services, casually tossing around psychological buzzwords like "coping skills" and "CBT." Their answers seemed plausible — carefully phrased and jargon-laden — yet strangely vague. We found ourselves stuck in tautological loops: "Do you use chatbots for emotional support?" we asked. The reply: "Yes, all the time." We followed up: "How exactly do you use them?" "Well, for my mental health," was the answer.

Subtle inconsistencies started to add up: locations in the US that didn't match the time of day visible on-screen, participants who were consistently off-camera,

incongruous demographic details. Slowly, we came to an unsettling realization. It seemed likely that the people we were interviewing about chatbots *didn't* actually use them regularly — and that they might not even be located in the US.

So, we shifted tactics. We probed for more details; we framed the same question differently a couple of times to see if the answers changed. Then in one case, after nearly 25 minutes of careful questioning, gentle probing, and attempts to encourage more specificity, we paused. With curiosity rather than accusation, we openly acknowledged the mismatch we perceived: "It would be great if you could be honest with us. You're going to get your payment either way." After a brief pause came an unexpected breakthrough. With evident relief, our participant confessed: "Okay, I don't need to lie. To tell you the truth, I don't use these tools very much [for mental health]," he said.

This confession transformed the tenor and scope of the interview. Our conversation shifted dramatically, from a scripted performance of a participant trying to sound "legitimate" to a candid exploration of how generative AI (genAI) facilitates new forms of gig work and creative participation in qualitative research studies. Our participant explained how individuals like him routinely use ChatGPT to anticipate likely interview questions based on recruitment descriptions and rehearse responses in the hopes of qualifying for higher-paid research studies, effectively gaming the research process itself:

"It is very simple, what do you do first? You read the idea of your study. What is the story [that the researchers] are looking for? [··· Then,] you can talk to ChatGPT. [···] So you say: 'I'm having an interview with [researchers] actually looking for people like this. So I want you to act like one of them.' Then you get questions [and you use these questions to get answers···] You review that, and you read through it, and you get [ready. Some···] people have [this Q&A] right next to their screen and start reading through word by word verbatim."

Far from being angry or frustrated, we were fascinated. We had finally figured it out: we were not hearing lived experience — we were hearing a ChatGPT "script," prepped to anticipate our questions and deliver just-plausible-enough answers in the voice of someone it imagined we wanted to talk to. Our pivot sparked a moment of refreshing honesty and opened up a lesser-known use of genAI: individuals strategically leveraging genAI to become believable research participants.

Gig Work, GenAI, and New Economies of Participation

Our situation may seem like textbook "scamming": individuals, most commonly crowdworkers, leveraging AI to convincingly fake expertise in qualitative research interviews. Yet it opens a space for a much richer and complicated story. On the one hand, participants were not maliciously "scamming" us in the classic sense — to us,

it seemed that they were making strategic use of social media to find opportunities to participate in research studies that pay substantially more than their typical gig labor marketplaces. These interviews were not just scams, they were survival strategies — ones enabled and shaped by the evolving digital landscape of gig labor. On the other hand, such crowdworkers were essentially mediating between researchers and ChatGPT's simulated participants, enacting what Shivani Kapania and her collaborators call "surrogate effect," whereby AI-generated responses stand in for authentic human experiences. This surrogate mediation complicates traditional notions of qualitative inquiry, raising questions about whose voices are ultimately represented in research findings.

These complexities resonate deeply with Data & Society's recent publication *ScamGPT: GenAI and the Automation of Fraud*, authored by Lana Swartz, Alice E. Marwick, and Kate Larson, which highlights the fact that genAI doesn't just facilitate traditional fraud — it complicates distinctions between authentic and deceptive behavior, reshaping social interactions and economic relations in unexpected ways. AI-enhanced scams often thrive in contexts of economic uncertainty and precarity. As evident in our case, these scams blur the lines between legitimate gig work and deception, illustrating how genAI is reshaping participation in qualitative research studies in particular, and digital labor markets more broadly.

Participant deception predates genAI; anthropologists have long grappled with the layered realities of informant dishonesty. Steven Nachman's seminal article "Lies My Informants Told Me" opens with the observation that lying is a universal human behavior, yet one that remains curiously underexplored in ethnographic research. He further complicates the picture by noting that "the most accomplished liars in the community are also sometimes the most accomplished truth tellers." Building on this, Jan Beek and colleagues, in "Mapping Out an Anthropology of Defrauding and Faking," argue that acts often labeled as fraudulent or deceptive are better understood as contextually situated performances — responses to economic constraints, institutional opacities, or shifting moral terrains. These are not simply lies, but strategies.

In the age of genAI, the fakes come with cleaner grammar and a keener instinct for what researchers want to hear. What might once have been a muddled story is now fluently rendered by ChatGPT, tailored to match the expectations baked into our study designs. The vagueness is still there, but softened — less a red flag, more a shimmer — requiring a keener, more reflexive ear to detect. These moments ask us to rethink deception not as failure, but as a signal: a trace of the pressures, hustles, and hopes that shape how people show up in research. The question is no longer just what is true, but what these stories, however calculated, tell us about the shifting socioeconomic landscapes our participants are navigating.

Since that honest conversation with our participant, we have been contending more explicitly with new ethical and methodological questions:

- **Ethical reconsent:** If a participant reveals deception mid-interview, should we pause and ask for re-consent, explicitly related to new research on how crowdworkers leverage genAI for their own survival and creativity?
- **Epistemological reflection:** Can "fake" responses provide meaningful insights into precarity, gig work, and the role of AI in the everyday lives of crowdworkers?
- **Parallel studies:** Should we explicitly develop parallel research tracks — one on genuine chatbot use for mental health, and another on how genAI is reshaping gig economies, deception, and qualitative research?

In taking these questions seriously, we hope to move beyond simplistic narratives of deception and fraud to explore: What counts as authenticity when AI helps perform it? What truths emerge when deception is also a survival strategy? What passes as data now demands a deeper reckoning with the blurred lines between sincerity, hustle, and simulation. In this shifting terrain, our task is to listen more closely to the performances themselves — what they reveal, what they obscure, and what they make possible. For qualitative research, the challenge is no longer just how to collect authentic stories, but how to make sense of truths that arrive mediated, curated, and sometimes co-written by machines.